

*To Alexius Meinong and Bertrand Russell*



# Contents

<i>Acknowledgments</i>	5
FRANCK LIHOREAU	
Introduction	7
TERENCE PARSONS	
Fictional Characters and Indeterminate Identity	27
BRENDAN MURDAY	
Two-Dimensionalism and Fictional Names	43
ERICH RAST	
Classical Possibilism and Fictional Objects	77
ALBERTO VOLTOLINI	
How Creationism Supports Kripke's Vichianism on Fiction	93
GRAHAM PRIEST	
Creating Non-Existents	107
FRED ADAMS	
Sweet Nothings: The Semantics, Pragmatics, and Ontology of Fiction	119
R. M. SAINSBURY	
Fiction and Acceptance-Relative Truth, Belief and Assertion	137
ROBERT HOWELL	
Fictional Realism and Its Discontents	153

FREDERICK KROON	
The Fiction of Creationism	203
GRANT TAVINOR	
Virtual Worlds and Interactive Fictions	223
MANUEL REBUSCHI & MARION RENAULD	
Fiction, Indispensability and Truths	245
<i>References</i>	287
<i>Notes on Contributors</i>	301
<i>Index</i>	305

# Acknowledgments

My work on the book project was carried out at the Philosophy of Language Institute (IFL), New University of Lisbon. It was supported by a research fellowship from the Portuguese Science and Technology Foundation (FCT), and developed within the FCT-funded IFL project “Context & Communication” (PTDC/FIL/68643/2006). In these respects, I am particularly indebted to António Marques and João Sàágua.

I have received valuable advice on essay selection, comments, suggestions and encouragement from a number of friends and colleagues, including Luiz Carlos Baptista, Thomas Behrens, Éric Cléménçon, Gabriele De Angelis, Rijn Dutoit, Manuel García-Carpintero, Erich Rast, Manuel Rebuschi, and the authors of the essays. In addition, I truly appreciate the careful reading and extremely valuable suggestions by Mick Sweeney on certain parts of the book.

I am grateful to Ontos Verlag for publishing this volume. Rafael Hüntelmann has been an understanding and astute editor, and the Ontos Verlag staff has shown professionalism and efficiency from the outset. I thank them all for making my work on the project a worthwhile and enjoyable experience.

And finally, above all, and again, I want all the authors of the essays to know how grateful I am that they so generously and enthusiastically accepted to be part of this project.



# Introduction

FRANCK LIHOREAU

*A literatura, como toda a arte, é uma confissão de que a vida não basta.*

Fernando Pessoa

The eleven essays collected in the present volume are all partially, if not entirely, concerned with the connection between fiction and truth. This question is of utmost importance to metaphysics, philosophy of language, philosophical logic, and epistemology, as it raises in each of these areas and at their intersections a large number of issues related to creation, existence, reference, identity, modality, belief, assertion, imagination, pretense, etc. All these topics, and many more, are addressed in this collection, which brings together original essays written from various points of view by philosophers of diverse trends. These essays constitute major contributions to the current debates that the question of the connection between fiction and truth continually enlivens, and give a sense of the directions in which research on the question is heading. In this Introduction I give a synoptic description of the essays, highlighting how the ideas they discuss find their natural place in those debates. (In doing so, I shall rely on definitions of, and distinctions between various “labels” whose quasi-stipulative character is meant for the sake of clarity and relevance rather than exhaustivity. For surveys of recent relevant work, one may want to consult Thomasson, 1999; Woods & Alward, 2002; Everett, 2005; Friend, 2007; and Fontaine & Rahman, 2010.)

We sometimes praise people for telling stories that we commonly know to be true. We often blame people for telling stories that we commonly know

to be false. But we hardly ever praise or blame people for telling stories that we commonly know to be fictive. This observation makes particular sense when we consider excerpts from literary works of fiction, like:

- (1) Holmes, who was usually very late in the mornings, save upon those not infrequent occasions when he was up all night, was seated at the breakfast table. (*The Hound of the Baskervilles*)
- (2) When they left Tostes at the month of March, Madame Bovary was pregnant. (*Madame Bovary*)
- (3) On meeting Anna Karenina, as he was Alexey Alexandrovitch's enemy in the government, he tried, like a shrewd man and a man of the world, to be particularly cordial with her, the wife of his enemy. (*Anna Karenina*)

Unless we are ignorant of the fact that the stories told in the works are “made up” and intended by their author to be thus read, this sort of discourse—which Currie (1990) simply calls “fictive” and which Bonomi (2008) associates with what he calls “textual” uses of fictional discourse—will not naturally be assessed in terms of its truth or falsity. The link between truth and fiction becomes manifest, however, in other sorts of fiction-involving talk. One is the sort of talk that Currie calls “metafictive” and Bonomi “paratextual”, exemplified by such sentences as:

- (4) Sherlock Holmes was a detective.
- (5) Sherlock Holmes was clever.
- (6) Sherlock Holmes lived at 221B Baker Street in London.
- (7) Watson lived at the same place as Holmes.
- (8) Anna Karenina was a woman.

These claims, which are somehow “internal” to the relevant fictions, are intuitively assessable as true or false. They are intuitively true in the fiction and given the content of the fiction. Another sort of fiction-involving discourse corresponds to Bonomi's notion of a “metatextual” sentence, a notion that overlaps Currie's notion of “transfictive” discourse and exemplified by:



- (9) Holmes is more clever than Poirot.
- (10) Anna Karenina and Raskolnikov were both Russian.
- (11) Sherlock Holmes is more clever than any real-life detective.
- (12) Anna Karenina and Tolstoy were both Russian.
- (13) Professor James Moriarty made his first appearance in *The Final Problem*.
- (14) When authors create fictional characters, they describe them with more or less physical details, but some characters in 19th-century novels are described with a greater wealth of physical details than is any character in any 18th-century novel.

This time, the claims are somehow “external” to the relevant fictions, in that they do not bear so much on the contents of the works of fiction as on the works of fiction themselves, and do not so much say something true of the worlds of the fictions as of the real world. So, at least as far as paratextual and metatextual claims are concerned, it makes intuitive sense to think of them as being true or false. But how could such claims ever be true (or false) if, as everyone will agree, it is also true that:

- (15) Sherlock Holmes does not really exist.
- (16) James Moriarty does not really exist.
- (17) Emma Bovary does not really exist.
- (18) Anna Karenina does not really exist.

and so on, and if, more generally, being fictional seems to be incompatible with being real?

The literature on the matter can be divided into two main camps: the “realists” and the “irrealists”. *Realists* hold that there are such things as fictional entities, and maintain that one can refer to, quantify over, attribute and deny properties to them. For some realists, those entities precede human activity; for others not. In any case, since fictional entities do not really exist, realists must explain how we can apparently make true claims *about things that do not exist*. *Irrealists* deny that there are such things as

fictional entities, hence deny that one can ever refer to, quantify over, or attribute and deny properties to such entities. Since on their view there are no fictional entities, irrealists must explain how we can apparently make true claims *about nothing at all*.

The first five essays—by Parsons, Murday, Rast, Voltolini, and Priest—all assume the truth of some form of realism. Realism can be understood as saying that objecthood exceeds actual existence, that is to say, the things that actually exist form only a proper subset of the things that are. This idea can be developed in various ways. In this respect, *neo-Meinongians* like Parsons (1980), Routley (1980) and Priest (2005) *inter alia*, offer what is perhaps the most straightforward and spectacular form of realism. According to them, some objects are existent, some are nonexistent, and a fictional character like Sherlock Holmes is a nonexistent object. A negative existential claim like that in (15) is thus understood as being about an object, that referred to by “Sherlock Holmes”, and as simply denying, truly, the existence of that object. To introduce nonexistent objects, realists of this kind sometimes rely on a distinction between “nuclear” properties (like *being golden*, *being a mountain*, etc.) and “extranuclear” properties (like *being existent*, *being impossible*, *being admired by Christie*, etc.), to formulate a “Generating Principle” like this one:

For any class  $C$  of nuclear properties, there is an object  $o$  such that for any property  $P$ ,  $o$  has  $P$  iff  $P$  is in  $C$ ,

thus following a strategy that Terence Parsons endorsed in his book *Nonexistent Objects*. A principle like this one guarantees that what objects there are will include such nonexistent objects as the golden mountain, viz. that object that has exactly the nuclear properties making up the class  $\{\textit{goldenness}, \textit{mountainhood}\}$ . The result carries over naturally to the case of nonexistent fictional objects like Sherlock Holmes, *modulo* the restriction of the relevant class of properties to those that can be assigned to that character *based on what is said in the relevant stories*. We can encompass this restriction by saying that a fictional character is that object that has all and only those properties attributed to the character in the story in which it originates. Then, such claims as “Sherlock Holmes is a detective” or “Sherlock Holmes lives at 221B Baker Street” count as true because being a detective and living at 221B Baker Street are indeed prop-

erties that the character in question is understood to have in the Sherlock Holmes stories. (It should be noted, however, that Meinongians need not follow the general strategy just outlined; in particular, they do not have to accept a distinction between characterizing and non-characterizing properties. Priest, for instance, dispenses with this idea altogether in his book *Towards Non-Being*.)

Fictional characters can have many potential originating stories, and these stories will eventually disagree on whether those characters have certain properties. For instance, there seems to be no unique potential source for Sherlock Holmes, as we may well speak of “the Holmes of the Conan Doyle series” as opposed to “the Holmes of the TV series”, or of “the Holmes from the early stories in the series” as opposed to “the Holmes from the later stories in the series”, and as the relevant series or stories may well disagree on some nuclear properties of their main protagonist, for instance, on whether Holmes is bald or not. Should we then conclude that there can be more than one object for a fictional character, say, a bald, and a non-bald Sherlock Holmes? In his essay “Fictional Characters and Indeterminate Identity”, **Terence PARSONS** abandons a previous suggestion from *Nonexistent Objects*—that there are many Sherlock Holmes and that the name “Sherlock Holmes” is ambiguous between the many Sherlock Holmes there are—, and explores the option of maintaining that there is at least and at most one Sherlock Holmes, even though there is no answer to the question whether he is bald or non-bald. To this effect, building on some considerations from his book *Indeterminate Identity*, he avails himself of the notion of *worldly indeterminacy*, the idea that there is genuine indeterminacy of identity in the world, to be precise, indeterminacy in states of affairs. Roughly, there are states of affairs; some are determinate; some are indeterminate. Objects can enter indeterminate states of affairs. This calls for an amendment of the above Generating Principle along the following lines:

For any two definitely disjoint classes  $C_1, C_2$  of (nuclear) properties, there is an object  $o$  such that (i) for any property  $P$ ,  $o$  determinately has  $P$  iff  $P$  is determinately in  $C_1$ , and (ii) for any property  $P$ , it is indeterminate whether  $o$  has  $P$  iff  $P$  is determinately in  $C_2$ .

Then, the answer to the question whether an object  $o$  has a property  $P$  can be indeterminate if the very state of affairs of the object in question having the property in question is itself indeterminate. Now, fictional objects can enter indeterminate states of affairs too. This is what happens in the case of the disagreement concerning Holmes' hair. The state of affairs of his being bald is indeterminate. The question of whether he is bald—as well as the question of whether he is non-bald—thereby has no answer.

It is the neo-Meinongians' contention that paratextual claims like (4) to (8) have full propositional contents and are truth-evaluable as they are. It might be thought, however, that instead of a sentence like “a famous detective lived at 221B Baker street in London”, it would be somehow more appropriate to say “According to the Sherlock Holmes stories, a famous detective lived at 221B Baker street in London”. This kind of thought motivates the underlying idea of the “fictional operator” approach to fictional truth, whose best-known exposition is found in Lewis (1978):

Let us not take our descriptions of fictional characters at face value, but instead let us regard them as abbreviations for longer sentences beginning with an operator “In such and such fiction ...”. Such a phrase is an intensional operator that may be prefixed to a sentence  $\phi$  to form a new sentence. But then the prefixed operator may be dropped by way of abbreviation, leaving us with what sounds like the original sentence  $\phi$  but differs from it in sense. (pp. 262–263)

This idea has been endorsed by authors who espouse realist views, like Lewis himself and Howell (1979), as well as by authors who espouse other views, like Currie (1990). Lewis' own proposal was to construe fictional operators as modal operators and to give them a standard possible worlds interpretation: basically, “According to fiction  $f$ ,  $\phi$ ” is true exactly if  $\phi$  is true at all worlds that are relevantly determined by  $f$ , e.g. all worlds “where  $f$  is told as known fact rather than fiction” (Lewis, 1978, p. 268). The resulting picture seems clear enough, but gets more complicated as soon as the issue of the meaning of names in fiction is brought onto the table. According to the *direct reference* view of proper names—also referred to as “Millianism”—, the meaning of a name is totally exhausted by the object it denotes, in other words, its bearer. According to the *descriptivist* view, by contrast, there is more to the meaning of a name than simply

its bearer, and that meaning is best analyzed as being equivalent to that of an appropriate (definite) description. Now, a question that might arise is whether any one of these two views can successfully be extended to proper names in fiction and combined with a fictional operator approach within a realist framework. This question is addressed in the next two essays in the volume, albeit in quite different ways.

An obvious difficulty for a direct reference theorist who endorses a fictional operator account is this. Since a name like “Holmes” has no bearer, a sentence like (5) expresses an incomplete, partially empty propositional content—on a structured proposition view:  $\langle \text{____}, \textit{being clever} \rangle$ —rather than a complete singular proposition. Then, Lewis would say, “we may abandon it to the common fate of subject-predicate sentences with denotationless subject terms: automatic falsity or lack of truth value, according to taste” (Lewis, 1978, p. 263). The problem, however, is that the true prefixed sentence “According to the Holmes stories, Holmes is clever” certainly does not express the content:  $\langle \text{According to the Holmes stories}, \langle \text{____}, \textit{being clever} \rangle \rangle$ , for after all, the Holmes stories do not attribute cleverness to nothingness. As a consequence, either the direct reference view is incompatible with a fictional operator approach, or else such operators have a semantic behavior that differs significantly from that of modal operators of the usual kind. In his essay “Two-Dimensionalism and Fictional Names” **Brendan MURDAY** explores the second option. According to him, all names are directly referential as long as they do not occur within the scope of fictional operators. This idea is implemented by saying that names have referential contents determined by “reference-fixers”—which *need not* be thought of as definite descriptions—that are included in the propositions expressed by sentences in which the names occur. Fictional objects do not actually exist, so nothing in the actual world satisfies the reference-fixer for fictional character’s names, like “Holmes” when it occurs in an actual world utterance of the simple, unprefix sentence “Holmes is clever”. The name is then empty. Yet, if the fiction turned out to be true, something would satisfy the reference-fixer for that name, which would thereby be filled, and cleverness would be truly attributed to that thing. This is what a complex sentence like “According to the fiction, Holmes is clever” tells us, which can be interpreted as a complex instruction of the form: “Go to the world of the fiction, find out what individual *d* satisfies the reference-fixer of ‘Holmes’ and what property *P*

satisfies the reference-fixer of ‘is clever’ at that world, then check whether the proposition  $\langle d, P \rangle$  turns out true at that same world”. Fictional operators thus behave differently from usual modal operators, as they require that a world distinct from the evaluation world be considered as actual. To be precise, for Murday, “According to the fiction” is equivalent with “At the world of the fiction considered as actual”. To spell out this last idea, Murday relies on a two-dimensional semantic framework inspired by Kaplan (1989), which has expressions evaluated with respect to two parameters instead of just the usual single world-parameter. Fictional prefixes can then be defined as two-dimensional modal operators that shift the values of both parameters by designating the world of the fiction as actual and having the unprefixes sentences to which they apply evaluated at that same world. (Predelli (2008) also gives a two-dimensionalist treatment of fictional operators, though for the distinct purpose of accounting for their interaction with indexical modal and temporal operators.) An advantage of this view, Murday claims, is that it allows for a version of *actualism*—the view that everything there is is actual—that leaves room for things that are merely possible, things that are not but could have been, had another world turned out to be actual.

Another alternative, though, to the actualist claim that nothing there is is non-actual simply is to endorse its straight denial, that is, to endorse “possibilism”. According to *possibilism* in its classical form, the things that have actuality or existence form only a proper subset of the things that are *tout court*. What is actual exists and what is non-actual and only possible does not exist, yet could exist. This is another way of saying that objects that do not actually exist may still exist in various other ways. Being a fictional object is one of these ways. This is the stance adopted by **Erich RAST** who, in his essay “Classical Possibilism and Fictional Objects”, sets out to combine fictional operators with an existence predicate and an actuality operator in order to account for the semantic contribution of fictional names, in a descriptivist setting this time. If we go for a definite description analysis of fictional proper names, we can indeed make it part of the meaning of the name “Sherlock Holmes” that there is a unique individual in that domain that actually goes by that name, is nonexistent, yet exists according to the Holmes stories. A simple name-predicate sentence like (4) will then merely express that that unique individual has the property of being a detective—in symbols:  $\iota x@[Sx \wedge \neg Ex \wedge \Box_f Ex]Dx$ —,

without thereby entailing its actual existence. By combining this analysis of fictional names with a non-standard theory of predication that admits of two kinds of negation, “inner” (predicate) negation, and standard truth-functional negation, it becomes possible to express, of an object, that it is both not the case that it has a property and not the case that it lacks that property—in symbols:  $\sim Pa \wedge \sim \neg Pa$ —, and possible for this to come out true in some cases. The resulting treatment allows Rast to predict what he takes to be the expected truth-values of various paratextual as well as metatextual statements. Rast also explores how his treatment can be extended so as to account for other ways, besides the fictional, in which non-actual objects can exist (in belief, in the future, etc.), and how it might accommodate the replacement of the initial definite description theory by a direct reference theory of names.

The variants of realism mentioned so far are all compatible with the claim that fictional characters can be concrete individuals and that, for instance, Holmes is a flesh-and-bones, albeit non actually existing, detective. They are also compatible with the claim that the fictional depends on the real to some extent, that is, in so far as it means that *what properties* a fictional object has depends on the behaviour of the people who tell or listen to the stories about that object in the real world; *but not* if it means that *what fictional objects* there are depends on the behaviour of the story-tellers and their audience. The domain of objects available for reference and quantification in the real world is given “in advance”, or so to say, regardless of what agents in the real world do. This is likely to sound too radical a consequence to many, specially to those who find it natural to say that Arthur Conan Doyle “created” Sherlock Holmes, and to think that had Doyle decided otherwise, Holmes would not have been a detective, or that if Doyle had not written his detective stories, there would have been no Holmes”. The naturalness and pervasiveness of this metatextual way of talking and thinking about fictional characters might well be the main reason for the current popularity of so-called “creationist” approaches to fiction. *Creationists* (also referred to as “Artefactualists”), like Thomasson (1999) and Voltolini (2006), are moderate realists who take the above kind of talk and thoughts seriously, not merely metaphorically, and accordingly profess that *ficta* must be assigned full ontological status as abstract artefacts created by authors of fictions. They thus defend the idea that fictions and fictional characters are creations, and like any other creation, they de-

pend on the real in several respects. For Amie Thomasson, for instance, a fictional character is an abstract cultural object that has a dependent existence, as it is brought into existence by the act of an author writing a work of fiction, and remains in existence by the continuing existence of copies of that work.

The creationist's best selling point is, as we just mentioned, that compared with more radical forms of realism, it offers a rationale that straightforwardly explains our ordinary and pervasive uses of "creationist" or "creation-connoting" locutions. In his essay "How Creationism supports Kripke's Vichianism on Fiction", **Alberto VOLTOLINI** insists on a further merit of creationism: rightly understood, it vindicates the view that true claims about fictional characters are true because they are stipulated to be so (and not because those claims truly attribute certain properties to real individuals at the end of a referential chain associated with the characters' names, as a causal theory of reference applied to names in fiction would have it). The *prima facie* plausibility of this stipulationist view, or "Vichianism" as Voltolini calls it, is threatened by the possibility of discovering one day that the use of what we previously thought to be the name of a purely made-up character turns out to be causally grounded in a referential chain that does end in a real individual. According to Voltolini, the threat can be removed by endorsing a version of fictional creationism that allows for a distinction between two kinds of use for names in fiction: a *pretending* use of the names whereby one refers to nothing but simply pretends to refer to something; and a *characterizing* use of the names whereby one does refer to something, an abstract created entity, by characterizing it through features that are "internal" to the relevant fiction. On the pretending use, the sentence "Sherlock Holmes is a detective" is true because it is true with respect to the corresponding "make-believe world" where things are just like Doyle pretends that they are. By contrast, on a characterizing use, the same sentence tells us that the character Sherlock Holmes is a detective character, and this is true only because Doyle once decided to make that character a character of that kind. By endorsing a version of creationism along these lines, hence a moderate form of realism, the above threat to the view that true fictional claims are true by stipulation can be alleviated, by noticing that indeed there is *a priori* no guarantee that such claims, in their pretending use, involve names whose reference



is not causally grounded in real individuals, yet that the view automatically holds of the same claims when made in their characterizing use.

Although creationism might appear to many to be the most plausible form of realism for the reason that it gives credit to our ordinary and widespread uses of “creationist locutions”, even radical realists might agree that in some sense at least, authors *create* their characters. Realists of a Meinongian persuasion would certainly not paraphrase this by saying that authors *bring* their characters *into existence*, since for those realists, characters are nonexistent objects. But the creationist idea could alternatively be paraphrased by saying that the extent of the nonexistents that can meaningfully be talked about, referred to or quantified over in our world “supervenes” on the actions of the relevant existents in that world, in particular, the story-tellers and their story-telling activities. For **Graham Priest**, this idea of a supervenience of the fictional on the real is, at least, worth exploring. “Noneism”, the radical realist, neo-Meinongian account that Priest developed in his book *Towards Non-Being* (2005), had it that the domain of objects available for reference or quantification was the same at each world, some of these objects being existent, others not. In his essay “Creating Non-Existents”, he shows how to modify the original, “constant-domain” noneist account so as to get a “variable-domain” semantic account on which the supervenience idea can be captured: if two worlds have exactly the same existents and exactly the same properties for these existents, then they also have the same domain of objects available for reference and quantification and the same properties for these available objects. Priest goes on to contrast his original picture with the alternative supervenience picture in terms of their philosophical implications. They have different implications with respect to the “baptism” mechanism for fictional objects. The original account has it that Doyle gave the name “Holmes” to a nonexistent object picked out by (mental) pointing, but the object was already available to be pointed at. On the supervenience account, Doyle did not select Holmes, but created him in the sense that Doyle’s story-telling activities resulted in that nonexistent object becoming available for reference and quantification. Also, applying a (suitably modified) notion of identity as indiscernibility *à la* Leibniz may have different implications on the two accounts. If two identical stories happen to be told independently about a fictional character to which they give the same name, they turn out to bear on the same object if the identity as in-

discernibility condition is satisfied. On the original account, the answer to the question of the identity of the object is indeterminate: each of the two story-tellers selects an object and they both call their selected object by the same name, but there is no guarantee that they selected the same object. On the supervenience account, by contrast, whether we think that the object and its properties supervene on the content of the story—therefore give a positive answer to the question whether the object is the same—or on its being told by a particular story-teller instead—therefore give a negative answer to the question—, the question always receives a determinate answer. Priest concludes his essay by examining some other aspects of the supervenience picture, thus making it a serious contender in the realist camp.

The next five essays in the volume—by Adams, Sainsbury, Howell, Kroon, and Tavinor—all contest the truth of realism in some form or another, and assume the partial, if not the entire truth of irrealism in some form or another. Recall that irrealists deny that there are fictional entities, hence fictional entities that can be referred to or quantified over. At first glance, this amounts to rejecting what Sainsbury and others, borrowing a term from Fine (1982), call the “literalist” interpretation of such paratextual sentences as (4)–(8): the idea that there are things that literally have the properties attributed to them by these claims. But if this interpretation is rejected and the claims are not literally true, the challenge faced by irrealists, then, is to explain why and how they nevertheless can *seem* true to us, as they certainly do.

A natural reaction, here, would be to “blame it on the pragmatics”. This is what **Fred ADAMS** does in his essay “Sweet Nothings: The Semantics, Pragmatics, and Ontology of Fiction”, where he defends an irrealist position grounded on a pragmatic and cognitive approach to the meaning of names (developed in previous collaborative work of his with Dietrich, Fuller and Stecker). He starts by assuming a strong version of the direct reference view on names, on which a name has a meaning only if it has an actual, concrete, physical thing as its bearer, and merely and simply has no meaning otherwise. Since empty names, be they fictional or not, have no referent, they have no meaning. From this, two things follow. First, sentences with fictional empty names are neither true nor false. On a structured proposition view, the contents associated with sentences like

“Holmes lives at 221B Baker Street”, or “Holmes does not exist”, are respectively:  $\langle \text{____}, \textit{living at 221B Baker Street} \rangle$ , and:  $\neg(\exists x)[x = \text{____}]$ . Since no actual person is named by “Sherlock Holmes”, the slot cannot be filled. The content of the sentences is incomplete, and cannot be evaluated for truth or falsity. The problem is that we do evaluate the sentences that way, in actual fact, we judge them true. Second, two sentences that only differ in the fictional empty name they contain will be associated with the same incomplete content. Thus, the sentences “Sherlock Holmes does not exist” and “James Moriarty does not exist” both end up expressing the same aforementioned content:  $\neg(\exists x)[x = \text{____}]$ . The problem, of course, is that the two sentences seem to express two different things, one about Sherlock Holmes, the other about James Moriarty. To explain away the difficulties, Adams appeals to pragmatic and cognitive features of language. The basic idea is that language users associate informational and cognitive “storage files” with the names they use. This is true of filled as well as empty names. Thus, people have in their cognitive file for such names as “Sherlock Holmes”, “James Moriarty”, or “Vulcan”, a number of associated descriptions, say, “the detective living at 221B Baker Street”, “the archenemy of Sherlock Holmes”, “the planet between Mercury and the Sun that is responsible for the perturbations in the former’s orbit around the latter”, respectively. Now, according to Adams, although all above claims lack a truth value, they nonetheless trigger “extraction” of the files corresponding to the relevant names and exploit the extracted information so as to convey a number of truths—e.g. that there never was a detective by the name of “Sherlock Holmes” living at 221B Baker Street—or falsities. This explains the apparent truth or falsity of the claims. That also explains the apparent difference in what is conveyed by the claims: the bundles of descriptions that our cognitive files associate with different names, in particular with empty names like “Sherlock Holmes” and “James Moriarty” are not the same. The triggering of the corresponding files will thus result in different things being pragmatically conveyed, say, that there never was a detective by the name of “Sherlock Holmes” living at 221B Baker Street, nor anyone named “James Moriarty” that was an enemy of a detective named “Sherlock Holmes”. Adams defends this pragmatic irrealist explanation, and contrasts it methodically with semantic realist accounts of the creationist kind, focusing on Thomasson’s view alluded to above, as well as on the view, recently held by Martinich &

Stroll (2007), that talk about fictional characters is made true in virtue of “fictional facts”, facts which are ultimately institutional in nature.

Another way to explain away the literalist reading of fictional discourse is to concede that sentences of fiction and non-fiction containing empty terms are indeed true, rather than just seemingly true, but to insist that they are not true *tout court*, but somehow relative to the people that judge them so. This is the explanatory line followed by **Mark SAINSBURY** in his essay “Fiction and Acceptance-Relative Truth, Belief and Assertion”. Drawing on some of the key ideas defended in his book *Fiction and Fictionalism*, he maintains that an irrealist can successfully explain the intuitive judgment that a claim like (6), “Sherlock Holmes lived at 221B Baker Street”, is true, provided that she avails herself of a distinction between two kinds of propositional attitudes, acceptance on the one hand, belief on the other hand, and that she acknowledges the idea of acceptance-relative truth. Acceptance is defined in a non belief-entailing way, so as to make it possible to say that just like sales assistants easily accept that the customer is always right in the local context of interacting with customers, yet hardly believe it, likewise people will often accept (6), but will never believe it. Also, acceptance is context- or purpose-relative, in that one may well accept (6) and take it as true in the context of reading or rehearsing a Sherlock Holmes story, but clearly not in a context where what is at stake is the real world, say, while inspecting the residential records of real-life Baker Street in London. This allows for an account on which the sentence in question apparently makes for a true claim because it often is a true claim, except that the truth in question is not truth *simpliciter*, but truth implicitly relative to accepting the appropriate presuppositions. This role of presuppositions is attested to by the fact, which Sainsbury already stressed in previous work, that depending on where it occurs in a sentence, the name “Sherlock Holmes” will trigger different presuppositions and yield different judgments about the truth of that sentence. Thus, “Holmes met Gladstone” could be intuitively true: Doyle need only have written a slightly different story, and the sentence would trigger the presuppositions of that story. By contrast, “Gladstone met Holmes” is intuitively false because “Gladstone”, in that position, triggers the presupposition that we are talking about the real-life Gladstone, hence that Sherlock Holmes does not exist. This illustrates how we can do justice to both aspects of the ambivalence that we may feel about

the truth or falsehood of sentences like (6). Based on these ideas, Sainsbury also re-defines other truth-related notions, like acceptance-relative assertion, and offers a detailed characterization of acceptance as opposed to belief.

Although both Adams and Sainsbury seem to assume an irreducible gap between the realist and the irrealist positions, the gap might not be as large as it may appear. Thus, the idea that our intuitive judgments concerning the truth or falsity of fictional discourse may depend on what presuppositions or assumptions we make is deemed crucial by **Robert HOWELL**. In his essay “Fictional Realism and Its Discontents”, he breaks with his earlier realist commitment as he puts forward what he takes to be a heir of realism and irrealism. The latter is right that there are no fictional objects, but the former is right too that claims about them are literally true. To be precise, paratextual claims like (8), “Anna Karenina was a woman”, are to be considered literally true, but only in so far as they “express such truths only relative to the assumptions that we make (usually non consciously) in the course of reading the relevant fictions”. When, in the course of our reading, we run into a textual equivalent of (8), we cognitively register the name “Anna Karenina” and the predicate “is a woman”, and in doing so, we experience the name as being a proper name and the predicate as being a predicate that expresses the property of being a woman. Given that this is how our experience of reading the sentence goes, we form the unconscious metalinguistic judgment—about the words in the text we read—that:

- (19) There is an object  $x$  that is rigidly denoted by the name “Anna Karenina” and satisfies the predicate “is a woman” which expresses the property of being a woman;

*Semantic descent* can then take place, by which we move from that unconscious metalinguistic judgment to its object-language equivalent, viz. the conscious judgment expressed by claim (8). And we take this last claim to be in some sense literally true, because we already took the claim in (19) to express a truth about “the world of the novel” that we take the text to concern and whose domain of objects the judgment quantifies over. And this, in turn, is made possible by the further, usually unconscious assumption that there exists a unique concrete world, with its domain of objects, that the text of the novel truly describes. Given these unconscious

assumptions, our conscious judgment in (8) expresses the singular proposition  $\langle \textit{Anna Karenina, being a woman} \rangle$ , where *Anna Karenina* is that assumedly unique object, in the assumedly unique world of the novel, which is assumedly (rigidly) designated by the name “Anna Karenina”, and that proposition will be true at the world of the novel. It is because we are usually blind to the fact that we make singular claims about fiction only against a set of background assumptions, that we can take a claim like (8) to express a literal truth. Howell extends this account to other paratextual, as well as metatextual fiction-involving claims, and stresses how, by focusing on the etiology of the literalist readings, it retains the best of realism and irrealism.

Yet another way to explain away the literalist readings is to dispense totally with them, by considering intuitive reference to the “truth” of a fictional discourse to be merely and simply a *façon de parler*, a fiction by itself. This corresponds to the so-called “fictionalist” approach to fiction, made famous by Kendall Walton (1990) and people influenced by his work. According to Walton, works of fiction are props in authorized games of “make-believe” whose rules stipulate that this-or-that is to be imagined, e.g. that the names that occur in the work have real-world objects as referents and that the sentences in the work describe these objects. In saying “Holmes lived at 221B Baker Street”, what one expresses is that it is appropriate to say so in the game of make-believe authorized by Doyle’s stories, that to say so is to speak truly in the game in which such stories are props. But truth in a game of make-believe is not for real, it is only fictional.

Fictionalism seems to have become the irrealist counterpart of creationism in terms of popularity. Recall that an advantage of creationism over more radical forms of realism is that it offers a straightforward explanation for our ordinary and pervasive use of creationist locutions. This by no means entails that the idea that characters are somehow created by their authors is simply dismissed by fictionalists, though. Indeed, in his essay “The Fiction of Creationism”, **Frederick KROON** contests the creationist’s possible monopoly of that idea, and argues that not only can an important part of our talk about fictional characters’ being “created” be given a fictionalist reading, but “creationist” locutions, rightly understood, also speak for a fictionalist point of view. His analysis focuses on metatextual claims of literary criticism like that in (14). This sentence

does seem to express something true, and this fact can be used in what was presented above as the main argument for creationism. Taken at face value, the truth of that sentence entails that of fictional creationism: since it quantifies existentially over characters, there must be characters, hence fictional objects; and since it says that characters are created by authors, those fictional objects must be created artefacts. However, as Kroon points out, the line of reasoning behind the argument overgenerates. A particularly interesting example in this respect involves claims made by experts concerning the phenomenon of imaginary companions, these for instance:

- (20) Fun and companionship are the primary reasons most children create imaginary companions.
- (21) [Some] imaginary companions are more closely modelled after playmates of the child's own age, size and gender.

Just because they mention imaginary companions, seem to quantify over these, and involve creationist locutions, it should not follow that their possible truth supports creationism about imaginary companions, an implausible sounding position according to Kroon. He concludes from this overgeneration problem that creationism about fictional characters is not supported by the *prima facie* truth of discourse involving creationist locutions. To account for the meaning of these locutions, Kroon suggests that talk about fictional characters be understood, precisely, on the model of talk about imaginary companions. Just as children pretend in their games that their imaginary companions exist, likewise authors in their writings pretend that their fictional characters exist; and just as psychologists continue the children's pretense in order to describe, "from the inside", imagination at work by using such locutions as "made up by so-and-so", likewise literary critics continue the authors' pretense in their comments when they use locutions like "being created by so-and-so". The legitimate use of creationist locutions can thus be accounted for without thereby having to concede the truth of creationism. In other words, there is room for what Kroon calls "fictionalist creationism".

Fictionalism is multifaceted and its resources make it fairly polyvalent, as **Grant TAVINOR** clearly illustrates in his "Virtual Worlds and Interactive Fictions". In this essay, he sets out to show that a fictionalist approach based on Walton's prop-based analysis of fiction allows for

a neat analysis of videogames as interactive fictions. This involves giving credit to the rather intuitive thought that videogames, like other kinds of digital artefacts, are simply another kind of fictional work, just like books, films, etc. Such a connection between videogames and fictions, however, can be contested on the grounds that the virtual worlds of videogames and the fictional worlds of novels and films differ in their level of interactivity: one can battle with a dragon character in a fantasy videogame while one cannot battle with a dragon character in a fantasy novel. According to Tavinor, to oppose the virtual and the fictional on such grounds is to fail to notice that these two notions belong to two distinct dimensions of the media. Media like books, films, etc., can be used to represent things. The things in question can either be real—in the case of history books or documentaries—, or fictional; in the latter case, the represented things are only imagined to exist. On the other hand, some media support representations that are structurally and interactively isomorphic with what they represent, in which case they may be deemed virtual. Now, this is compatible with the idea that some media may support structurally and interactively isomorphic representations of fictional items. Videogames are, precisely, cases of such virtual fictions. It is fairly natural to describe a player's engagement in a virtual fiction of this sort by mentioning her interactions with virtual items, for instance her battling with a dragon. To account for this way of talking about how a player may relate to the game she plays, Tavinor exploits the core ideas of the prop-based approach to fiction in order to maintain that there is properly speaking no causal interaction with the dragon in the videogame: what there is is a virtual fictive prop, that is, an animation on the screen, that represents a dragon and undergoes a series of graphical transformations in accordance with the moves of the controllers in the hands of the player. So, it is true only of an interactive fiction that the player battles with a dragon, which is another way of saying that “players do not causally interact with the worlds of videogames, it is only fictional that they do”.

All the essays presented so far were concerned, directly or indirectly, with issues to do with the connection between fiction and truth. The essay that closes the volume, “Fiction, Indispensability and Truths”, is no exception, but it adds a further dimension to the question. In this essay, **Manuel REBUSCHI** and **Marion RENAULD** hold that there is much in-



terest in reflecting on how a fiction might tell us something true about the real world and the people who, just like us, are part of it. To make sense of this idea of the truth *of* a fiction, they propose to apply to fictions a variant of Quine and Putnam's well-known "Indispensability Argument" for the existence of mathematical entities. As they read it, the argument involves two transitions: one from indispensability to truth, the other from truth to ontological commitment. What they question in the argument thus read is the idea, which corresponds to Quine's ontological commitment criterion, that acknowledging a discourse as true bears some ontological commitment to the existence of the entities in the intended domain of the discourse. According to Rebuschi and Renaud, developments in natural language semantics tend to indicate that although Quine's criterion might well apply to highly idealized forms of discourse for which an external and representational approach to truth, focusing on how language relates to the world in order to be meaningful, seems sufficient, that criterion is irrelevant when it comes to accounting for a number of cases for which an internal and procedural approach, focusing on how language-users process language in order to build meaning, is better suited. They conclude that there is no natural step from truth to ontology. However, they retain the idea that the indispensability of a kind of discourse can ground acknowledgement of its truth. Fictions are indispensable in that in certain respects, they are useful to us in a way that nothing else can be. Of course, they are not theoretically indispensable. The indispensability is not theoretical, it is practical. Based on analyses from literary criticism and philosophical aesthetics, Rebuschi and Renaud argue that fictions, novels in particular, are indeed useful to us in that they play an incomparable role in helping us advance our understanding of reality, by drawing our attention to various aspects of life that ordinarily go unnoticed, and by suggesting different, sometimes unusual, sometimes extraordinary, interpretative frames for understanding the world and others. With respect to this hermeneutical role, fictional discourse knows of no non-fictional equivalent. Part of what makes for this irreducibility is that fictional discourse aims at grasping reality in its full complexity, extends the realm of the meaningful far beyond our actual reality, relies heavily and constitutively on the symbolic and the metaphorical, fully exploits the stylistic and poetic dimensions to trigger multiple readings of reality that can thereby hardly be paraphrased, and leaves it to the reader and its context to fill in the blanks not only in the

told fiction, but also in respect of which of several possible interpretations of the same fiction is correct. Being irreducibly useful in such respects, fictions can be said to be indispensable, hence true, except that the truth in question consists not in being descriptively and theoretically linked to reality, but in being interpretatively and practically linked to it.

# Fictional Characters and Indeterminate Identity

TERENCE PARSONS

In *NEO (Nonexistent Objects)*, Parsons, 1980) I developed a Meinongian theory of objects that includes objects that do not exist. The first objective was to formulate such a theory, and a second was to find a use for it. The most obvious use appears to be as the underpinning for a theory of fictional objects.<sup>1</sup> It is a truism that Sherlock Holmes does not exist, and I explored the idea that this statement can be taken in its most straightforward sense, which is to deny the existence of a certain fictional character (a certain object) referred to as “Sherlock Holmes”. The theory that resulted faced certain difficulties of detail, difficulties that have always bothered me, but that I could not improve on using the theory at hand. I will try to do better here.

I begin by reviewing the theory of nonexistent objects with its application to fiction. Then I state the difficulty, which has to do with how many Sherlock Holmes’s there are according to the theory. (The theory seems to say that there are many.) I then discuss a new way to handle this question by using the notion of indeterminacy; this permits one to hold that there is exactly one Sherlock Holmes. This development raises new objections, which I will try to meet.

---

<sup>1</sup> By “fictional” I mean “appearing in fiction”. So calling an object fictional does not imply that there is no such object.

# 1 The Original Theory

## 1.1 The Ontology

I begin by reviewing relevant parts of the theory from *NEO*. The Meinongian metaphysical theory tries to maximize the objects that there are; not the objects that exist—I suppose that there is consensus about these—but all objects, including those that do not exist. The way I hit upon was first to suppose, following Leibniz, that objects are individuated in terms of the properties that they have; objects with exactly the same properties are identical, and objects that disagree concerning their properties are distinct.

**(Individuation)** Objects are identical iff they have the same (nuclear) properties.

I then proposed that for any set of properties whatever, there is an object that possesses exactly those properties.

**(Generating Principle)** For any class of (nuclear) properties, there is an object having exactly the properties in that class.

If you take the set of properties possessed by an actually existing object, the Generating Principle says that there is an object having exactly the properties in that set. In this case, this is the existing object that you began with. If you take some other set of properties, say the set consisting only of goldenness and mountainhood you get a highly incomplete nonexisting object, one possessing only those two properties. This may be the object that Meinong had in mind when he referred to *the golden mountain*. Similarly, the properties of roundness and squareness yield the round square, a nonexisting object that is both incomplete and impossible. (“Impossible” has a special meaning here; an object is impossible iff it has some (nuclear) properties that no existing object could possibly have. So to say that an object is impossible is not to say that there is no such object. It is saying that it is an object that has properties that cannot co-occur in anything that exists.)

It is soon apparent that the properties appealed to in the Generating Principle must be limited in some way, in order to avoid Russell’s famous objections in “On Denoting”. If you begin with goldenness and mountainhood and existence, you seem to be committed to a golden mountain

*that exists*, which runs counter to the policy of leaving the existing objects untouched, and only adding nonexistent ones. Meinong's response to this objection was essentially that you cannot use the predicate "exists" in this way. The properties appealed to in the generating principle were later called *nuclear*, and the rest were called *extranuclear*. On one view, existing is not a property at all; on another view, it is a special kind of property, an *extranuclear* property. But only nuclear properties are to be used to generate objects. Meinong thought that genuine extranuclear existence has a corresponding "watered-down" version, the nuclear property of *being existent*. So there is such an object as the *existent* gold mountain, but this object does not *exist*, thus avoiding the paradox. It is apparent at this point that it is important to clarify the distinction between nuclear and extranuclear properties, and there is an effort to do this in *NEO*. Here, I can take for granted that there is such a distinction, because its exact form is not crucial to the problem I wish to discuss. I will simply use the rough guideline that nuclear properties are ordinary properties such as goldenness and mountainhood, and extranuclear are special ones, such as existence. At times I will speak as if nuclear properties are all of the properties that there are; extranuclear *predicates* are predicates that fail to pick out properties. It won't matter much whether there are no extranuclear properties, or whether they are properties of some special kind. The important principle is that it is nuclear properties that figure in the basic assumptions above.

## 1.2 Objects of Fiction

The next step is to apply this theory to fiction. What is required here is to find a way to identify each apparent fictional character with a unique object. Best of all is to find a natural way to do this. An obvious thing to try is to identify each character with the object that possesses exactly those (nuclear) properties that the character is understood to have in its story:

**(Fictional Characters)** A fictional character *c* is the object *o* which has exactly the (nuclear) properties attributed to *c* in the story/stories in which *c* originates.

So Sherlock Holmes is to be that object that has exactly the properties attributed to him in the Sherlock Holmes stories, and no others. This means that it will end up being literally true that Sherlock Holmes is a detective, because that is a nuclear property that he is understood to have in the stories about him. It will not be true that he exists, even though he does exist according to the story; this is because “exists” is an extranuclear predicate, and the formula for determining which object Sherlock Holmes is appeals only to nuclear properties. We know that he does not in fact exist, because we know that there is no object among existing objects with exactly the nuclear properties that Holmes has. Further, this will automatically be true of each character of fiction which originates in fiction, because such characters are all “incomplete”: for each such fictional object, there will be many nuclear properties such that it does not possess that property, nor does it possess the nuclear negation of that property:

**(Incompleteness)** An object  $o$  is incomplete iff there is some (nuclear) property  $p$  such that  $o$  does not have either  $p$  or non- $p$ .<sup>2</sup>

Fictional characters originating in fiction are incomplete because each actual work of fiction is finite, and it attributes only a limited set of nuclear properties to its characters. For example, Holmes has the property of being a detective, as well as the property of being a non-dragon (since it is understood in the stories that he is not a dragon). But he lacks both the property of having had tonsillitis, as well as the property of not having had tonsillitis. All existing objects must have one of the other of those nuclear properties. (Most of them, such as rocks, have the latter.)

### 1.3 Native and Immigrant Characters

One additional factor will be relevant below. I distinguish between characters that are native to a story from those that are immigrants to the story. A character is native to the story which defines that character; so Holmes is native to (at least) the Conan Doyle stories, in which he is created. The city of London, however, is not created by the Conan Doyle stories, nor by any other work of fiction. London is an immigrant appearing in those

---

<sup>2</sup> Non- $p$  is the nuclear negation of  $p$ . See *NEO*, pp. 19–20, 105–6, 227–8.

stories. Immigrant objects are not subject to the principle that identifies fictional characters; otherwise the theory would contradict known fact. For example, if a story were to portray London as being located on the Seine, our principle would identify the nuclear property of being located on the Seine as a nuclear property of London, whereas London in fact lacks that property. It is only characters that are native to a story that can be assigned properties based on what is said in the story.

## 2 How Many Sherlock Holmes Are There?

Let me turn now to a problem that has nagged at me over the years. The proposal I have been discussing identifies a fictional character in terms of the properties it has in its originating story. But does each character have a *unique* originating story? If not, which story are we to use? If a fictional character originates in multiple texts, which are to be definitive? Suppose that there is an initial primary text, followed by multiple spin-offs, unauthorized extensions, and/or alternative stories. How do these non-primary texts affect which object we refer to when discussing a character that occurs throughout them?

One concrete instance of this question is: How many Sherlock Holmes's are there? In the actual case of Sherlock Holmes, one thinks first of the Conan Doyle series of stories about him as being his source. So perhaps he is native to these stories and to no others. If so, he actually has the properties attributed to him there, and no others. But several other options are available, and we need some principled reason to rule them out. Why not make the source bigger? The series of movies starring Basil Rathbone and Nigel Bruce feature Holmes as a main character, so why not count the movies in determining his properties? This raises a problem, because the plot of the movies contradict that of the Conan Doyle stories in many ways, one being that in the movie Holmes is not bald, whereas in the stories he is.<sup>3</sup> So do we include baldness among his properties, or not? I argued in *NEO* that we can tolerate some inconsistencies in the source stories of fictional characters so long as they occur in ways that do not interfere with our understanding the story; this is especially so if the

---

<sup>3</sup> This is based on illustrations accompanying the original stories. Actually, he is portrayed there as being *partially* bald; I say "bald" in the text for brevity.

ingredients of the inconsistencies are remote from one another. But in the case of Holmes' baldness the inconsistency is obvious, and we end up not knowing which properties Holmes has, and thus the theory does not yield any such object as Holmes.<sup>4</sup>

There are several ways to address these problems. One tempting way is to suppose, on grounds of artistic priority, that Holmes is native to the Conan Doyle series, and merely an immigrant character in the movies. If this is so, our Fictional Character principle does not refer to any movie in judging Holmes' properties, and the problem is apparently solved. But this solution is not a stable one. For consider now the stories of the series itself. Why not imagine that Holmes is native to the first story and an immigrant to the rest? After all, there is nothing in the first story to indicate that it is part of a series—so far as a reader is concerned, the original story is complete.

We can probably rule this out on the grounds that readers clearly understand the later stories as a continuation of the earlier ones, and not independent stories in which Holmes appears as an immigrant. But this cannot solve the problem, because we can readily imagine intermediate cases where the distinction is harder to make. Suppose that after Doyle stops writing, the publisher hires someone else to continue the series, and new stories are added that are somewhat at odds with the others. In many such cases it is not at all clear what to say.

It may be tempting at this point to conclude that this kind of problem simply shows the absurdity of the original idea of admitting nonexistent objects at all, or trying to find fictional objects among them. I think this would be hasty—because the same kinds of problems arise on views about fiction that do not admit nonexistent objects. An example is Amie Thomasson's view in *Fiction and Metaphysics*. Thomasson holds that fictional characters are abstract artefacts that have dependent existence. They are brought into being by the creative activities of authors, and maintained

---

<sup>4</sup> I am not worried about supposed anomalies such as the death of Holmes in one of Conan Doyle's stories, which is contradicted in the next story. The later story in this case leads to a *reinterpretation* of the earlier story, in such a way that it is not true in that story that Holmes died. A similar case arises from the fact that in Tolkien's *The Hobbit*, the evil ring is lost, whereas in the *Trilogy of the Rings* we discover that Frodo, who narrated *The Hobbit*, lied about this, and kept the ring. This is just additional information to be used in interpreting *The Hobbit*: the narrator lies, and what actually happens in that story is that Frodo keeps the ring.



in being by the continuing existence of copies of their story. This is a quite different idea, and it seems to face the same kinds of problems described above. Is Sherlock Holmes a character whose being is dependent on the writings of Conan Doyle? Or on the joint writings of Doyle and the makers of the movies? Or on the first story in the Conan Doyle series? These issues will arise on almost any approach, regardless of the ontology.

### 3 A Contextual Solution and Difficulties with It

In *NEO* I suggested a contextual solution to this problem. I noted first that in discussions about literature we sometimes appear to distinguish “the Holmes of the Conan Doyle series” from “the Holmes of the movie”, speaking as if these were different characters, and attributing different things to them. We even distinguish “the Holmes of the early stories in the series” from “the Holmes of the later stories”. Relying on this practice, I suggested that there is no unique source for Sherlock Holmes—any old body of literature involving Holmes will do. Because of the Fictional Character principle, this yields not one Sherlock Holmes, but many. Which one is under discussion at a given time will depend on the context of discussion. In essence, I suggested that the name “Holmes” is ambiguous, referring sometimes to one of these, and sometimes to another. Since the context generally makes it relatively clear which Holmes is under discussion, and since they are all so similar, the ambiguity is harmless.<sup>5</sup>

This now seems to me to be wrong for at least three reasons. First, it seems to me that most statements like “The Holmes of the early Conan Doyle stories is not nearly as clever as the Holmes of the later Conan Doyle stories” are not intended to be about what Holmes is really like, but rather about what is so in the stories. Unless we are making definite predications about two distinct objects, the point being made is not relevant. Second, I suspect that it is wrong to suppose that when we discuss “the Holmes of the early stories” as opposed to “the Holmes of the later stories” we intend to be discussing different characters. This way of talking parallels talk about real objects, where we are *not* distinguishing different objects at all. For example, in New Testament exegesis we

---

<sup>5</sup> For statements made out of context, one can always employ the semantic trick of using supervaluations to account for apparent agreement.

speak as if we distinguish “the early Paul (before his conversion)” from “the post-conversion Paul”, yet it is clear that we do not intend to be contrasting two distinct people at all. Instead, we intend to distinguish what one and the same person was like at two different times. If the usage is the same with fictional characters, we are not distinguishing one character from another, but rather distinguishing what the same character was like in different parts of literature. The third problem is that when we speak of Sherlock Holmes we suppose ourselves to be speaking of a unique character, who in some sense appears in diverse bodies of literature. Certainly, if I ask you a question about Sherlock Holmes, your natural reaction is not “which one?”

I will thus proceed below to see if it is possible to maintain that there is only one Sherlock Holmes, and that this is the actual truth, and not just an appearance of language.<sup>6</sup>

## 4 A Solution in terms of Indeterminacy

If there is only one Holmes, we need to ask what he is like. What should we say about the question of whether Holmes is bald, or not? I think there is no answer to this question. At least, I can’t come up with one, and it seems wrong to me to expect one. And so the right (meta-)answer is that there is no answer.

Notice that I am not saying that Holmes is neither bald nor non-bald. In the Meinongian theory under discussion, that is a definite answer, not a rejection of an answer. The ontological theory of *NEO* is formulated from a bivalent point of view, which supposes that there are exactly four options: Holmes is both bald and non-bald, he is neither bald nor non-bald, he is bald and not non-bald, or he is non-bald and not bald. In the first case he is impossible, in the second case he is incomplete with respect to baldness, and in the third and fourth cases he is purely bald or purely non-bald. Each of these is an answer. And none of them seem right to me regarding Holmes. I want to say instead that there is no answer at all. This raises the issue of indeterminacy; a phenomenon which results in meaningful questions not having answers.

---

<sup>6</sup> Such as the appearance of oneness that the method of supervaluations can yield.

## 4.1 The Nature of Indeterminacy

Traditionally, there are two ways to conceive of indeterminacy. One way is semantic: some questions do not have answers because the language in which the question is formulated is defective. For example, there is no answer to the question of how tall my sister is, because I have no sister. The question presupposes that I have a sister, and since I don't, it has no answer. Many people think that semantic indeterminacy of this sort is the only kind there is. This is because they think it makes no sense to think of the world itself as indeterminate. This merits a long discussion, which is not included here. Instead, I am going to suppose that there is indeed real indeterminacy in the world, and use this assumption as a basis for theorizing. For simplicity, I will assume that there is nothing semantically deficient in our talk about fictional objects.<sup>7</sup>

I hold that worldly indeterminacy is indeterminacy in states of affairs. For example, it may be indeterminate whether a certain object has a certain property or not. Such a state of affairs—of object  $o$  having property  $p$ —is then indeterminate. If “ $n$ ” is a name naming the object  $o$ , and “ $P$ ” a predicate picking out the property  $p$ , then the sentence “ $Pn$ ” has no truth-value, because the state of affairs that it expresses is indeterminate. The idea is straightforward enough. It raises lots of metaphysical and logical questions which I pursued in *Indeterminate Identity* (Parsons, 2000), and will not address here. We must, however, reformulate the theory from *NEO* with which we began, in order to allow for indeterminacy in that framework.

Our original Generating Principle was formulated within a bivalent framework, and we need to see what it says within a framework admitting indeterminacy. In fact, it needs extension. As it stands, the Generating Principle says:

Suppose  $C$  is a class of (nuclear) properties. Then there is an object  $o$  such that:  $(\forall p)(o \text{ determinately has } p \text{ iff } p \in C)$ .

This says nothing about an object having some properties indeterminately. For that we need the more powerful principle:

---

<sup>7</sup> Of course, there are semantically deficient questions we can raise about fiction. An example would be: “How tall is the unicorn in the Conan Doyle stories?” There is no unicorn in the Doyle stories, so the question is defective. But questions like “Is Sherlock Holmes bald?” are not semantically defective, and these are the ones we will address.

**(Generating Principle 2)** Suppose that  $C_1$  and  $C_2$  are definitely disjoint classes of (nuclear) properties. Then there is an object  $o$  such that:

- $(\forall p)(o \text{ determinately has } p \text{ iff } p \text{ is determinately a member of } C_1)$ , and:
- $(\forall p)(\text{it is indeterminate whether } o \text{ has } p \text{ iff } p \text{ is determinately a member of } C_2)$ .

Do we also need a new Individuation principle? I am supposing here that the Leibniz principle of identity is definitional of identity. An identity “ $x = y$ ”, when embedded in a framework allowing for indeterminacy, has the same truth-value status as the statement that  $x$  and  $y$  agree in all of their (nuclear) properties. This is our original Individuation principle, which in symbols is:

$$\text{(Individuation)} \quad x = y \stackrel{\text{def}}{=} (\forall P)(Px \leftrightarrow Py)$$

Our original individuation principle is retained, and its effect differs only in that the *definiens* can now lack truth-value, and thus so can identity statements. This will happen when  $x$  and  $y$  do not definitely disagree in any property, but where one, say  $x$ , has a property such that it is indeterminate whether  $y$  has it. (That is, just in case there is no property  $p$  such that  $x$  definitely has  $p$  and  $y$  definitely lacks  $p$  (or vice versa), but where there is some property  $p$  such that  $x$  definitely has  $p$  and it is indeterminate whether  $y$  has  $p$  (or vice versa).)

## 4.2 Application to Fictional Objects

I suppose that fictional characters can enter into indeterminate states of affairs. In the case of Holmes, the question of whether he is bald has no answer because the state of affairs of his being bald is indeterminate.

This is not the same as incompleteness. If Holmes’ source story definitely fails to specify whether he is bald, then he will be neither bald nor non-bald (he will definitely not have baldness, and he will definitely not have non-baldness), and thus he will be incomplete with respect to being bald. This is not indeterminacy; it is incompleteness. Indeterminacy is not when there are definite answers to “Is he bald?” and “Is he non-bald”

(namely, “no” and “no”); it is when there is no answer to these questions at all. What I am proposing is that since the Conan Doyle stories clearly indicate that Holmes is bald, and the movie clearly indicates that he is non-bald, then there is no answer to the question of whether he is bald (and no answer as to whether he is non-bald). The proposal is this: take the whole series of plausible sources of Sherlock Holmes stories, and consider the properties attributed to him in those sources taken as a whole. If a nuclear property  $p$  is attributed to him, he has it; if non- $p$  is attributed to him, he has that (and if both are attributed to him, he has both, and if neither, he has neither). He is actually to have any nuclear property attributed to him in this way, and actually lack any property which is not attributed to him in this way. But there will be properties for which the stories disagree, and thus there is no answer as to whether or not they are attributed to him. In that case, he is indeterminate with respect to that property. I think that this does not undermine the identification of what object Holmes is; instead, it simply makes indeterminate some questions asked about him. The result is that there is a unique Sherlock Holmes, and the uncertainties about how to identify him reappear as indeterminacies about what he is like. We do indeed all discuss the same thing when we discuss Sherlock Holmes, and the questions about him that confound us turn out to be questions that legitimately have no answers.

This solution I propose to the problem of the many Sherlock Holmes needs some clarification. There are two forms that it might take:

**(Proposal 1)** Holmes is to be identified as the object having those properties attributed to him in some candidate source and on which no candidate sources disagree, and lacking those properties which all such texts definitely leave open; in case there is a property,  $p$ , about which the texts positively disagree, then there is no matter of fact about whether the object has property  $p$ .

**(Proposal 2)** Holmes is to be identified as the object having those properties attributed to him on the basis of treating the whole corpus of literature as a single story, so that he would have exactly those properties attributed to him in that story, he would lack those properties which the story definitely does not attribute to him; in case there is a property,  $p$ , such that the

story is indeterminate about whether he possesses  $p$ , then it is indeterminate whether he has  $p$ .

The first idea treats the multiple sources independently, whereas the second combines them into one big story and treats that as the source.

When I say regarding the first option that the stories are treated independently, I do not mean to deny that some texts influence the content of others. E.g. I take for granted that each of the Rathbone/Bruce movies presents a story whose content relies on the Conan Doyle series to flesh out many aspects of Holmes's history. The story associated with the movie then is determined in part by the text of the Conan Doyle series. It is *this* (affected) story that is to be used in Proposal 1 above.

I should say that I am not at all sure whether the second proposal makes sense. This is because I am not at all sure that you *can* combine the Doyle series with the movie to get a story at all. If you can do so, you will need some way of treating the aspects in which the two sources rather explicitly contradict one another. I am taking for granted here that if that happens then if one of the sources clearly indicates that Holmes has  $p$ , and the other clearly indicates that he does not have  $p$ , then in the combined story it will be indeterminate whether he has  $p$ . Not that he will be incomplete with respect to  $p$ —for that, you need the combined story to clearly indicate that he does not have  $p$ , and that he does not have non- $p$  either. This is determinate incompleteness with respect to  $p$ , whereas the proposal is that he be indeterminate with respect to  $p$ .

Because of the doubtful nature of the second proposal, I will confine myself to discussing the first.

## 5 What is Sherlock Holmes like?

The proposed solutions naturally bring their own problems with them. An obvious problem has to do with whether it makes sense to say that we have identified Holmes at all. On the proposed view, identity of objects is determined by them having exactly the same nuclear properties. Once it becomes indeterminate which properties an object has, doesn't it become indeterminate which object it *is*? Have we identified anything at all?

It may help to clarify the relations between Sherlock Holmes on this view with the various other Sherlock Holmes's proposed in the earlier

theory. Consider first the object Sherlock Holmes according to Proposal 1. Call this object  $H$ . To get  $H$ , we first specify  $C_1$  and  $C_2$  as follows:

1.  $C_1$  is the class that contains a property  $p$  iff either Holmes is  $p$  according to both the Conan Doyle stories and the movies, or Holmes is  $p$  according to the Conan Doyle stories and Holmes is not non- $p$  according to the movies, or Holmes is  $p$  according to the movies, and Holmes is not non- $p$  according to the Conan Doyle stories.
2.  $C_2$  is the class that contains a property  $p$  iff either Holmes is  $p$  according to the Conan Doyle stories and Holmes is non- $p$  according to the movies, or Holmes is  $p$  according to the movies, and Holmes is non- $p$  according to the Conan Doyle stories.

Our Generating Principle 2 yields the object  $H$  as described above.

Now let us compare object  $H$ , with the previous “Sherlock Holmes of the Conan Doyle series”, call him  $H_{CD}$ , and the previous “Sherlock Holmes of the movies”, call him  $H_M$ . Recall that  $H$  is the object that has the properties from the combined sources,  $H_{CD}$  has the properties from the Conan Doyle series, and  $H_M$  from the movies. In some respects all of these objects are alike; e.g. each is a detective, and each is a non-policeman. In some respects they are all incomplete (e.g. regarding Holmes’ tonsillitis). But in additional respects they differ:  $H_{CD}$  is bald and  $H_M$  is not bald. What then of  $H$ ? Since the sources disagree on bald, there is no answer to the question of whether  $H$  is bald. This will in general be the case;  $H$  will agree with both  $H_{CD}$  and  $H_M$  where they agree with one another, and will be indeterminate when they disagree. This does not mean that  $H$  disagrees with  $H_{CD}$  and with  $H_M$  for some property; it means that it is indeterminate whether there is disagreement. Particularly, this is not false, it is indeterminate:

$$H \text{ is bald} \leftrightarrow H_{CD} \text{ is bald.}$$

(The biconditional used here is true when the sides are either both true or both false or both indeterminate, it is false when one is true and the other false, and it is indeterminate in all other cases, in particular when one side has a definite truth value and the other is indeterminate.)

This lack of determinate disagreement between  $H$  and the other two candidates has the following consequence: although  $H_{CD}$  and  $H_M$  are distinct objects, it is indeterminate of each of these whether it is identical with  $H$ . This is a consequence of the Individuation principle:

$$x = y \stackrel{\text{def}}{=} (\forall P)(Px \leftrightarrow Py).$$

Applying this to  $H$  and  $H_{CD}$  yields:

$$H = H_{CD} \text{ iff } (\forall P)(H \text{ is } P \leftrightarrow H_{CD} \text{ is } P).$$

Based on what we said above, some of the instances of the universal generalization on the right are indeterminate and the rest are true, and this makes the generalization itself indeterminate. So the identity is also indeterminate. So if one asks whether the Holmes of the Conan Doyle novels *is* Holmes, the answer is that there is no answer. This strikes me as being correct.

## 6 More Problems with Indeterminate Identity of Fictional Characters

I said above that there is exactly one Holmes, namely  $H$ . But now it turns out that it is indeterminate whether  $H$  is  $H_{CD}$  and indeterminate whether  $H$  is  $H_M$ . Won't that mean that it is false or indeterminate whether there is exactly one Holmes?

The answer to this will depend on what you mean by there being exactly one  $A$ . In a non-classical framework involving indeterminacy, different characterizations of "exactly one" which are classically equivalent are no longer equivalent. This is why answers will vary. But there is at least one natural formulation which gives the desired result:<sup>8</sup>

There is exactly one  $A \stackrel{\text{def}}{=} \text{Both of the following conditions hold:}$

- There is at least one  $A$ :  $(\exists x)Ax$ ;
- There is at most one  $A$ :  $(\forall x)(\forall y)(Ax \wedge Ay \rightarrow x = y)$

---

<sup>8</sup> See Parsons (2000) for details.



It is then a matter of technical detail to show that both of these are true when we use for  $A$  the defining conditions for Holmes given above. (That is, we use “ $Ax$ ” to mean “ $x$  satisfies the revised generating principle for  $C_1$  and  $C_2$  as defined above”.) The first is true because the revised generating principle produces an object that satisfies “ $Ax$ ” as specified above, and the second is true because the principle yields a unique object—that is, yields objects  $a$  and  $b$  if and only if  $a$  and  $b$  determinately have the same properties, and indeterminately possess the same properties.

## 7 Loose Ends

The attempt to find a single Sherlock Holmes has been somewhat complicated, and has involved overcoming serious obstacles. Unfortunately, others remain. The method I have used has supposed that we can divide potential sources into those that count toward the source and those that do not. This may not be so. The worry is illustrated by reference to another Sherlock Holmes movie, *The Seven Percent Solution*. In this movie, the character called Sherlock Holmes is not smart at all; he bumbles around foolishly, while Dr. Watson manipulates situations in such a way that it appears as if Holmes comes up with brilliant solutions to the crimes. If we were to count this movie as one of the candidate sources, it would then turn out that it is indeterminate whether Sherlock Holmes is clever. This strikes me as false as a matter of intuitive data; whatever Holmes turns out to be like, he can't fail to be clever. But how can we rule out a movie that purports to be about Sherlock Holmes? The only way that seems right to me is to suppose that Holmes is not a character native to this movie. Instead, the movie is about him in the same way that the Conan Doyle stories are about London; he is used in the story as an already familiar thing, and various things are said in the story about him that are not true of him. This is no problem because in determining which object Holmes is, we only count the stories to which he is native. The movie is not one of these.

Even if the above works in the case of *The Seven Percent Solution*, there are sure to be other cases that are less clear. At the moment, I don't know what more to say about them.<sup>9</sup>

---

<sup>9</sup> This paper was first presented in March, 2002, at a conference in Lyon, France, at the École Normale Supérieure des Lettres et des Sciences Humaines. I wish to thank the organizers and participants.

# Two-Dimensionalism and Fictional Names

BRENDAN MURDAY

## 1.1 The Problem

The prefix “According to the fiction” (or some equivalent) is often used when we talk about fiction. I am interested in the semantics of this prefix, as it seems to pose a problem for at least some approaches in the philosophy of language. Consider the English sentence:

(1) According to the fiction, Sherlock Holmes is clever.

Intuitively, this sentence expresses a truth: there is a story whose main character is Sherlock, and in that story, Sherlock is clever. Now consider:

(2) Sherlock Holmes is clever.

If we grant the assumption that “Sherlock” is an empty name,<sup>1</sup> (2) is false (or at least it fails to express anything truth-evaluable).<sup>2</sup>

Supposing that (1) expresses a truth and (2) does not,<sup>3</sup> how exactly does the prefix behave such that applying it to the false statement (2) generates a true statement? One might have thought that “According to the

---

<sup>1</sup> That assumption is certainly controversial: there are views according to which “Sherlock” *isn't* an empty name, such as descriptivism about names or the view that fictional entities are actual abstract objects. I will return to this briefly in §2.2, but generally this paper is aimed at those who think that “Sherlock” is an empty name.

<sup>2</sup> I will suppose that ordinary sentences with empty names in the subject position are false, but nothing in the chapter depends on it.

<sup>3</sup> See §2.1 for a little more about this supposition.

fiction” expresses a sentential operator analogous to common sentential operators such as negation (as expressed by “it is not the case that”) or modal operators (as expressed by “necessarily” and “possibly”). However, there is a critical difference between them. The familiar modal operators (on a popular understanding of them, at least) and negation apply to propositions, and produce new propositions. But the fictionalist operator does not seem to apply to propositions. A closer look at the behavior of these operators will bring out the difference. Negation applies to a proposition, and when so applied, produces a new proposition. Supposing that propositions are structured objects and that names are directly referential:<sup>4</sup>

(3) Jack is happy.

expresses the proposition:

(4)  $\langle \text{Jack}, \textit{happiness} \rangle$ .

A sentence formed by applying the negation prefix to (3) such as:

(5) It is not the case that Jack is happy.

expresses the proposition:

(6)  $\langle \text{Negation}, \langle \text{Jack}, \textit{happiness} \rangle \rangle$ .

The crucial point of interest is that applying the negation prefix to (3) leaves the semantics of the embedded clause unaltered. (4) is a *constituent* of (6)—the negation operator applies to the proposition expressed by the embedded sentence.

Modal operators (on a common interpretation, at least) resemble negation in this respect. The sentence:

(3) Jack is happy.

expresses the proposition:

(4)  $\langle \text{Jack}, \textit{happiness} \rangle$ .

To determine whether the sentence:

(7) Necessarily, Jack is happy.

---

<sup>4</sup> Two assumptions I will take on board in this paper, as I note below in §2.2 and §2.3.

expresses a truth, we look to see whether (4) is true at every accessible possible world. So, as with negation, the modal operator expressed by “necessarily” in (7) applies to (4), the proposition expressed by the embedded sentence.

The fictionalist operator seems to behave differently than negation and modal operators. Consider again:

(2) Sherlock is clever.

Since there is no such person, a Millian might say that (2) expresses the proposition:

(8)  $\langle \emptyset, \textit{cleverness} \rangle$ .<sup>5</sup>

Applying the fictionalist prefix to (2) produces:

(1) According to the fiction, Sherlock Holmes is clever.

But (1) does *not* express the proposition:

(9)  $\langle \textit{According to fiction } F, \langle \emptyset, \textit{cleverness} \rangle \rangle$ .

After all, (1) is supposed to express a truth, but (9) is *false*—the Conan Doyle stories do not attribute *cleverness* to nothingness.

There is an additional (and perhaps more compelling) reason for a Millian to reject (9) as the account of the fictionalist operator: the Millian might say that sentence:

(10) Pegasus is clever.

also expresses (8), but if the content of the prefixed:

(11) According to the fiction, Pegasus is clever.

---

<sup>5</sup> To clarify, I do not mean to claim that (2) expresses the proposition *the null set is clever*. (8) is meant to be a degenerate singular proposition, not the singular proposition attributing *cleverness* to the null set. If one were worried about how to represent the latter, we might do so with  $\langle \{\emptyset\}, \textit{cleverness} \rangle$ , while representing the content of (2) as  $\langle \emptyset, \textit{cleverness} \rangle$ . If we pursued this strategy, presumably we would want to reformulate the content of (3) as  $\langle \{\textit{Jack}\}, \textit{happiness} \rangle$ . For more on using the null set as a constituent of propositions in order to make sense of the content of sentences with empty names, see Everett (2000).

has the fictionalist operator applying to the proposition expressed by (10), then (1) and (11) would both express (9). That would be bad, since (11) is intuitively false while (1) is intuitively true.

We cannot resolve these problems by claiming that the content of (1) is:

(12)  $\langle$  According to fiction  $F$ ,  $\langle$  Sherlock, *cleverness*  $\rangle$   $\rangle$ .

If it were so easy for Sherlock to appear as a constituent of the proposition, we would just as easily be able to express the singular proposition  $\langle$  Sherlock, *cleverness*  $\rangle$  in asserting (2). If we want to say that (1) is true but (2) is false, we will need a more complicated understanding of the fictionalist prefix, a proposal on which the prefix forces us to reinterpret the semantic value of the embedded sentence.

So how *does* the fictionalist prefix behave? If a sentence like (1) is true (and intuitively it is), perhaps that is because (1) is telling us that “cleverness” applies to the referent of the name “Sherlock” *as used in the fiction*. Our ordinary uses of “Sherlock” are empty, but if the fiction turned out to be true the name would refer.

A sketch of the proposal I wish to advance (a full presentation will come in §5) is that the fictionalist prefix “According to the fiction” needs to be interpreted two-dimensionally in order to accurately depict its semantic behavior. In addition to *evaluating* the embedded sentence at the fiction, we are also designating the fiction *as actual*. To evaluate (1), then, we treat the fiction *as actual* and ask whether “Sherlock is clever” would be true at the fiction. If the fiction were actual, “Sherlock” would rigidly designate an individual, and at the fiction that individual is clever, so “Sherlock is clever” is true at the fiction considered as actual.<sup>6</sup> On my proposal, then, (1) is true in virtue of that fact; I will propose that the prefix “According to the fiction” is equivalent to “At the fiction considered as actual”.

## 1.2 Dialectic

A great deal here needs to be clarified. I will start in §2 by noting some of the assumptions made in setting up the problem. In §3 I will offer an explanation of two-dimensionality as I construe it, and in §4 I will offer at

<sup>6</sup> The term “considered as actual” comes from Davies & Humberstone (1980).

least a partial statement of the metaphysics I presuppose in applying two-dimensional modal semantics to discourse about fiction. Once I have thus shed some light on the background issues, I will present a more detailed account of the positive view in §5 and in subsequent sections address some objections.

## 2.1 Unprefixed Sentences Express Falsehoods

The first assumption made in formulating the problem is an intuition I am taking for granted—that the following two sentences are semantically nonequivalent:

- (1) According to the fiction, Sherlock Holmes is clever.
- (2) Sherlock Holmes is clever.

In particular, I want to say that (2) semantically expresses a false proposition, though I certainly want to maintain that (1) expresses a truth. I think (2) is false because there is no such person as Sherlock, and hence there is no such clever person.

Not everyone will be impressed by that argument—on some views,<sup>7</sup> “Sherlock” picks out an abstract object, and the verb picks out a relation other than instantiation (we might think of that relation as “being depicted as”). So there are ways to maintain that (2) semantically expresses a truth. But I suspect that if ordinary speakers are explicitly offered both (1) and (2) and asked whether both are true, they will not want to say that (2) is true. There is a tendency to *treat* (2) as true, but I suspect that is because we often presume that a speaker who utters (2) really meant to convey the proposition that (1) expresses; I think the inclination to say that (2) is true depends on conflating (1) and (2).

Of course we could find ways of taking (2) to literally express a truth. However, as a general methodological principle it seems best to me to take semantics at face value, and at face value (2) expresses a singular proposition attributing *cleverness* to an individual. If we want to appeal to speaker meaning or implicature to explain why (2) *merely seems* true we can certainly do so, but a cleaner account of the semantics of simple sentences makes it far easier to explain the semantics of complex sentences. If an

---

<sup>7</sup> See for instance van Inwagen (1977).

ordinary speaker thought that (2) were true, I think we could dissuade her by reminding her that Sherlock is not real, and so there is no such clever person. We would not be able to similarly dissuade someone who thinks that (1) is true. I am willing to thus accept that (1) literally expresses a truth but that (2) does not.

## 2.2 Fictional Names Purport to Refer to Concrete Individuals

The problem also arises only if we reject a descriptivist semantics for “Sherlock”. Like many, I take descriptivism to be discredited for names in general, and I find it desirable to give the same semantic story about fictional names that we endorse for names in general (a point to which I will return shortly). I am presupposing a directly referential semantics for names, though I will not present any arguments for it here.

More controversial than the rejection of descriptivism is my assumption that fictional names purport to refer to concrete individuals and not abstract objects. I have alleged that in (2), “Sherlock” fails to refer. But if fictional characters are actual abstract objects, then “Sherlock” has an extension even as used in (2)—the name simply picks out the abstract object. This view could be (and has been) developed in a variety of ways, so constructing an objection to the general strategy would be too large a project for me to attempt here. However, I should at least say something about why I do not find it a satisfying dissolution of the problem with which this paper began.

I assume that ordinary names like “Barack Obama” designate concrete individuals, not abstract objects. I think it desirable to say that fictional names and ordinary names have the same semantic function—they are all names. So I am tempted to say that if ordinary names purport to pick out concrete individuals, we should think that fictional names also purport to pick out concrete individuals. Fictional names *fail* to do so while typical ordinary names *succeed*, but they all *purport* to do the same thing—designate concreta.

I do not expect my opponent to find that persuasive; it is an explanation of my theoretical motivations, but not a decisive argument. I can, however, say a little more. I want to address an opponent who thinks that ordinary



names and fictional names are semantically distinct—she holds that the former purport to designate concreta, while the latter purport to designate abstracta. She might defend that distinction by observing that when an ordinary name is coined, the user is attempting to designate something that already existed independently of the baptism. But when a fictional name is coined, the user is not attempting to designate a pre-existing entity; we might say that in creating a work of fiction she is *creating* a referent, but she is not attempting to refer to something that existed prior to the creation of the fiction. There is thus a difference in the way that fictional names and ordinary names are introduced, and that might support the claim that the kinds of names have different semantic functions.

I want to respond as follows. I see a similarity between fictional names and the theoretical terms of false theories. I am attracted to a directly referential view of natural kind terms just as I favor a directly referential account of names (having been persuaded by Kripke that a descriptivist view is problematic in both cases). Given that, empty theoretical terms like “phlogiston” pose the same kind of problem to direct reference as fictional names like “Sherlock”. But “phlogiston” purports to designate a concrete substance, not an abstract substance, since the scientific theory of which it is a part was not developed as a work of fiction—it was offered as a theory about concrete physical events. An attempt to defend direct reference against the problems raised by “Sherlock” would, I hope, also defend that view against the problems raised by “phlogiston”—a single solution that solves both empty term problems is more appealing on theoretical grounds. And so the suggestion that fictional names successfully refer to abstract objects strikes me as a dissatisfying solution to the empty name problem.

Admittedly, someone might have independent reasons for endorsing the view that fictional names designate abstract objects—I have construed the view as motivated by a desire to say why fictional names do not threaten direct reference, but one might have other reasons for favoring this approach to fictional names. I will not pursue the matter further here, however; at this stage, I merely want to indicate why my setup of the initial problem should not be straightforwardly dismissed, and I hope that what I have said in this section at least shows that we should *explore* solutions to the problem that maintain that fictional names purport to refer to concrete individuals.

### 2.3 Propositions are Structured Abstract Objects

In setting up the problem, I assumed that simple name-predicate sentences express Russellian propositions. “Obama is happy” expresses a structured proposition that has that guy and the property of being happy as constituents. There are of course other accounts of propositions; one might prefer a more minimalist account on which propositions are mere functions from possible worlds to truth-values, with no commitment to abstract objects that have concrete individuals as constituents. Here again I do not want to focus on defending a particular view of propositions for fear of diverting attention from my primary area of interest; I will simply note that like many I think Russellian propositions offer notable advantages over the minimal account of propositions as mere functions. Structured propositions allow for an easy explanation of the nonequivalence of necessarily true propositions; they permit an attractive explanation of compositionality, since complex expressions have their simpler parts as constituents; and they allow for the distinction between direct reference and rigid descriptions, since we can ask whether the individual herself is a constituent of the proposition. Observing these advantages, I will simply take the Russellian view of propositions for granted.

### 2.4 Plantingan Existentialism

One might concede that fictional names purport to refer to concrete individuals and endorse the Russellian view of propositions, yet still think that the initial problem is based on a misconception. I claimed that:

(2) Sherlock Holmes is clever.

expresses a degenerate singular proposition, one lacking any individual corresponding to the name. Someone might reply that (2) expresses a *non*-degenerate singular proposition—it attributes *cleverness* to an individual, though admittedly that individual does not (actually) exist.

If one held such a view, then (2) would express a false proposition, but it would not contain an empty name. It would attribute *cleverness* to a man; the proposition is false, since the man is not in fact clever. (How could he be, since he does not exist?) (1) still expresses a truth, since in the fiction that man does exist and is clever, and thus (1) and (2) would

remain nonequivalent. But the problem of empty names would not arise, since in both propositions “Sherlock” picks out an individual. The prefix “According to the fiction” would thus function no differently than an ordinary world-shifting prefix such as “At world  $\beta$ ”, which we can make sense of using ordinary one-dimensional modal semantics. In other words, on this alternative view the prefix used in (1) would apply directly to the proposition expressed by (2), so the prefix would function like ordinary modal operators.

I find this strategy deeply counterintuitive. Sherlock does not exist. There is no such person. Singular propositions contain persons as constituents, so if there is no such person, there is no such singular proposition. If no such singular proposition exists, how could we be expressing it in uttering a sentence? The key assumption I rely on here is a thesis Plantinga dubs “existentialism”<sup>8</sup>—there is a singular proposition about  $X$  *only if*  $X$  exists. Not everyone finds this thesis compelling, and I will not attempt to persuade those people here, except to say that if one thinks that singular propositions have the relevant individual as a constituent, it seems to me intuitively compelling to think that the singular proposition depends for its existence on the existence of the constituent individual.

Having specified the major assumptions involved in setting up the problem, I now turn to an explanation of two-dimensional semantics in order to subsequently explain how a two-dimensionalist account can solve the problem. Readers who are already familiar with two-dimensionalism will be able to skim §3, though some of what I will say concerns non-canonical claims about the relationship between the two dimensions and what falls under the category of “semantic content”.

### 3 Two-Dimensional Semantics

#### 3.1 What Should Count as Two-Dimensional?

Orthodox modal semantics dictates that expressions are evaluated with respect to a parameter; possible worlds provide the values for that parameter. An expression has an extension relative to a world, and if we change the world of evaluation, the expression may or may not be mapped to

---

<sup>8</sup> See Plantinga (1979).

a different extension. I am proposing that an expression counts as two-dimensional when the following conditions hold: first, the evaluation of a two-dimensional expression requires the specification of *two* parameters. Second, a two-dimensional expression is non-atomic; it can be rendered as an operator applying to a one-dimensional expression (or to another two-dimensional expression).<sup>9</sup> Third, the operator functions by shifting one or both parameter-values in the double-index in ways that differ from the pattern of parameter-shifting familiar with one-dimensional operators. It will be helpful to give a few examples; let me stress that the *second* parameter in the double-index represents the parameter usually invoked in one-dimensional modal semantics, and the *first* parameter is the addition. For one-dimensional operators, the first parameter (whose value is “ $\alpha$ ” in the table) has no substantive role, and could be omitted.

### One-dimensional Operators:

*Possibly*  $S$  at  $(\alpha, \beta)$  iff for some world  $\gamma$  such that  $R\beta\gamma : S$  at  $(\alpha, \gamma)$ .

*Necessarily*  $S$  at  $(\alpha, \beta)$  iff for every world  $\gamma$  such that  $R\beta\gamma : S$  at  $(\alpha, \gamma)$ .

*At world*  $\delta$ ,  $S$  at  $(\alpha, \beta)$  iff  $S$  at  $(\alpha, \delta)$ .

In all of these cases, the only shifting occurs in the second parameter-value, and further, the first parameter-value does not determine the shifted value of the second parameter. We could omit the first parameter on both sides without compromising the explanation of one-dimensional operators.

---

<sup>9</sup> Two-dimensionalists do not always present their view in terms of operators, and this might make one suspicious of my characterization of two-dimensionalism (and indexicality) in those terms. But while two-dimensionalist views *are not* presented in terms of operators, they *could* be formulated as such without distortion. Suppose one wants to say that “water” is a two-dimensional term, such that if our world is actual, “water” rigidly designates H<sub>2</sub>O, but if Twin Earth is actual, “water” rigidly designates XYZ. That view might be reformulated as the view that our term “water” is a complex expression with an operator applying to a non-two-dimensional description (perhaps the description is deferential, or perhaps it can be captured as “the transparent liquid ...”). The operator rigidifies that description with respect to some world designated as actual. Hence if Earth is actual, “water” rigidly designates the Earth-satisfier of the non-two-dimensional description, and *mutatis mutandis* if Twin Earth is actual. Formulating two-dimensionalism in terms of operators and double-indices is compatible with common invocations of two-dimensionalism, has the added benefit of casting the net wide enough to encompass views like Stalnaker’s, and additionally helps to clarify the relationship between indexicality and two-dimensionalism in general.

**Two-dimensional Operators:**

According to fiction  $F$ ,  $S$  at  $(\alpha, \beta)$                     iff         $S$  at  $(F, F)$ .  
 At world  $\delta$  considered as actual,  $S$  at  $(\alpha, \beta)$     iff         $S$  at  $(\delta, \delta)$ .

What is distinctive about the shifting here is that *both* parameter-values are shifted. One-dimensional operators only shift the parameter-value representing the point of evaluation. These operators, however, are two-dimensional because the parameter-shifting cannot be explained simply by appealing to the second parameter-value.<sup>10</sup>

The distinction between considering a world as actual and considering a world as counterfactual is one of the hallmarks of two-dimensionalist semantics. The simplest explanation of these technical terms can be given by relating these terms to the doubly-indexed modal semantics invoked in the examples of operators just given. Let “ $\alpha$ ” designate the world we inhabit. When a philosopher asks whether one of our ordinary utterances is true, without any kind of qualification attached to “true”, we are typically evaluating that utterance at the double-index  $(\alpha, \alpha)$ —we are relying on the commonsense fact that  $\alpha$  is the actual world, and we are asking whether the proposition generated by the utterance at our world  $\alpha$  is true of our world.

When the philosopher asks instead whether the utterance says something true *of some world*  $\beta$ , however, we are evaluating the utterance at the double index  $(\alpha, \beta)$ . We continue to accept that  $\alpha$  is the actual world, and use that world to generate the proposition that is expressed by the utterance. But once that proposition has been generated, we can then de-

---

<sup>10</sup> Kaplan’s treatment of indexicals is another use of the two-dimensional machinery. The two parameters in this case depict a context of utterance and a point of evaluation, respectively. If we think of “I” as an operator  $O$  applied to the definite description “the speaker”, then Kaplan’s theory tells us that the extension of  $O$  (*the speaker*) when evaluated at  $(\alpha, \beta)$  is the extension of *the speaker* when evaluated at  $(\alpha, \alpha)$ . That is, no matter what point of evaluation we use to evaluate a use of “I”, we find the extension by going back to the context of utterance to find the extension of “the speaker”. Since the shift in the value of the second parameter is determined as a function of the first parameter-value, an explanation of the parameter-shifting will have to invoke the first parameter, unlike any one-dimensional operator. Kaplan does not present his theory in terms of operators, but this formulation is equivalent to his in explaining how the extension of an indexical is determined in modal contexts. Of course it is important to note that Kaplan will demand that construing “I” as a complex of an operator and a definite description does not illuminate the *semantic content* of the indexical. This characterization of two-dimensionalism deals only with extensions, not with content. Two-dimensionalism by itself is entirely neutral about what should count as semantic content and what should not.

termine whether that proposition is true or false of any world we like. So in this case we evaluate the generated proposition at world  $\beta$  instead of at  $\alpha$ . This is what it means to “consider  $\beta$  as counterfactual”.

It is common for a two-dimensionalist to “consider world  $\beta$  as *actual*”. When we do this, we are roughly asking whether someone who inhabits world  $\beta$  is expressing a truth. To put the point more colorfully, we imagine that  $\beta$  turns out to be the world we inhabit—we imagine that  $\beta$  is the actualized world. We then find the proposition that is generated by the  $\beta$ -world utterance (an utterance we are now imagining to be ours), which may well differ from the proposition that  $\alpha$ -denizens express. Once we have this  $\beta$ -world proposition, we can ask whether that proposition is true of world  $\beta$  in order to assess whether  $\beta$ -utterances would be uttering truths. The double-index corresponding to considering  $\beta$  as actual is  $(\beta, \beta)$ .

Finally, to *designate*  $\beta$  as actual is to imagine that  $\beta$  is the actualized world without yet committing to any particular world of evaluation, which we could represent as  $(\beta, \_)$ . We identify the proposition generated by the utterance in  $\beta$ , and we can then evaluate that proposition at any world we choose. (Should we evaluate the proposition at  $\beta$ , we are considering  $\beta$  as actual, but we can in addition evaluate the proposition at other worlds, in which case we are *designating*  $\beta$  as actual but not *considering*  $\beta$  as actual.)

The three technical terms thus correspond to the following specifications of parameter-values:

Evaluating $S$ at $\beta$ considered as counterfactual	= Evaluating $S$ at $(\alpha, \beta)$
Evaluating $S$ at $\beta$ considered as actual	= Evaluating $S$ at $(\alpha, \beta)$
Designating $\beta$ as actual	= Evaluating $S$ at $(\beta, \_)$

A vacant place has been left in the final case because designating  $\beta$  as actual does not by itself commit us to any particular value for the second parameter, since no world of evaluation has yet been specified. One-dimensional modal semantics handles the case of considering a world as counterfactual, but two-dimensional semantics is necessary to account for designating or considering a world as actual.

### 3.2 Content and Semantics

Different implementations of the two-dimensional machinery make different assumptions about what counts as semantic. In Kaplan (1989)'s framework, the first dimension represents something like semantic meaning, and the second dimension represents semantic content. In some more recent developments<sup>11</sup> of the two-dimensional machinery, both dimensions are equally semantic. Stalnaker (2001) characterizes the second dimension as semantic but the first dimension as *meta*-semantic.<sup>12</sup>

The lesson I want to draw from this is that one can employ the two-dimensionalist framework without thereby being committed to any view about whether the first dimension captures something semantic or not. In this paper I am suggesting that the Millian can invoke two-dimensionalism to help with the problem of fictional names; in advancing that suggestion I must maintain that semantic *content* is captured solely by the second dimension, as is the case in Kaplan's two-dimensionalist treatment of indexicals. A two-dimensionalist about "water" may say that the first dimension captures the sort of descriptive information about water that any ordinary speaker understands—water is the transparent liquid in our lakes, for instance. But that may be compatible with Millianism, so long as we hold that the first dimension does not capture the semantic content of "water"; the Millian might say instead that the first dimension captures something *pre-semantic*<sup>13</sup> in that it captures the reference-fixing conditions that a substance must satisfy to count as the referent of "water" as we use the term.<sup>14</sup> If it turned out that we inhabited a world with XYZ instead of H<sub>2</sub>O, then it would be XYZ that satisfied the reference-fixing conditions associated with "water", and thus our utterance of "water is wet" would be about XYZ. It is that fact that the two-dimensionalist wants to address in saying that "water is wet" is true of an XYZ world considered as actual (though she will nevertheless admit that "water is wet" is probably false of an XYZ world considered as counterfactual, since there is no H<sub>2</sub>O to be wet in such a world). That point about two-dimensionalism is perfectly compatible with putting a Millian spin on the two dimensions—we can

---

<sup>11</sup> See for instance Jackson (1998).

<sup>12</sup> For a helpful discussion of the various ways of interpreting the two-dimensional framework, see Gendler & Hawthorne (2002, pp. 45–53).

<sup>13</sup> I owe this term to André Gallois.

<sup>14</sup> For a suggestion along these lines, see Gendler & Hawthorne (2002, p. 53, note 116).

simply specify that the first dimension captures something pre-semantic, and only the second dimension captures semantic content.

## **4 Modal Metaphysics**

In this section I want to specify some assumptions about the ontology of possible worlds. I will not try to argue for an ontology here, as the two-dimensional view of the fictionalist prefix does not depend on any one modal ontology, but I do not want to unduly confuse the reader who might have different assumptions about the ontology of possible worlds.

### **4.1 Actualism**

Like many who invoke possible worlds, I do not want to be committed to the Lewisian pluriverse of concrete worlds, but I find the language of possible worlds useful. It is not always clear whether a possible world is assumed to be a concrete thing or an abstract object. If possible worlds are concrete entities, the typical actualist presumably rejects a plurality of possible worlds; she may instead believe in “ersatz worlds”, actual abstract objects that represent ways that things might have been. If possible worlds are abstract objects, the typical actualist may well endorse a plurality of possible worlds—those actual abstract objects that represent possibilities are themselves the possible worlds. In the interest of disambiguation, I will use the word “universe” to designate the sort of large concrete entity of which we are a part—a universe is the collection of real concrete and abstract objects. Lewis’s view could thus be described as a plurality of universes. I will use the term “ersatz world” to characterize an actual abstract object that represents a way the universe might have been. I am an actualist in that I believe in only one universe, though I believe in a plurality of ersatz worlds that represent the many different ways the universe might have been.

### **4.2 Ersatz Worlds**

I will suppose a linguistic ersatzism according to which an ersatz world is a set of propositions. Roughly speaking, an ersatz world is the set of



propositions that would be true had that ersatz world been actualized, i.e., had the universe turned out such that the ersatz world in question was an accurate representation of reality. I say that this is rough because there will be a familiar problem with alien singular propositions. Suppose there could have been some simple individual who does not in fact exist.<sup>15</sup> Then there are not in fact any singular propositions about that individual. For such singular propositions to exist, the individual herself would have to exist. If there are no such actual singular propositions, then no such singular propositions are members of the set of propositions that constitutes the ersatz world. However, if that ersatz world were actualized, there would be such true singular propositions. So if the ersatz world were actualized, there would be certain true propositions that are not among the set of propositions that constitute the ersatz world. This result obtains because we have assumed that there could have been simple individuals that do not in fact exist, and we have assumed that in the absence of such individuals, there are in fact no propositions that contain such individuals as constituents.

## 5 The Positive View

Having made our way through the background, I now return to the main task: offering a detailed proposal of the semantics of an  $\alpha$ -denizen's utterance of:

- (1) According to the fiction, Sherlock Holmes is clever.

I have claimed that our utterances of (1) are true (of our world) just in case:

- (2) Sherlock is clever

is true at the fiction considered as actual. But while an account of truth-conditions is interesting, I have said little about the semantic content we express in uttering (1). I turn to that now.

---

<sup>15</sup> I say "simple" to rule out a case in which an individual is an unactualized combination of actual parts. The problem arises when we have an alien *simple*.

## 5.1 Semantic Content and Two-Dimensional Operators

Propositions are generated by combining something pre-propositional (reference-fixers) with the actual world. (2) has some associated complex of reference-fixers such that were the fiction actualized, an utterance of that sentence would generate a singular proposition. What the  $\alpha$ -denizen is saying in uttering (1) is that such a proposition (the proposition generated by combining the reference-fixer with the world of the fiction  $F$ ) would be true when evaluated at  $F$ .

The prefix “according to the fiction” thus has two effects. First, it tells us to designate the fiction as actual; this helps to determine what sort of proposition we should evaluate. Second, it tells us to use the fiction as the point of evaluation—we are interested in knowing whether the right sort of individual is clever *at the world of the fiction*. This dual shift is typical of considering a world as actual, as defined in §3.1. When an  $\alpha$ -denizen *designates*  $F$  as actual, she is shifting the first parameter-value of the double index from  $\alpha$  to  $F$ . This leaves open which world of evaluation should be used. The  $\alpha$ -denizen is hypothesizing that  $F$  is actual for the purposes of seeing what would follow. When an  $\alpha$ -denizen *considers*  $F$  as actual, however, she is shifting both parameter-values of the double index to  $F$ . To consider  $F$  as actual is *both* to designate  $F$  as actual and to use  $F$  as the world of evaluation.

A non-degenerate singular proposition results from combining the reference-fixer of “Sherlock is clever” with the fiction  $F$  designated as actual. The content of an  $\alpha$ -utterance of (1) cannot simply be a non-degenerate singular proposition, given the absence of any individual answering to the name, but it might contain as constituents the reference-fixer and the fiction  $F$  (*qua* world designated as actual), with the instructions that they are to be combined in the relevant way.

An analogy may help; consider two contents, both of which determine the number 4 as an extension:

$$(13) \langle 4 \rangle .$$

$$(14) \langle \text{product}, \langle 2, 2 \rangle \rangle .$$

Suppose that there were some domain containing 2 but not 4 (though naturally this supposition will not be plausible as a characterization of the

domain of any possible world). Then (13) does not exist at that domain, but (14) does.

If a speaker wants to say something counterfactual about the number 4 (such as that it exists at  $\delta$ ), she can do so by attributing a modal property to (14) even though she cannot express (13) or attribute modal properties to (13) directly. Similarly, an  $\alpha$ -denizen cannot refer directly to Sherlock, but she can talk about him indirectly by appealing to the  $F$ -satisfier of the reference-fixer of “Sherlock”.

The proposal I wish to advance is that the  $\alpha$ -denizen accomplishes exactly this by uttering:

(1) According to the fiction, Sherlock Holmes is clever.

The content of (1) includes the reference-fixer of “Sherlock is clever” (which will presumably capture the reference-fixers of “Sherlock” and “clever” together with the structural relations specifying that substance satisfying the reference-fixer of “Sherlock” exemplifies the property satisfying the reference-fixer of “clever”), and the instructions that this complex reference-fixer should be combined with the fiction  $F$  to generate something truth-evaluable. The content of (1) will not be a singular proposition; just as (14) determines the extension 4 without including 4 as a constituent, so the  $\alpha$ -denizen can express a nonsingular proposition that is true in virtue of facts about what a singular Sherlock-proposition would be like even though in fact there is no such singular proposition.

How exactly should we depict the content of (1)? There are some hazards in trying to be precise here, since I do not want to assume too much about the nature of reference-fixers, particularly whether or not the reference-fixers of terms are adequately captured with definite descriptions. I am inclined to say that the content of (1) has two constituents. The first constituent captures the reference-fixer of the embedded clause “Sherlock is clever”; this reference-fixer will presumably feature the reference-fixers of the terms “Sherlock” and “clever” with the right structural relationship. The second constituent of the content of (1) will specify which world is being considered as actual.

Supposing that we characterize the content of (1) in this way, should the Millian feel vindicated? One might think not—after all, on this construal, no individual is a constituent of (1). However, I think the Millian can accept this result. Names have referential contents and *pre-semantic*

reference-fixers that determine those contents. Operators that shift the world designated as actual produce strange effects, but those effects are not attributable to the names, they are due to the operators. So the fact that the name in (1) does not function referentially does not reflect on the semantic contribution of the name; it is a result of the unusual operator governing the clause in which the name occurs. In some respects, then, this proposal resembles the Fregean view that propositional attitude verbs trigger the indirect sense of the embedded terms.

One might well prefer *ceteris paribus* a theory that does not require reference-fixers to show up in the proposition on *any* occasion—it is admittedly a cost of this theory that the content of (1) includes reference-fixing conditions, since those conditions are typically pre-semantic rather than semantic. But it is hard to see how an  $\alpha$ -denizen can speak truly in uttering (1) without accepting *something* counterintuitive. We can reason about the consequences of some other world(s) being actual, and to make sense of such talk we are forced into some surprising conclusions about semantics. The worry concerning this point deserves attention, though, so I will consider it with more care.

I suggest that in simple name-predicate sentences, we can accept the Millian semantics for the name. But when the name is embedded under an operator that shifts the world designated as actual, including operators specifying that a world or fiction is *considered* as actual, the name ultimately is not functioning directly referentially. The fictionalist prefix is such an operator, so the name in (1) does not purport to refer directly.

Such a proposal might be interpreted in a number of ways. We might classify it as a lexical ambiguity, saying that some utterances of “Sherlock” attempt unsuccessfully to directly refer, while other uses are not even purportedly directly referential. We might treat the proposed view as a semantic pluralism by saying that there are multiple elements that deserve to count as meanings of a name—the extension and the reference-fixing description—and they are all present in some way for any utterance featuring the name. But I prefer a formulation that evokes Frege, at least in one respect—I want to treat the prefix “According to the fiction” as triggering an oblique context such that embedded terms express something other than their customary content.

I think it would be a distortion to classify such a view as an ambiguity theory or as a pluralism. According to Frege, in a propositional attitude

context, the referent of a name is its customary sense, but we should not claim that Frege holds that names are lexically ambiguous. On his view, all unembedded occurrences of the name will feature the same sense and reference, and all embedded occurrences will take the customary sense as their reference. The Fregean view is quite unlike a typical case of lexical ambiguity; first, the place of the name within the sentence's syntactic structure dispels any uncertainty about whether the name has its customary or indirect reference. Second, the indirect reference is determined by the customary sense, so there is a relationship between the indirect and customary semantic values. Contrast this with lexical ambiguity: the syntactic structure of "Alice likes bats" does nothing to tell us what "bat" means in this utterance, and neither candidate meaning is determined as a function of the other.

It is a little more plausible to think of Frege as a pluralist, but there is an important distinction to be made. Frege associates a sense and a reference with an utterance of a term, and in that way we might say that there are two elements that could deserve to be classified as semantic values. But the claim that sense and reference are both semantic values of an unembedded occurrence of a term is distinct from the claim that the reference in embedded contexts is the term's customary sense. We might call Frege a pluralist in virtue of his view that every occurrence of a term has both a sense and a reference.<sup>16</sup> But we should not call Frege a pluralist simply in virtue of his view that the indirect sense differs from the customary sense, since he would not claim that a particular occurrence of the term has both the customary sense and the indirect sense.

Thus I think a Fregean view about oblique contexts is best distinguished from ambiguity or pluralist views, and I want to treat prefixes that shift the world designated as actual as triggering a kind of Fregean oblique context. The content of unembedded occurrences of names is accurately captured by the direct reference theory, I will suppose. But the content of a name within the scope of "According to the fiction" is something other than the customary content, not because there is something semantically strange about names, but because there is something semantically strange about the prefix. There are many differences between the view I am proposing and Frege's—for instance, I am using a semantics

---

<sup>16</sup> Or we might not—I have no stake in the interpretive question whether a distinction between sense and reference *ipso facto* counts as a kind of semantic pluralism.

that explains content in terms of structured propositions, with no intent to draw a sense/reference distinction. Furthermore, while Frege thought that propositional attitude contexts triggered indirect sense and reference, I am limiting the appeal to oblique contexts to those that shift the world designated as actual. Nevertheless, I find the comparison with Frege useful in explaining how to square my general sympathies toward the Russellian intuitions about names with the proposal being advanced here, particularly in motivating my claim that the proposal should not be treated as a lexical ambiguity theory about names.

One might have qualms about my proposal—where does the non-customary content come from? I suggested that when embedded under a two-dimensional operator, the content of the name features a function of the reference-fixing condition as opposed to the Russellian content we would expect in customary contexts. But if the customary semantic content of such terms is simply the extension, the reference-fixing condition *determines* the semantic content without itself being *part* of the semantic content. So how does the reference-fixing condition of the natural kind term factor into the semantic content of the sentence containing the two-dimensional operator, if in general it is not part of the semantics of the name?

In my estimation, this question exposes the greatest drawback of the account I favor. On a lexical ambiguity approach, one could say that the semantics of a natural kind term which is embedded in a two-dimensional context differs from its semantics in an unembedded context. There would be no reason to suppose that the semantic content in a specialized context is a function of the semantic content in other contexts, since in general the candidate contents of lexically ambiguous terms are entirely independent. On a pluralist approach, the semantics in embedded contexts *is* a function of the semantic content in other contexts—there are multiple contents in all contexts, and for whatever reason different contexts make one or another content more salient.

The Fregean-style approach is the only one of these options forced to say that the reference-fixing condition is typically non-semantic, but that it can enter into the semantic content through the presence of an operator governing the name. However, it should not be a devastating concession to accept that the two-dimensional operator might have this effect, because direct quotational devices might be taken to behave in a similar way.

The expression “snow” is not part of the semantic content of the sentence “snow is white”, but it is part of semantic content of the sentence “‘snow is white’ has three words”. If we can treat the quotational device as a function applying to the clause “snow is white”, then it behaves by taking something pre-semantic from that clause and bringing it into the semantic content of “‘snow is white’ has three words”. So quotational devices may establish a precedent for a linguistic feature that brings a customarily pre-semantic element into the domain of semantics.

In addition, we can observe the familiar problems with the alternatives. We have good reason to favor a referential semantics for unembedded names, but we know that the problem of empty names will require some revisionary move to explain how we can speak truly about things that do not actually exist. The account proposed here is simply one way of meeting the burden of that revisionary account—if we grant that the two-dimensional operator can bring the reference-fixing condition of names and natural kind terms into the semantic content of the complex sentence, an attractive picture emerges.

## **6 Against Modal Approaches to Fictional Content?**

The fact that fictions may be incomplete and/or inconsistent has led many to think that we should abandon a possible-worlds approach to fictional content. But while there are challenges to such an approach, there are good reasons to pursue it, and obstacles to developing an alternative account.

One advantage stems from the fact that the possible-worlds framework is now entrenched in explaining truth for nonfiction. If we could adapt that framework to account for fictional truth,<sup>17</sup> we would have an appealing uniformity in our semantics. A uniform treatment will be helpful if the line between fictional and non-fictional claims is sometimes blurry, as it may be for theoretical claims. False theories aim to be nonfiction, but they turn out to be false (and sometimes impossible) stories about the world. If we are going to use possible worlds to account for the truth and falsity of theoretical statements, then we will have to reconcile modal semantics

---

<sup>17</sup> As Lewis (1978) does.

and impossible statements—why not at that point make use of that reconciliation for impossible fictions as well? We might also be motivated by a desire to account for historical fictions, in which there is a blending of fact and fiction—a unified account of truth that applies to both fiction and nonfiction would be desirable here.

Alternatives to the worlds-based approach will struggle to achieve this kind of uniformity. Consider first an *abstracta* approach to fictional content, according to which the content of a fiction is an abstract object, and fictional characters are themselves abstract objects.<sup>18</sup> On such a view, “Sherlock” and “Santa” designate abstract objects that inhabit our world, as opposed to the names purporting to designate merely possible (or merely impossible) concrete individuals.<sup>19</sup> Whatever the merits of such an approach, its proponent has to admit that there are two different explanations for the truth of a sentence of the form “X is *F*”. “Brendan is happy” is true just in case Brendan exemplifies *happiness*, but “Sherlock is clever” *isn’t* true just in case Sherlock exemplifies *cleverness*, because the abstract object is the wrong type of entity to exemplify such a property. We can make this clearer by considering “Sherlock is a concrete individual”. According to the fiction, Sherlock *is* concrete, but if “Sherlock” designates an abstract object, the sentence says something false. The *abstracta* approach needs some way to distinguish real property-exemplification from fictional property-exemplification, such as positing an otherwise-unmotivated ambiguity in “is”; the worlds-based approach avoids the need for such a move.

A similar point arises when we consider a pretense approach to fictional content. On this approach, “Sherlock is clever” does not express a truth, but we may *pretend* that it does (perhaps by pretending that the sentence expresses a Russellian proposition in the first place). But while this approach has its appeal, we are still owed an account of just what content we are pretending that the sentence expresses. If we are just pretending that some alien proposition exists, then the view collapses into the possible-worlds approach with an actualist interpretation. So whatever the

<sup>18</sup> I have previously noted van Inwagen (1977) as an instance of this strategy; see also Kripke (1973) and some of the papers cited in Everett & Hofweber (2000, p. xv, note 6).

<sup>19</sup> There is a Meinongian flavor to the claim that Sherlock is a merely possible/impossible concrete individual, but that flavor should not be taken seriously. All I mean is that there is in fact no such individual as Sherlock, but if there were it would be a concrete individual.



benefits of a pretense approach, it does not promise a substantially different treatment of fictional *content*, though perhaps it does offer a different account of fictional truth (viz., “Sherlock is clever” is fictionally true because we pretend that it says something true, as opposed to the explanation in terms of the sentence expressing at the fiction a proposition that is true when evaluated at the fiction).

Thus far in this section I have attempted to point out an advantage of the possible-worlds-based approach as a way of deflecting one source of worry about the two-dimensional account of fictional truths. But it is now time to address specifically the problems we would incur due to the incompleteness and inconsistency of some fictions.

## 6.1 Incompleteness

Fictions are typically incomplete—there are propositions whose truth-value is undetermined by the content of the fiction. Since possible worlds are complete, it seems that in general a fiction cannot be associated with a single possible world. One solution to this problem is to associate a fiction with a *set* of worlds whose members are all and only the worlds compatible with the content of the fiction.<sup>20</sup> When we ask whether a proposition is true at the fiction, we will supervaluate—a proposition is true at the fiction iff it is true at every world in the set. If the proposition is true at some worlds in the set but false at others, the proposition is indeterminate at the fiction.

This provides instructions on *evaluating propositions* at incomplete fictions; to consider a fiction as actual, however, we also have to designate the fiction as actual in order to determine what proposition is *generated* by a sentence. How does a sentence generate a proposition when an incomplete fiction is designated as actual? Similar remarks apply here. If a sentence generates the same proposition at every world in the set, then that proposition is part of the content of the fiction. If a sentence generates proposition *p* at one world in the set, but *q* at another (where *p* and *q* are non-identical), then it is indeterminate whether the proposition is part of the content of the fiction. It will be indeterminate whether a certain individual is an element of the domain of the fiction if two worlds in the set

---

<sup>20</sup> Lewis (1978) is well-known for advancing this suggestion.

have different domains. In such a case, it will be indeterminate whether denizens of the fiction could generate singular propositions about such an individual.

## 6.2 Inconsistency

Fictions may also be impossible—for instance, in the Sherlock stories Watson has one wound, but the stories conflict regarding the location of that wound.

We might argue here that it is merely indeterminate according to the fiction where Watson's wound is (or whether he has one or two wounds, perhaps), and avoid the need to say that the fiction is genuinely inconsistent. But there is a temptation to say that when a proposition is explicitly expressed by the fiction, we should say that the proposition is true according to the story. This forces us to admit that some stories are inconsistent, but we would be hard-pressed to avoid this result, because of stories such as "Sylvan's Box",<sup>21</sup> which prominently features an empty box with something in it. We should not claim that it is simply indeterminate whether the box in that story is empty or has something in it. The point of the story is that some impossible situation obtains, and while we may not be able to understand what an empty box with something inside would be like, we certainly know that the fiction posits such a thing. So we should accept that fictions can be impossible.

The hope that fictions can be explained simply in terms of possible worlds seems to fade with this admission. But we have independent motivations for positing *impossible* worlds, and we might well appeal to them here in saying that fictions are associated with sets of possible *and/or impossible* worlds.

Some may find impossible worlds distasteful, but I think the case against them is not compelling. If impossible worlds were thought to be impossible *universes*, they would certainly be worrisome, but if we are merely positing impossible *ersatz* worlds, they are not ontologically repugnant. If an ersatz possible world is a consistent set of propositions, then an ersatz impossible world is an inconsistent set of propositions—there is nothing ontologically problematic here. One might be worried

---

<sup>21</sup> See Priest (1997).

about how to understand truth at an impossible world, but it seems we have a need to make sense of this independently of a modal understanding of fiction. For instance, there are counterpossible conditionals that we can evaluate, and intuitively they are not all true. Consider:

(15) If  $2 + 2$  equaled 5, then at least one of the Peano axioms would be false.

(16) If  $2 + 2$  equaled 5, then cows would fly.

(15) is intuitively true, but (16) is not. So we should not say that all counterpossible conditionals are trivially true. If we want to maintain the standard treatment of counterfactual conditionals in terms of world-based semantics, it is tempting to posit impossible worlds here. Similarly, we may be motivated to posit impossible worlds in considering counterlogical conditionals, if we assume that one logical system is true and the others are false.

With these motivations, there is no reason to think that impossible worlds will obey classical logic, so we should not expect that every proposition is true when evaluated at an impossible world. Thus we might well think that there is an impossible world at which the propositions explicitly expressed by the Sherlock fiction are true, but at which other propositions may be false. The worlds associated with the Sherlock fiction will be those worlds (all of which will be impossible, given the explicit inconsistency concerning Watson's wound) at which every proposition explicitly expressed by the Sherlock fiction is true.

With these strategies, we have some instructions on how to consider a fiction (whether impossible and/or incomplete) as actual.

## 7 Why are Fictional Names Empty?

In general, when  $\alpha$ -denizens use the name "Sherlock" outside of the scope of a fictionalist operator, I have claimed the name expresses a degenerate content, and thus fails to determine an extension no matter which world is the world of evaluation. I think the name is empty because there is no individual in the domain of  $\alpha$  that satisfies the reference-fixing conditions on the name, but others have suggested different reasons for thinking that

fictional names are empty; it will be worthwhile to look at some of these reasons.

## 7.1 Access

One motivation for taking fictional names to be empty will not trouble us here: some think that for a name to refer, the user must have some sort of access to the referent. The precise nature of this access has to be specified, of course; one candidate is some sort of acquaintance relation or connection by historical chain to a user with an acquaintance relation to the referent. For present purposes, we can sidestep worries about specifying the relationship required for reference, since the view I am defending is not committed to reference without access— $\alpha$ -denizens do not assert singular propositions about alien individuals, even when they invoke operators that shift the world designated as actual.

## 7.2 Individuation

Fictional names frequently present a much more serious problem—the stories do not specify a sufficient condition for being Sherlock. Suppose that a sufficient condition for being a particular individual is a matter of having a certain DNA sequence, along with further facts.<sup>22</sup> Conan Doyle's stories never tell us what Sherlock's DNA sequence is. If we label a person with one particular DNA sequence "X" and a person with a different particular DNA sequence "Y", then X and Y are not identical, yet each may be a candidate for being Sherlock. It is indeterminate whether Sherlock is X or Y; so there is no one individual, actual *or* alien, who would count as Sherlock were he to exist.

This suggests that the fictional name "Sherlock" fails to refer *regardless* of which world is designated as actual. Even if a world at which X plays the Sherlock role is actual, "Sherlock" does not seem to refer to X, since Y could have played the Sherlock role, and necessarily X and Y are non-identical. I suggested that "Sherlock" *would* refer if the fiction turns out to be actual, so my view is incompatible with the view that "Sherlock" is empty in virtue of failing to individuate a referent.

---

<sup>22</sup> The further facts will be invoked to distinguish identical twins, of course, but we needn't worry here about just what those further facts are.

The problem arises because fictions are incomplete. Some facts are left unspecified by the fiction, and it is implausible to think that those further facts are correctly specifiable in only one way. The incompleteness of fictions is one significant obstacle to giving an account of fictional content in terms of possible worlds. This is an issue that merits its own discussion, so I will return to it in §8.2, but to anticipate that discussion a little, one way of handling the problem is to account for fictional content in terms of *sets* of worlds. Suppose that world  $w_1$  makes true all the claims of the fiction, and further has it that  $X$  is Sherlock, while  $w_2$  makes true all the claims of the fiction, and further has it that  $Y$  is Sherlock. The content of the fiction is associated with a set of worlds, where both  $w_1$  and  $w_2$  are elements of that set. The claims that are true according to the fiction are those claims that are true at every element of the set. Claims that are true at some elements and false at others, such as the claim that Sherlock is  $X$ , are neither true nor false according to the fiction—it is indeterminate whether Sherlock is  $X$  according to the fiction.

That tells us something about how to handle incompleteness when *evaluating* propositions for truth-value, but my view now requires that we can *designate* a set of worlds as actual. What does that mean? I suggest that to designate a set of worlds as actual is to designate some unspecified element of that set as actual—we suppose that some one of a number of worlds turns out to be actualized. There are now two kinds of indeterminacy resulting from the incompleteness of fictions. When we *evaluate* the claim that Sherlock is  $X$  at the fiction, we do not get a determinate truth-value, since some worlds in the set hold that Sherlock is  $X$  and other worlds in the set hold that Sherlock is  $Y$ . Additionally, if we designate the fiction as actual but do not specify which completion of the fiction is designated as actual, it is indeterminate whether a denizen of the fiction uses “Sherlock” to refer to  $X$  or to  $Y$ . So the incompleteness of fictional content leads to two kinds of indeterminacy: indeterminacy can arise both in evaluating a specified content, and in generating a content from an utterance.

It is important as always to be clear about the distinction between our utterances about the content of a fiction, and the fictional utterances made by the characters within the story. In asking what propositions constitute the content of the fiction, we are asking what content would be expressed by unprefixing sentences made within the story. Our unprefixing utterances

using the name “Sherlock” generally express degenerate propositions; it is the content of Watson’s utterances that pose a challenge here. I claim that Watson would express a singular proposition when he makes Sherlock-utterances. *Which* singular proposition does he express? It is indeterminate whether they are singular propositions containing  $X$  as a constituent or  $Y$  as a constituent.

The problem has been narrowed down to a question about the implications of incompleteness for content-generation, which arises because a set of worlds is designated as actual. More needs to be said here, but I will wait until §8.2 to expand on it.

### 7.3 Actualism

Another reason to think that “Sherlock” fails to refer is that there is no such individual. To put it simply (if not entirely clearly), actualists hold that everything there is is actual. Some formulations of actualism leave no room for aliens, i.e., things that are merely possible. I claim that there are aliens,<sup>23</sup> so my view is incompatible with such versions of actualism.

I think it is highly intuitive that there could have been more simple things than there in fact are, and that seems to commit us to aliens. But I do not want to endorse possibilism, either; a possibilist could evade entirely the problem that is the subject of this paper, simply by claiming that  $\alpha$ -speakers can use terms to designate aliens. I am happier with a view Karen Bennett (2005) labels “non-domain-inclusion actualism”, a view according to which some merely possible entities are not constituents of the domain of the actual world. Instead, we might say that there could have been things that in fact simply do not exist. If we adopt non-domain-inclusion actualism, we can hypothetically consider how things would be if some other world turns out to be actual, but maintain that these other worlds are not real in the same way that  $\alpha$  is. More needs to be said here about the ontology of possible worlds, of course, and that is a problem I lack the space to address here. But this outlines the broad strategy enough to explain why I am not motivated to think that “Sherlock” fails to refer on the grounds of actualism.

---

<sup>23</sup> Again, this should not be interpreted as an endorsement of Meinongianism. It is merely the claim that there could have been entities that are not actual, nor are they constructible out of actual entities.

## 7.4 Authorial Intentions

When Conan Doyle wrote the stories, he did not intend to be talking about anyone in particular, though Sherlock was inspired by a real individual. Since no individual is the author's intended target, some say either that the name stands for something else entirely (e.g., an abstract object created when the author writes the story), or that the name just fails to refer (perhaps the author was merely *pretending* to refer). As I indicated above, I have my doubts about these alternative approaches to fictional content. I prefer to say that it is indeterminate which among many individuals would count as Sherlock.

## 8 Unembedded Empty Names

I want to close by briefly considering two potential problem cases; in §8.1 I will address sentences containing comparisons across fictions. In §8.2 I will say a little about negative existentials.

People (like myself) who treat fictional names as empty names are confronted with a variety of problems, and I have not attempted to address them all here. Thus far I have focused on the contrast between (1) and (2). I have sought to explain how:

(1) According to the fiction, Sherlock Holmes is clever.

expresses a truth by offering a complicated analysis of the semantics of the prefix, while claiming that:

(2) Sherlock Holmes is clever.

simply expresses a degenerate false singular proposition. My strategy has been to abandon any intuition that (2) is literally true in order to say that the Millian semantics for names is correct for names that are not embedded under the fictionalist prefix. But there are some examples of unembedded fictional names for which this approach may not seem particularly compelling, and in those cases I can at least hint at a response.

### 8.1 Cross-Fiction Comparisons

There is no fictionalist prefix in:

(17) Sherlock is more clever than Poirot.

The basic strategy I have adopted for (1) and (2) might suggest that I would concede that (17) literally expresses a false degenerate proposition, but that a sentence like:

(18) According to the fiction, Sherlock is more clever than Poirot.

expresses a truth. Unfortunately for me, it is hard to argue that (18) expresses a truth, since Poirot is not a character in the Sherlock stories, nor vice versa. If I cannot say that (18) expresses a truth, then I do not have a ready explanation why (17) might even *appear* to say something true—I would not be able to claim that (17) is a sloppy way of trying to express some truth. So one might worry that the two-dimensionalist account will not be able to help the Millian provide a satisfying account of (17).

I do not find this example especially troubling; I will not take the space to discuss it in great detail, but I can offer the following as a quick response. Taken literally, (17) expresses a degenerate (and false) proposition. To semantically express a truth, we would need to say something like:

(19) If certain stories were true, Sherlock and Poirot would exist, and Sherlock would be more clever than Poirot.

A sentence like this might be taken to invoke a kind of fictionalist prefix:

(20) According to a fiction constructed by merging two well-known stories, ....

The speaker is effectively constructing a new fiction (which no doubt will be incomplete, and perhaps impossible) by merging the Conan Doyle stories and the Christie stories. It will be indeterminate what is true in the hybrid fiction, but that does not pose any new problems—it is indeterminate what is true in each of the original fictions too. So I do think that (17) is a sloppy way of attempting to express a truth, but that truth is one that would be properly expressed by something like (20).

## 8.2 Negative Existentials

Unembedded fictional names in negative existentials also pose a problem for Millianism. There are two scenarios that must be considered separately, as they raise different difficulties; it is the second that I have alluded



to in two places above. First, however, imagine a fiction in which it is clear that no actual individual is identical to the fictional character—perhaps a story exclusively about the Ents would suffice, since it seems that the Ents as envisioned by Tolkien are not members of any actual species. Imagine that we then utter a sentence about one of those Ents:

(21) Treebeard does not exist.

What proposition does (21) express? Not a singular proposition, since there is no such individual to be a constituent of the proposition. However, since (21) expresses a truth, we cannot dismiss it as expressing a degenerate singular proposition either.

Two-dimensionalism can help here, though the account will have to be more complicated than what we have seen above. I suggest that in effect, an utterance of (21) is an assertion that had the fiction been actualized, there would have been a singular proposition. Admittedly, that singular proposition would have been false when evaluated at the fiction—Treebeard *does* exist should the stories turn out to be true. But such a proposition would have been true when evaluated *at our world*  $\alpha$ . In the notation of double-indexing, I want to say that our utterance of (21) is a shorthand way of saying something like:

(22) “Treebeard does not exist” is true with respect to double-index  $(F, \alpha)$ .

That is, if a Treebeard-world were actualized, there would be a proposition denying the existence of Treebeard, and that proposition would be true when evaluated at  $\alpha$ . It is critical to note that the proposition would be *false* when evaluated at the fiction (i.e., when “Treebeard does not exist” is evaluated at  $(F, F)$ ), since Treebeard *does* exist in stories about Ents. But just as we can evaluate a proposition at another world, so too we can speak of what truth-value would be determined by the evaluation of a fictional proposition at some world outside of the fiction. So this is the truth an  $\alpha$ -speaker is conveying in uttering (18)—take the proposition that would be generated by the reference-fixer of “Treebeard doesn’t exist” combined with the Tolkien fiction designated as actual, and say that it would be true when evaluated at  $\alpha$ .

This proposed approach to impoverished-world negative existentials may not be intuitively very appealing. Unlike sentences of the form “If

the fiction turns out to be actual,  $S$ ”, there is nothing at the surface level of “Treebeard doesn’t exist” to suggest that we are shifting the world designated as actual. Even if one is untroubled by the appeal to reference-fixers in the previous section of this chapter, this approach to impoverished-world negative existentials may thus seem *ad hoc*. In response to the charge of ad-hocery, I offer a pair of observations.

First, impoverished-world negative existentials are deeply puzzling. No solution is very attractive, so we should expect to have to swallow a bitter pill to make sense of these utterances. Second, this account of impoverished-world negative existentials is also tied to a phenomenon that we have independent need to explain—designating another world as actual, as when we utter sentences of the form “If world  $\delta$  turns out to be actual,  $S$ ”. *Somehow*, we are able to express thoughts about what would have been true should we inhabit a world that contains more individuals than does  $\alpha$ . Intuitively, we are doing something like this in uttering an impoverished-world negative existential: we are envisioning a merely possible individual and explicitly stating that the individual is *merely* possible. Syntactically, it is odd to think that “ $X$  does not exist” involves shifting the world designated as actual, but intuitively, it explains something that might otherwise appear incoherent—supposing that there is an individual only to then assert that there is no such individual. We could not truly say that at some possible world, there exists an individual who does not exist at *that* world. But we *can* truly say that had some other world been actual, there might have existed an individual who does not in fact exist at our world. If we allow that we can shift the world designated as actual independently of using that world as the world of evaluation, we can provide an intuitively satisfying explanation of what is at issue with impoverished-world negative existentials.

I offer that as the beginnings of a way of understanding of one kind of negative existential. However, another kind of negative existential poses a very different problem. Consider a story that does *not* rule out the possibility that actual individuals are identical with the fictional character. There is a possible world in which every statement of the Nancy Drew stories is true, and in which the individual who solves the mysteries meets the identity conditions of Nancy Pelosi (after all, nothing in the stories specifies that Nancy Drew has some DNA sequence *other* than Nancy Pelosi’s). We want to say that our utterance of:

(23) Nancy Drew does not exist

is true. But the identity of Nancy Drew is indeterminate, and one of her possible identities is Nancy Pelosi. So one candidate for the content of (23) (supposing that “exist” expresses a property that can be exemplified) is:

(24)  $\langle$  It is not the case that  $\langle$  Pelosi, *existence*  $\rangle\rangle$ .

(24) is *false*; Pelosi *does* exist. It now appears that we cannot simply say that our utterance of (23) is true; at best, it seems to be neither true nor false. That would be unwelcome; we want to say that (23) says something unambiguously true. Suppose for simplicity that the fiction specifies so much about Nancy Drew’s biological features that the only possible candidates for Nancy Drew’s identity are actual individuals: Nancy Pelosi and Nancy Reagan, let us say.<sup>24</sup> In that case, in designating the fiction as actual, my proposal about negative existentials suggests that the proposition we are expressing in uttering (23) is something equivalent to what we might express in uttering:

(25) If the fiction were actualized, there would be an individual satisfying the reference-fixing conditions for counting as “Nancy Drew”, and such an individual does not exist at world  $\alpha$ .

But because both of the following propositions are false of our world  $\alpha$ :

(26)  $\langle$  It is not the case that  $\langle$  Pelosi, *existence*  $\rangle\rangle$ .

(27)  $\langle$  It is not the case that  $\langle$  Reagan, *existence*  $\rangle\rangle$ .

we would seem to have to say that (25) is false at  $\alpha$ . But we do not want to say that (23) is false of  $\alpha$ , so there is trouble for me if I claim that (23) expresses the same proposition that (25) expresses.

The worry is that some actual individual (such as Pelosi) could, in some other possible world consistent with everything specified by the Nancy Drew stories, be the individual that counts as Drew. But then it

---

<sup>24</sup> This supposition is incredibly implausible—surely there will be some *non*-actual individuals who fit any biological specifications a fiction could impose. I ask the reader to allow the supposition anyway, as it allows me to present the problem more clearly (and would only serve to aid my opponent at my own expense).

would be false to say of the individual who would be Drew if that possible world were actualized (i.e. Pelosi) that she fails to exist at  $\alpha$ —that individual *does* exist at  $\alpha$ . One way to address this is to specify that there are facts about the author's intentions concerning Drew's identity that were never explicitly expressed in the fiction: namely, the author lacked any intention of making Drew identical with an actual individual. One can tell fictional stories about real individuals, but one does so only by *intending* the story to be about that individual. In the absence of such an intention, I suggest that the story does not describe a possible life of any actual individual.

Is this solution *ad hoc*? I do not think so; we understand in reading those stories that they are meant to be fictional, and that is part of how we engage with them. Giving this role to the author's intentions is still compatible with treating fictional content as a set of worlds—we simply say that the Drew world-set includes some worlds and not others in virtue of what the author intended to express. It is in this way that I think the two-dimensional account of fictional names in the end needs to appeal to authorial intentions.

The suggestions here in §8 are of course underdeveloped, as my project here is not to show that Millians can successfully respond to all of the objections their opponents have raised. The main thrust of this paper is to suggest that the machinery of two-dimensional semantics offers another tool for the Millians to use in explaining how their view can accommodate fictional names.<sup>25</sup>

---

<sup>25</sup> Thanks to André Gallois, Mark Heller, Tom McKay, and an anonymous reader for helpful comments on various versions of this paper.

# Classical Possibilism and Fictional Objects

ERICH RAST

## 1 Classical Possibilism

For a brief period of time Russell considered a metaphysical position, henceforth called *classical possibilism* (POS), according to which there are various notions of existence and it is possible to make meaningful true statements about objects that don't actually exist.<sup>1</sup> Like many others Russell erroneously attributed this position to Meinong and finally rejected it for violating our "robust sense of reality".<sup>2</sup> In this article POS is laid out, it is shown on the basis of examples that POS can be used for expressing relationships between such sorts of nonexistent objects as fictional objects within the object language, and it is argued that the position does not violate our robust sense of reality as it is fully compatible with reductionist stances towards actually nonexistent objects. For the purpose of making examples easy to read a description theory of reference will be used to specify the meaning of proper names for various sorts of nonexistent objects. However, the account does not hinge on the acceptance of descriptivism and hints will be given at appropriate places on how to express the same metaphysical position in a more Millian fashion. An in-depth comparison of POS with other broadly-conceived Meinongian positions such

---

<sup>1</sup> See Priest (2005), pp. 105–6.

<sup>2</sup> Ibid., pp. 106–8.

as those of Lewis (1986), Salmon (1998), Routley (1980), or Priest (2005) is beyond the scope of this paper but similarities and differences between POS and these accounts will be mentioned whenever this is deemed appropriate.

Before going into the details a number of operational definitions are needed. According to *strict actualism*, henceforth referred to as *actualism*, it is not possible to truthfully ascribe an extralogical property to an object that doesn't actually exist. For example, according to the actualist doctrine the conjunction of (1) and (2) cannot be true.

- (1) Sherlock Holmes doesn't actually exist.
- (2) Sherlock Holmes is a detective.

In contrast with this, classical possibilists (in the sense the term is used here) claim that sentences like (1) and (2) can be true in the same model at the same time with respect to the same context. Two details about this distinction are important: First, the attribute "actually" cannot be left out; as their name indicates, actualists are committed to using actuality as an existence criterion in one way or the other. Second, it doesn't suffice to just say that for an actualist (1) cannot be, literally speaking, true. There are many solutions to the so-called problem of negative existentials that allow a sentence like (1) to come out true and are generally not considered possibilist. For example, if (1) is analyzed in a free logic using a non-denoting constant for "Sherlock Holmes" it may come out true in a model yet any simple predicative clause about Sherlock Holmes would come out false if it involved an *extralogical property*. For what it's worth such a theory should fall under the label "actualism". Proponents of positive free logic have even claimed that a statement like  $a = a$  can be true when  $a$  does not denote,<sup>3</sup> which seems to imply that  $a$  has the logical property of being identical with itself, yet such accounts are actualist in spirit. Hence, only *extralogical* properties should count when distinguishing actualism from possibilism. The idea is that according to actualism a statement ascribing an extralogical property to an object that doesn't exist must be non-true (false, without truth-value, having a third truth value, etc.), whereas it may come out true according to possibilism.

---

<sup>3</sup> See Bencivenga (1994), pp. 396–8.

As is well-known, a form of possibilism can readily be obtained in classical first-order predicate logic (FOL) when quantifiers are interpreted as being existentially unloaded (a term coined by Lewis and used by Linsky & Zalta, 1994) and one or more existence predicates are introduced into the language. Using  $\exists$  and  $\forall$  for the unrestricted quantifiers, relativized quantifiers may be introduced in the usual way as  $\forall^*xA := \forall x(Ex \supset A)$  and  $\exists^*xA := \exists x(Ex \wedge A)$ . Since we will use the same technique in first-order modal logic, two common objections need to be addressed and ruled out right from the start. First, many philosophers such as Kant in the *Critique of Pure Reason*, Frege (1983), and Ryle (1931–1932) have argued that existence is not a predicate or at least is not a real, irreducible predicate. Viable counter-arguments have already been laid out against this position elsewhere.<sup>4</sup> Second, logicians sometimes point out that sorting a logic by partitioning the domain using predicates and defining relativized quantifiers on the basis of them is old hat and not very interesting from a logical point of view, as if this fact alone would somehow speak against the account. It is hard to see how this line of thought could be turned into a convincing argument and we will ignore it in what follows.

An actualist can have as many existence predicates as he wishes, if he accepts that existence can be a predicate in the first place. However, according to the actualist doctrine all of these have the same extension and are redundant in the sense of being reducible to ordinary quantification by means of existentially-loaded quantifiers. Whenever a simple predicative clause  $Pa$  is true in the actualist system it follows that  $a$  actually exists. Hence, whatever existence predicate the actualist might ascribe to an object  $a$ , the whole predicative clause will only come out true if  $a$  actually exists. This means, of course, that there is no compelling logical reason for an actualist to have one or even more existence predicates, yet there might be other reasons for an actualist to use one or more existence predicates—be it only for the fact that from a grammatical point of view English “to exist” acts as a grammatical predicate.

Things look different when a possibilist perspective is taken, because various distinct possibilist positions can be introduced on the basis of one or more existence predicates in FOL. To give one example, Meinong’s position may be approximated by using three unary predicates  $E_1, E_2, \overline{E}_2$ ,

<sup>4</sup> See for example Salmon (1998, pp. 290–1), Routley (1980), and Priest (2005).

where  $E_1x$  is read *x exists*,  $E_2x$  is read *x subsists*,  $\overline{E_2}x$  is read *x has non-being*, and the following restrictions hold in intended models:

$$\forall x[E_1x \supset E_2x] \quad (3)$$

$$\forall x[(E_1x \vee E_2x) \supset \sim \overline{E_2}x] \quad (4)$$

$$\exists x \sim (E_1x \vee E_2x \vee \overline{E_2}x) \quad (5)$$

While this position is not defended here, this example illustrates the general usefulness of FOL with existence predicates for doing metaphysics, once Quine’s famous dictum “to be is to be the value of a bound variable” (Quine, 1948/1961) is rejected and quantifiers are taken as existentially unloaded means of counting objects. Relations between different *sorts* of actually nonexistent objects and existing ones can readily be expressed in classical FOL. For example, model constraint (3) expresses the rule that whatever exists also subsists, but not necessarily *vice versa*.

## 2 From FOL to Modal Logic

Various existence predicates can be reduced to a uniform existence predicate in combination with several modalities by switching from FOL to a constant-domain double-index first-order modal logic (NDML). In this section, some technical peculiarities of this language will be briefly introduced—the full logic can be found in the Appendix. Apart from the use of non-traditional predication theory in order to deal with incomplete fictional objects, the account remains classical.

NDML is similar to the logic in Kaplan (1989) and truth is relativized to context and index parameters respectively, where the latter encode times and worlds and the former times, worlds, agents, addressees and whatever else is required for the saturation of indexicals. Mapping functions with fairly self-explanatory names are used for retrieving the respective factors of contexts and indices. For example,  $world(c)$  yields the world of a context  $c$ . For descriptions a two-place quantifier (“iota quantifier”) may be defined as follows, where  $A\{\alpha/\beta\}$  is the same formula as  $A$  except that all free occurrences of  $\alpha$  in it have been replaced by  $\beta$ .

$$\iota x[A]B := \exists x[A \wedge \forall y(A\{x/y\} \supset x = y) \wedge B] \quad (6)$$



For each modality  $m$  there is a modal operator  $\Box^m$  and for finitely many agents  $a \in \text{Agt}$  there are finitely many attitudes. We will only consider rational belief in what follows and write  $\text{Bel}_a\phi$  for the fact that  $a$  believes that  $\phi$ . Standard operators FutA and PastA for minimal absolute tenses introduce a new index time in relation to the time of the context and are read *it will be the case that A* and *it was the case that A* respectively. The sentential operator @ shifts the world and time of the index to the world and time of the context and thus works as an actuality operator. In order to account for incomplete objects, non-traditional predication theory (NTPT) of Sinowjew (1970) and Sinowjew & Wessel (1975) will be used, which allows for two kinds of predication in the object language. Each positive predicate symbol  $P$  is given an inner negation form  $\neg P$ , where  $\neg P(t_1, \dots, t_n)$  is interpreted as the denial of the fact that  $t_1, \dots, t_n$  stand in relation  $P$  or that an object has the property expressed by  $P$  in the unary case, and the extensions of any predicate or relation symbol  $P$  and  $\neg P$  must be distinct at each context-index pair at which they are evaluated.<sup>5</sup> (The symbol  $\sim$  is used for the ordinary, truth-functional negation.) In what follows, only intended models will be considered in which  $E$  is understood as an existence predicate.

### 3 Fictional Objects

Suppose  $\Box^f$  is a fictional modality which, via an accessibility relation, picks out the worlds that are compatible with a certain work of fiction, say, Doyle's Sherlock Holmes stories. Let  $D$  denote the property of being detective and  $S$  denote the property of being called "Sherlock Holmes" according to nominal description theory.<sup>6</sup> We can then express the fact that Sherlock Holmes is a detective as  $\iota x@[Sx \wedge \neg Ex \wedge \Box^f Ex]Dx$ . This means that it is part of the meaning of "Sherlock Holmes" that he doesn't

<sup>5</sup> Originally having been developed independently, NTPT is closely related to (partial) logics of first degree entailment investigated by Anderson, Belnap, and Dunn. See Anderson et al. (1975, 1992); cf. Muskens (1995, chap. 5),  $N_4$  in Priest (2001, chap. 8) and the similar system in his contribution to this volume. Notice, however, that in terms of its conceptual underpinnings and original motivation NTPT differs considerably from partial and relevant logics, as it preserves bivalence and is not paraconsistent.

<sup>6</sup> The term nominal restriction theory is borrowed from Kent Bach. See Bach (2002) and Rast (2007, pp. 100–3) for arguments why the property of being called by a certain name doesn't violate the circularity prohibition of Kripke (1980).

actually exist now, but that he exists in the worlds compatible with Doyle's Sherlock Holmes novels.

So far so good. Now it seems striking that fictions are social constructs. If this is so, then the question which properties Sherlock Holmes has is not decided by Doyle's writings alone, but also by various experts on his novels. Moreover, many of Sherlock Holmes properties are decided by encyclopedic, everyday common sense knowledge about objects, persons, detectives, doors, streets, dressing habits of the late 19th and early 20th century, and so forth. All of these sources only partially determine Sherlock Holmes, and so in contrast to Lewis' view Sherlock Holmes can only be partially determined by these sources. In the setting of non-traditional predication theory this means that statements of the form  $\iota x@[Sx \wedge \neg Ex \wedge \Box^f Ex] \sim \phi(x) \wedge \sim \neg\phi(x)$  can come out true for a predicate  $\phi$  in intended models with respect to a given context-index pair, i.e. the property expressed by  $\phi$  can neither be ascribed to, nor denied of the given fictional object. It cannot be emphasized enough in this context that the question which properties cannot be ascribed to Sherlock Holmes and which ones he has is neither a matter of logic nor a matter of metaphysics. This question is really decided by the way people generally deal with the Sherlock Holmes stories, by experts on these stories, and by the background knowledge required to understand them. Sometimes there may be no authoritative source for determining whether Sherlock Holmes has a certain property or not, yet if we understand the socio-linguistic processes underlying Sherlock Holmes stories we sufficiently well understand the  $\Box^f$  modality. From a logical and semantical point of view we do not need to understand the modality in each and every detail in order to use it, though, just as we do not need to understand other extralogical expressions such as "freedom" in each and every detail in order to do truth-conditional semantics. So if someone would object to a possible worlds account of fiction that Sherlock Holmes' properties and the identity conditions of fictional objects in general are left somehow unclear and appear to be mysterious at times, the adequate reply is: Read the stories, watch the movies, and consult an expert! If you have done this and it is still unclear whether a fictional object has a certain property or not, then withhold from ascribing or denying that the object has that property!

To illustrate the expressive power of the account, let us take a look at an extended example. Suppose the actual world is denoted  $w_0$  and  $w_i$  stands

for a world compatible with Conan Doyle's Sherlock Holmes stories. Now consider the following utterances with their intended evaluations:

- (2) Sherlock Holmes is a detective. (true in  $w_0$ , true in all  $w_i$ )
- (7) Sherlock Holmes doesn't exist. (true in  $w_0$ , false in all  $w_i$ )
- (8) Sherlock Holmes exists. (false in  $w_0$ , true in all  $w_i$ )
- (9) Sherlock Holmes is a flying pig. (false in  $w_0$ , false in all  $w_i$ )
- (10) Sherlock Holmes is not a flying pig. (true in  $w_0$ , true in all  $w_i$ )
- (11) Sherlock Holmes loves his wife. (false in  $w_0$ , false in all  $w_i$ )
- (12) Sherlock Holmes doesn't love his wife. (false in  $w_0$ , false in all  $w_i$ )
- (13) Sherlock Holmes was cleverer than Hercule Poirot. (Salmon, 1998, p. 298) (by assumption true in  $w_0$ , false in all  $w_i$ )
- (14) Sherlock Holmes wasn't cleverer than Hercule Poirot.  
(by assumption false in  $w_0$ , false in all  $w_i$ )
- (15) Sherlock Holmes is a fictional character. (true in  $w_0$ , false in all  $w_i$ )

According to the position laid out above (2), (7)–(15) may hold within the same model and with respect to the same context, where the inner negation is used in (7), (10), (12), and (14) as in previous examples.<sup>7</sup> These assessments of truth-conditions are certainly not uncontroversial and there are alternatives. Thus, according to Salmon (1998, pp. 298–310), “Sherlock Holmes” denotes an abstract object in the actual world and so, taken literally, an utterance like (2) is untrue in  $w_0$  in his opinion. In contrast to this “Sherlock Holmes” in the above example denotes

---

<sup>7</sup> Assuming description theory, examples (7) and (8) are analytically true and false with respect to any fixed actual world (the world of the context) and therefore tautological and contradictory in any model with designated context. However, when we say that an utterance (with corresponding formula  $\phi$  in the semantic representation language) is true or false in  $w_i$  this must be understood as a shortcut for saying that  $\Box^f \phi$  is true with respect to the given initial context and index and not as a way of evaluating  $\phi$  with respect to *counterfactual contexts*. In our view it makes no sense to evaluate utterances with respect to non-actual worlds simpliciter, i.e. without embedding them into a suitable modality in the object language, and we also do not endorse Lewis' modal realism.

a concrete person that doesn't actually exist. Ascribing the property of being a fictional character to Sherlock Holmes is from this point of view another way of expressing the fact that Sherlock Holmes only exists as a fictional object. As Salmon lays out convincingly this position is not tenable when fictional characters are also considered abstract objects, "[...] since abstract entities make terrible musicians" (ibid., p. 303). While the given machinery is somehow neutral about this issue and allows for models in which sentences like (2) are false and Sherlock Holmes is an abstract object in the actual world, we submit that people do not ordinarily talk about Sherlock Holmes as if he was an abstract object. Instead, Sherlock Holmes is commonly ascribed most of the properties in the actual world that he also has in worlds of Conan Doyle's Sherlock Holmes stories except for being nonexistent and fictional in the former but existent and non-fictional in the latter, and this way of looking at Sherlock Holmes in the actual world seems to be generally more fruitful than considering him an abstract object.<sup>8</sup> The view that Sherlock Holmes is a person made of flesh and blood does not preclude the possibility of reducing him ontologically to another entity, perhaps an abstract entity, as part of a general reductionist programme, and in Section 5 we will sketch how to formulate such a programme.

## 4 Other Sorts of Possibilia

There are many other nonexistent objects apart from the fictional ones. For example, there are objects whose existence varies in the temporal dimension. Existence stipulations about past objects are generally more problematic than those about fictional ones. If it is taken as part of the meaning of "Socrates" that he has existed and no longer exists, then a speaker community needs to change language in retrospective once it is found out that Socrates has never existed, by the same token as it is possible at any time that water turns out to be XYZ. A case when language needs to be revised

---

<sup>8</sup> In defense of Salmon's position it must be mentioned, though, that there are cases when our talk about Sherlock Holmes as a detective in Doyle's stories and Sherlock Holmes as a fictional object is incompatible with certain background assumptions, for example the assumption that an object cannot be both a detective and a fictional detective, because "fictional" is an intensional adjective. We follow Kripke that in these cases the uses of the corresponding proper name are equivocal.

because our scientific beliefs have been revised may occur at anytime, because science is principally fallible; Millians and descriptivists are in the same boat with respect to such cases. However, for the descriptivist a statement like  $\lambda x@[Sx \wedge \neg Ex \wedge \text{Past } Ex] \neg Ex$  is true by linguistic apriori and many Kripkeans do not share this intuition.<sup>9</sup> Someone who fully accepts the consequences of Kripke's semantic argument will therefore remove the existence stipulations from the descriptions and instead formulate general metaphysical rules for past objects. In both cases (i) a speaker may refer to Socrates now and (ii) ascribe properties to him now even though (iii) he no longer exists. Since we are not convinced that Kripke's epistemic and semantic argument speak against *this particular application of description theory* and descriptions serve our illustrative purposes well, we make the existence stipulations part of the meaning of the term for a past object.

Names for future objects are usually introduced by description and if for example someone introduces the term "Spirit" by description as the model airplane he will build, then the plane really doesn't exist yet. As in the previous case, the question whether the existence stipulation is considered part of the meaning of the term or not depends on the Kripkean or Anti-Kripkean commitments one is willing to make. Notice, however, that any existing plane called "Spirit" is not the plane that the speaker has in mind when he introduces the name. *His* utterance of "Spirit will be great" may consequently be analyzed as  $\lambda x@[Sx \wedge \neg Ex \wedge \Diamond \text{Fut } Ex] \text{Fut } Gx$ , given the descriptivist premise that speakers will adjust the description associated with "Spirit" accordingly to  $\lambda x@[Sx] \dots$  when the plane comes into existence.<sup>10</sup> A Millian may achieve a similar effect by stipulating a rule like  $\forall x[Fx \supset \neg Ex \wedge \Diamond \text{Fut } Ex]$ , where  $F$  denotes the property of being a future object. In this case the Millian position seems to be favorable, as the Millian doesn't have to explain a language change when the future object comes into existence.

Doxastic objects exist in the worlds compatible with a given propositional attitude that is formally represented by a corresponding modal operator. Although there are many attitudes such as fear, making assumptions, doubting something, we will only consider rational belief in

<sup>9</sup> Notice that  $\Diamond \lambda x@[Sx \wedge \neg Ex \wedge \text{Past } Ex] \text{Now } Ex$  is satisfiable.

<sup>10</sup> Semantic ambiguity and the equivocality of proper names can in the present framework be analyzed by making the respective predicates like  $Sx$  (for being called "Spirit") context-sensitive. A discussion of the benefits and perils of this approach is left for another occasion.

what follows. Consider the famous wanderer in the desert, let's call him "Bob", who hallucinates and comes to believe that there is water at an oasis in front of him while pointing to that imaginary place. From an omniscient third-person perspective, Bob's belief may be described as  $Bel_b \exists^* y \lambda x [Ox \wedge P(I, x)] (Wy \wedge L(y, x))$ , where  $I$  is the first-person pronoun (see Definition 6 in the Appendix),  $Bel$  expresses a KD45 modality,  $L(x, y)$  is read  $x$  is located at  $y$ ,  $P(x, y)$  is read  $x$  points at  $y$ ,  $Ox$  is read  $x$  is an oasis, and  $Wx$  is read  $x$  is water. Suppose Bob believes the oasis he believes to be in front of him is called "Tisserdimine" and  $T$  expresses the property of being called by that name. Then Bob's *subjective reference* by means of the name can be expressed from a third-person perspective by a description of the form  $\lambda x @ [Bel_b (Ex \wedge Ox \wedge Tx)] \dots$  and his use of "Tisserdimine" can be modeled by this description. Since the property of being called by a certain name rarely suffices for being able to identify the referent of a given name, subjective identification criteria need to be added whenever an agent's attempt to identify the referent of a name is to be made explicit. In the present example, the property of being an oasis fulfills this role.

Having a notion of subjective reference might not seem necessary from the point of view of semantic externalism, but it is needed for expressing identifying reference as a condition between subjective reference and semantic reference. The motivation for introducing subjective reference—or speaker reference, as one may call it alternatively—and explaining it on the basis of a description theory of reference is thus primarily epistemic, and an elaborate descriptivist need not commit himself to the view that competent speakers need to associate identification criteria with a proper name *in virtue of being competent users of that name*. They activate and associate such criteria whenever they attempt to identify the referent of a referential expression, as long as identifying something as something involves discriminating something from other things by means of the properties it is supposed to have.<sup>11</sup>

There is a transition from subjective to intersubjective, yet still anti-realist reference when shared beliefs instead of beliefs by single agents are taken into account. Suppose we add groups of agents to our language and define  $Bel_G \phi$  as true with respect to  $c, i$  if all agents in the group  $G$  believe that  $\phi$ . (Since we have stipulated that there are only finitely

<sup>11</sup> Perhaps other modes of identification are required for explaining phenomena like essential indexicals and *qualia* in thinking.

many agents, this corresponds to a finite conjunction of individual beliefs.) Then it can happen that a group of people speaks the same sociolect with respect to a given proper name. For example, an utterance of “Tisserdimine is our rescue” with respect to this sociolect may be analyzed as  $\lambda x@[Bel_G(Ex \wedge Tx \wedge I_Gx)]Rx$ , where  $I_G$  represents identification criteria shared by members of the group.<sup>12</sup> Once again, it could be argued on Millian grounds that neither the existence stipulation nor the shared identification criteria are part of the term’s meaning. However, even if they are not made part of the meaning of the name (within the given sociolect) the Millian still has to take into account some identification criteria assumed by members of the group if he wants to explain their behavior with respect to their understanding of “Tisserdimine”.

Finally, there are also actually nonexistent objects in the metaphysical sense—entities that our world would contain had it evolved in a different way than it actually has, if parameters of scientific theories were different from what they are, or if different laws of nature held. These are generated by corresponding modalities with readings like *it is necessary that* or *it is possible that*. A thorough discussion of these sorts of actually nonexistent objects would go beyond the scope of this paper and we refer the reader to the relevant literature, for instance Lewis (1973, 1986) and Stalnaker (2003).

## 5 Reductionism

It might be desirable to reduce various forms of existence to actual existence and one of the main motivations for classical possibilism is precisely to be able to express reductionist positions within the object language in one and the same model instead of talking about relations between different models, as the Quinean is forced to do. Let  $\Box$  stand for any box-modality of the language such as  $\Box^f$  or  $Bel_a$  for a particular agent  $a$  and let  $\mathcal{R}$  stand for a special *encoding relation*. Generic reductionism can then be expressed as follows.

$$\forall x \exists y [(\neg Ex \wedge \Box Ex) \supset (Ey \wedge \mathcal{R}(y, x))] \quad (16)$$

<sup>12</sup> “Shared” is understood in a rather weak sense here. It may include the case that  $I_Gx$  abbreviates a large disjunction of the form  $I_{a_1}x \vee I_{a_2}x \vee \dots \vee I_{a_n}x$  for agents  $a_1, a_2, \dots, a_n \in G$ .

For every object  $x$  that doesn't actually exist but exists in some way there is an object  $y$  that actually exists and encodes  $x$ . It would go beyond the scope of this article to provide an actual theory of what it takes to stand in an encoding relation  $\mathcal{R}$ , but at least (16) illustrates that it is possible to be a reductionist in the present framework. Many potentially interesting metaphysical positions may be expressed in the current setting. To give another example, a form of anti-realism about fictional objects may be formulated as follows.

$$\forall x \exists y [(\neg Ex \wedge \Box^f Ex) \supset (Ey \wedge Bel_y \Box^f Ex)] \quad (17)$$

Whenever something exists as a fictional objects someone actually exists who believes that it exists as a fictional object. These examples show that a classical possibilist doesn't need to be committed to a huge ontology when "ontology" is understood in the narrow sense as an investigation of what it means to *actually* exist. Things that possibly exist, exist as fictional objects, have existed, will exist, or exist in worlds compatible with what someone believes don't actually exist and the meaningfulness of our talk about them does not imply that they cannot be explained entirely in terms of actually existing objects. As can be seen from the above examples, classical possibilism can be coupled with many kinds of reductionism to the actual and is therefore relatively neutral with respect to ontological commitments. The position acknowledges that actually nonexistent objects may have various extralogical and nontrivial properties, but it doesn't commit us to specific criteria for determining the properties of these objects.

## 6 Summary and Conclusions

Classical Possibilism is the view that objects that don't actually exist can exist in various other ways, for example as fictional objects, as doxastic objects, or as past objects. While we have used a description theory for laying out various sorts of objects that actually don't exist it has hopefully become clear that the position can also be adopted from a Millian perspective. Non-traditional predication theory has been used to allow fictional objects to be underdetermined and various sorts of actually nonexistent



objects have been discussed. While other broadly-conceived Meinongian positions like those of Routley (1980), Priest (2005) and in this volume, Parsons (1980) and Lewis (1986)<sup>13</sup> could not be addressed here in detail for lack of space, it must finally be mentioned that POS is, not very surprisingly, similar to Priest's Noneism. According to Priest, things that don't actually exist don't exist in any other way, which makes it harder to formulate rules like (3)–(5) or positions like (16). However, in a possibilist system the existence predicate behaves like any predicate, uses of the existence predicate could therefore in principle be replaced by other predicates, and any serious metaphysician must somehow differentiate between different sorts of possibilia. Whether one speaks of different ways to exist or uses other labels to classify objects that don't actually exist is primarily a terminological issue. The main difference between the present view and Priest's thus lies in his commitment to dialethism and corresponding use of impossible worlds, whereas in the present setting only the case  $\sim Pa \wedge \sim \neg Pa$  can be expressed.

The lesson to be drawn from the above examples is that classical possibilism is more appealing than it might seem at first glance because it allows for the formulation of metaphysical laws, including forms of reductionism to the actual, within the object language within one and the same model. Being able to formulate law-like metaphysical statements and meaning postulates of the kind laid out above is a good starting point for doing serious metaphysics.<sup>14</sup>

---

<sup>13</sup> See Linsky & Zalta (1991) for a thorough discussion of the Meinongian and Anti-Meinongian aspects of Lewis' metaphysics.

<sup>14</sup> Earlier versions of this article were presented at the 6th Barcelona Workshop on Reference in June 2009 and at the IFL, Lisbon. I would like to thank all the participants in these meetings and particularly David Braun and Robert Stalnaker for their useful comments and a fruitful discussion. This research was conducted under a Postdoctoral Fellowship from the Portuguese *Fundação para a Ciência e a Tecnologia*.

## Appendix: Language NDML

The language is a normal double-index modal logic similar to Kaplan's Logic of Demonstratives. For a more realistic treatment of indexicals, tenses, and spatial adverbs like "everywhere", additional entities and various pragmatic restrictions would have to be taken into account.

### Syntax

**Definition 1** [Alphabet] Variables are  $x, y, z$  and their indexed variants, constants are  $a, b, c$  and their indexed variants, terms are variables, constants, or any of the special terms  $I$  and  $here$ , predicate and relation symbols in positive form are sequences of one or more letters starting with capital letter and have a fixed arity (written  $P^n$ ). If a predicate or relation symbol  $\rho$  is in positive form,  $\neg\rho$  is the corresponding inner negation form.

**Definition 2** [WFF] For variable  $v$ , terms  $\tau_1, \dots, \tau_n$ , modal indices  $m$ , predicates  $\rho$  (in positive form or in inner negation form), and formulas  $A, B$ :

$$A := \rho(\tau_1, \dots, \tau_n) \mid (A \wedge B) \mid \sim A \mid \forall vA \mid \exists vA \mid \Box^m A \mid \Box_\tau^m A \\ \mid @A \mid \text{Now}A \mid \text{Fut}A \mid \text{Past}A$$

The iota quantifier is defined above in formula (6) and other modal operators and truth-functions are defined as usual. We write  $Bel_\tau\rho$  for  $\Box_\tau^0\rho$ ,  $1v@[A]B$  for  $1v[@A]B$  and leave out unnecessary parentheses.

### Semantics

**Definition 3** [Basic Entities] Contexts  $c \in C$  are ordered tuples of worlds, times, places, and agents and indices  $i \in I$  are ordered tuples of worlds and times. Accessor functions with self-explanatory names are used to retrieve the respective component of a context or index, e.g.  $world(c)$  is the world of the context  $c$  and  $time(i)$  is the time interval of an index  $i$ . Auxiliary sets are defined as  $W := \{world(x) \mid x \in (C \cup I)\}$ ,  $T := \{time(x) \mid x \in (C \cup I)\}$ ,  $L := \{place(x) \mid x \in (C \cup I)\}$ , and  $Agt := \{agent(x) \in D \mid x \in C\}$ .

**Definition 4** [Model] A model  $\mathcal{M} = (C, I, R, D, \llbracket \cdot \rrbracket)$  of NDML consists of a non-empty set of contexts, a non-empty set of indices, a set of accessibility relations containing a dyadic relation on  $W$  for each non-doxastic

modal operator  $m$  and a relation  $R_a^m \subseteq (W \times W \times \text{Agt})$  for each epistemic modal operator; the total domain  $D$ , and an evaluation function  $\llbracket \cdot \rrbracket$  as defined below.

**Definition 5** [Term and Predicate Interpretation] Terms and predicates are interpreted as follows (unless a more specific rule is given):

- $\llbracket \mathbf{v} \rrbracket^g(c)(i) = g(\mathbf{v})$ ,
- if  $\tau$  is a constant  $\llbracket \tau \rrbracket^g(c)(i) \in D$ , and
- if  $\rho$  is an  $n$ -ary predicate, then  $\llbracket \rho \rrbracket^g(c)(i) \subseteq D_1 \times \cdots \times D_n$ .

**Definition 6** [Special Terms]

- $\llbracket I \rrbracket^g(c)(i) = \text{agent}(c)$
- $\llbracket \text{here} \rrbracket^g(c)(i) = \text{place}(c)$

The operator  $\triangleleft$  is overloaded and yields a context or index variant with one of its members respectively changed to the value on the right hand side:

**Definition 7** [Context and Index Variants] Let  $\alpha \triangleleft t$  be the same context or index as  $\alpha$  except that  $\text{time}(\alpha) = t$  for time interval  $t$ , and correspondingly for other admissible members of contexts and indices.

**Model Constraint 1** [Inner Negation Constraint] In any NDML model and with respect to any context  $c$  and index  $i$ ,  $\llbracket \rho \rrbracket^g(c)(i) \cap \llbracket \neg \rho \rrbracket^g(c)(i) = \emptyset$ .

This means that  $\sim Pa \wedge \sim \neg Pa$  is satisfiable and can be abbreviated as  $?Pa$  but unlike in Priest's systems with impossible worlds  $Pa \wedge \neg Pa$  remains unsatisfiable.

**Definition 8** [Truth in a Model]

- $\llbracket \rho(\tau_1, \dots, \tau_n) \rrbracket^g(c)(i) = 1$  if all of  $\llbracket \tau_1 \rrbracket^g(c)(i), \dots, \llbracket \tau_n \rrbracket^g(c)(i)$  are defined and  $(\llbracket \tau_1 \rrbracket^g(c)(i), \dots, \llbracket \tau_n \rrbracket^g(c)(i)) \in \llbracket \rho \rrbracket^g(c)(i)$ ; 0 otherwise.
- $\llbracket \neg \rho(\tau_1, \dots, \tau_n) \rrbracket^g(c)(i)$  as in previous clause (as laid out above, inner negation is just a special sort of predication)

- $\llbracket \sim \rho(\tau_1, \dots, \tau_n) \rrbracket^g(c)(i) = 1$  if  $\llbracket \rho(\tau_1, \dots, \tau_n) \rrbracket^g(c)(i) = 0$ ; 0 otherwise.
- $\llbracket \forall vA \rrbracket^g(c)(i) = 1$  if for all  $v$ -variants  $h$  of  $g$ :  $\llbracket A \rrbracket^h(c)(i) = 1$ ; 0 otherwise.
- $\llbracket \exists vA \rrbracket^g(c)(i) = 1$  if there is a  $v$ -variant  $h$  of  $g$  such that  $\llbracket A \rrbracket^h(c)(i) = 1$ ; 0 otherwise.
- $\llbracket \Box^m A \rrbracket^g(c)(i) = 1$  if for all  $w$  such that  $R^m(\text{world}(i), w)$ :  $\llbracket A \rrbracket^g(c)(i \triangleleft w) = 1$ ; 0 otherwise.
- $\llbracket \Box_a^m A \rrbracket^g(c)(i) = 1$  if  $\llbracket \tau \rrbracket^g(c)(i)$  is defined and some  $a \in A$  and for all  $w$  such that  $R_a^m(\text{world}(i), w)$ :  $\llbracket A \rrbracket^g(c)(i \triangleleft w) = 1$ ; 0 otherwise.
- $\llbracket \text{Past}A \rrbracket^g(c)(i) = 1$  if there is a  $j$  that is like  $i$  except that  $\text{time}(j) < \text{time}(c)$  and  $\llbracket A \rrbracket^g(c)(j) = 1$ .
- $\llbracket \text{Fut}A \rrbracket^g(c)(i) = 1$  if there is a  $j$  that is like  $i$  except that  $\text{time}(j) > \text{time}(c)$  and  $\llbracket A \rrbracket^g(c)(j) = 1$ .
- $\llbracket @A \rrbracket^g(c)(i) = 1$  if  $\llbracket A \rrbracket(c)(i \triangleleft \text{world}(c) \triangleleft \text{time}(c)) = 1$ ; 0 otherwise.
- $\llbracket \text{Now}A \rrbracket^g(c)(i) = 1$  if  $\llbracket A \rrbracket(c)(i \triangleleft \text{time}(c)) = 1$ ; 0 otherwise.
- $\llbracket \text{Actually}A \rrbracket^g(c)(i) = 1$  if  $\llbracket A \rrbracket(c)(i \triangleleft \text{world}(c)) = 1$ ; 0 otherwise.
- $\llbracket \text{Everywhere}A \rrbracket^g(c)(i) = 1$  if  $\forall l \in L$ :  $\llbracket A \rrbracket(c)(i \triangleleft l) = 1$ ; 0 otherwise.

# How Creationism Supports Kripke's Vichianism on Fiction

ALBERTO VOLTOLINI

In this paper, I want to show that a reasonable thesis concerning truth in fiction, Fictional Vichianism (FV)—according to which fictional truths are true because they are *stipulated* to be true—can be positively endorsed if one grounds Kripke's justification for (FV), which traces back to the idea that names used in fiction never refer to concrete real individuals, into a creationist position on fictional entities that allows for a distinction between a pretending and a characterizing use of fiction-involving sentences. Thus, sticking to (FV) provides a reason for a metaphysically moderate ontological realism about fictional entities.

## 1 Fictional Vichianism

As is well known, in his masterpiece *The New Science* (1730–1744/1948), the Italian 18th-century philosopher Giovanni Battista Vico defended the thesis that *verum ipsum factum*, namely, the thesis that truth is the same as historical fact.

Vico's thesis is normally meant as a claim in epistemology: given a Cartesian form of scepticism about the outer world, the only things that can be known to be true are those made by humans. Yet, by itself, it may well be meant as a claim in ontology: the only things that are true are those

which concern *lato sensu* mind-dependent things, like things that humans make; in a nutshell, the only things that are true are historical facts.

Even then, the thesis remains quite controversial, for it relies on an analogously controversial ontological reading of the aforementioned Cartesian scepticism, on which there simply are no facts of the matter when it comes to things that are ultimately mind-independent. However, there seems to be at least one region of reality in which things are true just because they are made by humans, that is, in which things are, as artefacts, effectively mind-dependent things, or in other words, artefactual facts. In particular, this is the case of *fictional* truths, a kind of artefactual facts.

In this respect, once one switches from considering truth ontologically as factuality to considering truth linguistically as a property of (some interpreted) sentences, one may well envisage a Vichianist position on fiction, *Fictional Vichianism*:

**(FV)** Fictional truths are true because they are *stipulated* to be true, not because they are made true by an external reality.<sup>1</sup>

(FV) seems reasonable enough. Fiction precisely seems to be a case in which whatever is true or false depends on us, in particular on the choices that a certain author makes by telling a certain story in a certain way. In point of fact, if we take a sentence like:

(1) Holmes is a detective

there seems to be no reason why this sentence is true, but the fact that Conan Doyle decided to so characterize Holmes in his narratives: if it is true, then it has been so stipulated. Of course, Doyle might have written that Holmes was a folksinger or a politician; in such a case (1) would be false, as Doyle would have simply conjured up a completely different story. As Deutsch (2000) puts it, a fictional story is something made up. Conversely, if something apparently involving fiction does not stem from an author's decision or cannot be inferred either from her decisions or from the environment her decisions have contributed to sketch, there

---

<sup>1</sup> For a clear formulation of (FV), cf. the following passage: “as far as literally fiction is concerned, the author is the authority. There is nothing else—e.g. reality, or other texts—towards which he is responsible. [...] To read a text as a literary fiction is to decide not to look for any other justification over and above the text itself for the truth of the statements we read” (Santambrogio, 1992, p. 302; my translation).

is no way of assessing that thing either as true or as false. Alternatively put, any fiction  $F$  is ontologically *incomplete*, in the sense that for some pairs of states of affairs  $S$  and their complement not- $S$ ,  $F$  does not contain either. In this respect, fiction behaves in the very same way as mathematics does according to constructivists: there is no fact of the matter as to whether a mathematical series, or a numerical progression, proceeds in a certain way or not, until some calculation appropriate to the matter has been performed. As Wittgenstein once brilliantly put it:

But what are you saying if you say that one thing is clear: either one will come on  $\varphi$  in the infinite expansion [of  $\pi$ ], or one will not?

[...] What if someone were to reply to a question: So far there is no such thing as an answer to this question?

So, e.g. the poet might reply when asked whether the hero of his poem has a sister or not—when, that is, he has not decided yet anything about it. (Wittgenstein, 1978, V, §9)

## **2 Is There a Way of Grounding Fictional Vichianism?**

Intuitive as it may seem, one may well wonder whether (FV) is correct. Couldn't one entertain the apparently counterintuitive idea that what a piece of fiction says is true just because things so stand in the world that the piece of fiction contributes to select? Moreover, is this idea really so counterintuitive? Consider:

(2) The Earth is round.

Definitely, (2) is not only a real truth, but also a fictional truth, for instance as far as the Holmes stories are concerned. Yet, (2) is not explicitly stated by Doyle, and it cannot be inferred either from what Doyle explicitly stated, or from the environment that what he explicitly stated contributes to sketch—a certain 19th-century Britain, say. One would thus be tempted to say that (2) is true because this is how things are in the worlds of Doyle's stories, not because of any (explicit or implicit) stipulation on

Doyle's part. It thus seems that (FV) stands in need of justification. But what might justify it?

In *Naming and Necessity* (1980), Kripke seems to espouse (FV). A sentence like (1) does not come out true because it turns out that in the outer reality there is a Holmes-like detective. Even if by sheer coincidence it turned out that in the outer reality things were just like in the Holmes stories, Kripke says, this would not make the stories true for all that.

According to Kripke, there precisely is a specific reason why (FV) is correct. Since in general truth depends on reference and in particular a fiction-involving name does not refer to any real life individual, however similar to the corresponding character whom the relevant fiction is about, a fiction-involving sentence containing that name is not made true by what happens to that individual. In particular, "Holmes" does not refer to any Holmes-like concrete real detective; thus, (1) is not made true by what happens to *that* detective. In Kripke's own words:

The mere discovery that there was indeed a detective with exploits like those of Sherlock Holmes would not show that Conan Doyle was writing *about* this man; it is theoretically possible, though in practice fantastically unlikely, that Doyle was writing pure fiction with only a coincidental resemblance to the actual man. (Kripke, 1980, p. 157)

From Kripke's way of putting things, two consequences apparently follow. On the one hand, it may turn out that fictional truths are more than they *prima facie* seem, since we often erroneously believe that we are referring to a concrete real individual, while in actual fact we are make-believedly referring to something. Typically, this happens in the case of myths—by means of the name "Santa" we don't refer to a concrete real man, say, an ancient Italian bishop from Bari, as the referential chain linking together contemporary uses of that name does not lead back to that man—, or in the case of fake stories—by means of the name "King Arthur", *pace* Geoffrey of Monmouth, the author of the *Historia Regum Britanniae*, we do not refer to a 5th-century Romano-Briton conductor, as the referential chain linking together contemporary uses of that name does not lead back to that man (if there ever was such a man). All of this amounts to the relatively trivial discovery that not *all* fiction is intentional. Yet, on the other hand, at least *some* fiction is. As a matter of fact, it cannot turn out that what



we take to be a fictional truth is a real one. This is because it cannot turn out that we erroneously believe that we are make-believedly referring to something, while we are actually referring to a typically concrete real individual. (1) is a case in point. It cannot turn out that (1) is a real rather than a fictional truth, because it cannot turn out that, *pace* Doyle's opposite convictions, by "Holmes" he was referring to a concrete real individual.

Now, the first consequence is hardly contestable—witness the discovery any Western child ends up making, that Santa does not exist, that he is no concrete real individual. The second consequence, by contrast, sounds rather perplexing. If we can discover that a tale we took to be real is fictional, hence that a certain referential chain for a name does not end in a concrete real individual but, to use Donnellan (1974)'s terminology, in a block—say, the "Santa" myth, or the *Historia Regum Britanniae*—why couldn't we make the opposite discovery that a tale we took to be fictional is real, hence that a certain causal chain for a name does not end in a block but in a concrete real individual? Consider the "Vinland" tale, for instance. For a long while, people have believed that Erik the Red's report concerning a land named "Vinland" was a fictional tale about an imaginary land. Recent archaeological discoveries, however, have revealed traces of Viking villages on the Eastern side of the Canadian coast. This has shown that "Vinland" was not used by Erik and his companions as a name for an imaginary land, but was simply those Vikings' name for (at least part of) Northern America. Thus, the "Vinland" tale was not fictional at all.

The morale of this doubt is that there is no way of showing that what is taken to be an "empty" referential chain, that is, a chain ending in a block, really is such. If so, then there is no way of proving that the corresponding truth, which *prima facie* is a truth in virtue of its being stipulated as such, really has that character. Hence, (FV) is threatened.

On behalf of Kripke, one may reply to this doubt as follows. Of course we can mistake a real tale for a fictional one, insofar as we can mistake a "full" referential chain for an "empty" one, as in the case of "Vinland". Yet it remains that, once we establish that a tale has been made up by a certain author, it cannot turn out that that tale is real, insofar as the referential chain that author inaugurates definitely is an "empty" chain. As is the case with the "Holmes" tales. We know for sure that those tales are not reports about a concrete real individual: they were entirely made up by Doyle. This in turn depends on the fact that the name "Holmes" did not

name one such individual: it was altogether coined by Doyle in writing down the story he was creating. To put it otherwise, it may well turn out that we were wrong about the fictional character of the “Vinland” tale. Yet, it cannot turn out that we were wrong about the fictional character of the “Holmes” tales. For while even the author of the “Vinland” tale might have been mistaken on the origins of the name “Vinland”—Erik might have erroneously assumed that he was hallucinating while shouting “Here’s Vinland!”—, no such confusion might have ultimately affected Doyle, the author of the “Holmes” tales. So, there is no reason for being ultimately skeptical on the genuinely “empty” character of *any* putatively “empty” referential chain. As a result, there *are* cases in which (FV) is justified.

Yet, this reply does not seem to be enough either. For one may well hold that the real problem with Kripke’s justification for (FV) is not epistemological. Let us assume that there are, in Evans’ (1982) terminology, *existentially creative* make-believe games, namely, games in which one make-believes *that* there is an individual (named in a certain way) that does a lot of things. The “Holmes” tales are precisely cases of such make-believe games: once upon a time, there was an individual named “Holmes” that was a detective, and so on and so forth. Now, Evans distinguishes existentially creative make-believe games from *existentially conservative* make-believe games, namely, games in which, *of* a certain already constituted individual, one make-believes that that individual does a lot of things. The “Napoleon” tale in *War and Peace* is a typical example of an existentially conservative make-believe game: *of* Napoleon, that is, of the full-blooded French emperor, Tolstoy make-believes that, say, he had a certain toilette during his Russian stay. Now, the point is that an existentially creative make-believe game may well be grounded in an existentially conservative make-believe game. As psychoanalysts repeatedly tell us, one such grounding frequently occurs in dreams: a creative dream, such as a dream in which one fantasizes being raped by a fascinating woman, may well turn out to be a conservative dream *of* one’s mother disguised as such a woman. Yet the same often happens with hallucinations and illusions. As Donnellan (1966, p. 296) once made clear, cases in which one simply hallucinates that, say, that man over there has a walking stick while there is absolutely nothing over there, are quite rare; the most typical case is the illusory situation in which one mistakes, say, a rock for

a man and thinks that that man—which actually is a rock—has a walking stick. Thus, in this case, *of* a rock one has illusorily thought that it had a walking stick. So, it is no wonder that the same kind of grounding happens in purely make-believe games too. To borrow an example from Walton (1990), in make-believing that there is a bear grudgingly approaching, one ends up make-believing of a stump that it is so approaching. Thus again, can't it be the case that an indisputably "empty" referential chain is ultimately grounded in a "full" referential chain? And if this is the case, doesn't this leave (FV) without a real justification?

A caveat. The problem I just raised for (FV) is completely independent from the semantic account we give of sentences involved in existentially creative make-believe games. Let me explain.

As I have hitherto put it, from the linguistic point of view the distinction between existentially creative and existentially creative make-believe games seems to trace back to a well known distinction between a *de dicto* and a *de re* reading of a sentence involving an intensional operator, a make-believe, or pretense operator in such a case. Like any other sentence involving this kind of operator, a sentence of the form "S pretends that *a* is *F*" can be read either *de dicto*, as: "S pretends that: *a* is *F*" or *de re*, as: "of *a*, S pretends that it is *F*". If this were the case, the problem raised by the present opponent of (FV) would simply be the old Quinean problem, according to which in order for a *de dicto* reading of one such sentence to be true, some *de re* reading or other of that sentence must be true as well (see Quine, 1971, p. 106).

At this point, a defender of (FV) might simply reply that, once we endorse a directly referential account of proper names, according to which the semantic contribution of a (token of a) name to the truth-conditions of the (tokened) sentence in which it occurs is exhausted by its referent, the problem no longer occurs. For, once we use an allegedly empty proper name, e.g. "Holmes", in an existentially creative make-believe game, it becomes difficult for a "direct reference" theorist to accept that this use is accounted for by a *de dicto* reading of a sentence of the above form, in which a pretense operator occurs: say,

(3) S pretends that: Holmes is a detective.

For if that name *is* empty, then it simply makes *no* truth-conditional contribution to a sentence containing it, and this, on *any* reading.

As a matter of fact, it would be hard to circumvent this problem by claiming that on the *de dicto* reading, such a name is synonymous with a definite description, so that the embedded sentence in which that name occurs makes a Russellian, descriptive truth-conditional contribution to the sentence embedding it (on that reading)<sup>2</sup>. For not only is that partial adoption of descriptivism scarcely justifiable. What could the justification be for claiming that proper names come in two semantic categories, the full ones which are genuine directly referential devices, and the empty ones which are disguised definite descriptions? But also, that adoption would implausibly sever semantics from semantic competence. As we have just seen, we often fail to recognize either that a proper name is full or that it is empty. So, if partial descriptivism held, a consequence would be that we do not often recognize which semantic category a proper name belongs to—the category of a genuine directly referential device or that of a definite description.

To be sure, the intensionalist defender of the *de dicto* reading of sentences like (3) is not forced to adopt partial descriptivism. She may opt for a non-descriptivist account of a name like “Holmes”, on which this name could be both (i) empty and (ii) such that the embedded sentence in which it occurs still makes a truth-conditional contribution to the sentence embedding it (on its *de dicto* reading).<sup>3</sup>

However, if we want both to fulfill the above aims (i)–(ii) and still be non-descriptivist in respect of proper names, there is an easier way to do all these things. By following Recanati (2000), we can account for the difference between existentially creative and existentially conservative make-believe games in terms of there being a meaning-relevant vs. a meaning-irrelevant context-shift in the simple sentence at issue containing the name involved in those games. By “context”, I here mean a traditional Kaplanian context, namely a certain set of parameters providing a sentence endowed with linguistic meaning with determinate truth-conditions.<sup>4</sup> Thus, in the former case, the “Holmes” case, a “Holmes” simple sentence, say (1), will have no truth-conditions when linked to a real context having the real world as its “world” parameter, in which

<sup>2</sup> One such proposal can be traced back to Currie (1990).

<sup>3</sup> Perhaps applying to this case the analysis provided by Sainsbury (2009, pp. 36–8).

<sup>4</sup> See Kaplan (1989). This context is what people ordinarily labels “narrow context”, while Predelli (2005) calls it “index”.

“Holmes” refers to nothing, but will have genuine truth-conditions when linked to a pretend context having the world of the Holmes tales as its “world” parameter, in which “Holmes” refers to a concrete individual who is a detective, in conformity with the corresponding existentially *creative* make-believe game. A meaning-relevant context-shift will thus have occurred. In the latter case, the “Napoleon”-case, a “Napoleon”-sentence will express the one and the same truth-conditions both when linked to a real context in which “Napoleon” refers to the French emperor, and when linked to a pretend context characterized by an existentially *conservative* make-believe game in which “Napoleon”, again, refers to the French emperor. A meaning-irrelevant context-shift will thus have occurred.

Granted, if the directly referential approach to proper names is correct, this “context-shift” account of the difference between existentially creative and existentially conservative make-believe games fares better than the above intensionalist account, in its descriptivist as well as in its non-descriptivist version. Unfortunately, this is unlikely to satisfy the opponent of (FV). For the problem that originally arose remains untouched.

Indeed, the opponent of (FV) could simply say the following. Let us accept that the idea that a certain effectively empty name make-believable refers to something in an existentially creative make-believe game can be accounted for by saying that there is a pretend context in which that name refers to something. Yet, in order for that pretend context to subsist, it must again be grounded in a real context where that very name is not empty, but really refers to something. So, again, no justification for the claim that fictional truths are made-up truths has been found yet.

### **3 Creationism as the Solution to the “Grounding” Problem**

As is well known, creationism about fictional entities maintains that fictional characters are mind-dependent entities, i.e. abstract entities that are created in virtue of some *lato sensu* mental activity, typically the game of make-believe originally performed by a story-teller and subsequently shared by her audience.<sup>5</sup> Creationism comes in different versions; for the

---

<sup>5</sup> See Predelli (1997, 2002), Salmon (1998), Schiffer (1996, 2003), Thomasson (1999, 2003a, b).

time being, let us assume that creationism is correct in some version or other.<sup>6</sup>

Now, if this is the case, then we have found a way to legitimize Kripke's justification for (FV). The reason for this is that even after a fictional individual has been generated, it can still be the case that one erroneously believes to be referring to a concrete real individual while she is referring to that *fictum*. As a matter of fact, the generation of a fictional entity need not be transparent. This is precisely the case with Santa, or with any other mythological entity: an individual—a child—can think that the individual she is referring to by “Santa” is a concrete real individual, whereas in fact she is referring to a mythological character. As we saw before, this is also true of certain properly fictional entities, such as King Arthur. Yet it cannot be the case that *by using the name* of the generated *fictum*—“Santa”, “Arthur”, or even “Holmes”—, one erroneously believes to be referring to that *fictum*, whereas in point of fact one is referring to a concrete real individual.

The reason is simple. Once a *fictum* has been generated, there is no longer a chance that while using the name involved in its generation—“Holmes”, let us suppose—its user could be an unaware member of an already preexisting “full” referential chain for that name leading back to a concrete real individual. As we saw, this was the risk involved in using the name within an existentially creative make-believe game. Yet, for a creationist, we must distinguish this *pretending* use of that name, in which one merely make-believedly refers to an individual but in point of fact is referring to nothing, from *other uses* of that name, in which one is referring to a fictional entity. First of all, the *hypostatizing* use, the use in which one refers to that entity by characterizing it via features *external* to the fictions that involve it—paradigmatically, the use we typically make of:

- (4) Holmes is a fictional character.

But also the *characterizing* use, the use in which one refers to a fictional entity by characterizing it via features *internal* to the fictions that involve

---

<sup>6</sup> On the different versions of creationism, see my Voltolini (2009). It will immediately be seen that only those versions of creationism which accept that sentences like (1) are true in their characterizing use are the relevant ones. For clearly, a version of creationism that does not accept that cannot support (FV).

it—paradigmatically, the use we make of (1) when we are not engaged in a make-believe game.<sup>7</sup> Now, these latter uses linguistically support the generation of a *fictum*. Thus, by their means, a new “full” referential chain for the name involved has been established. Hence, there is no chance that by so using the name one slides back into an already pre-existing “full” referential chain for that name leading back to a concrete real individual. This, on the contrary, could be the case with the pretending use of sentences like (1): in this case, the “empty” referential chain regarding “Holmes” may always turn out to be grounded in a “full” referential chain for that name involving a concrete real individual. Thus, Kripke’s idea that the fiction-involving name “Holmes” does not refer to any concrete real individual is now vindicated. Hence, his justification of (FV) is legitimate.

The idea that one needs not only realism but more specifically creationism about fictional entities will strike many as surprising. How can it be that we need this kind of position in order to justify the idea that, to put it roughly, there is a gap between fiction and (outer) reality?

But the point is that in order to have fictional truths as a specific kind of truths—to repeat, truths which are such because they are stipulated as such—we precisely need to tell the pretending use from the characterizing use of fiction-involving sentences. For if sentences in the pretending use can be deemed as true, this is because they are true with respect to a make-believe world, namely a world in which things are precisely the way in which the relevant story-teller pretends that they are. This can clearly be seen if we adopt a “context-shift” approach to sentences in that use. Insofar as a sentence being pretendingly used, say (1), amounts to its being linked to a pretend context whose “world” parameter is represented by the relevant world of make-believe, that sentence in that context is true with respect to that world, the world of that context, iff the concrete individual named “Holmes” is a detective in that world. So, sentences in that use are not true because they are stipulated as such, but because things go in a particular way in the world which that use points to. Yet, once we turn to the characterizing use of a sentence, say (1) again, there is no way of explaining the fact that a sentence in that use is true, except by appealing to the idea that this is how the story—viz., a certain set of propositions—goes: in our case, to the idea that the proposition *that the fictional char-*

<sup>7</sup> For the difference between the pretending and the hypostatizing use, see Schiffer (1996, 2003); on the characterizing use, see Barbero (2005).

*acter* Holmes is a detective is part of the proposition set that constitutes the Holmes stories.<sup>8</sup> Of course, the story might have gone differently, had Doyle so decided; in which case another sentence in the characterizing use would be true insofar as the corresponding proposition would be part of the distinct proposition set that would then constitute the Holmes stories.<sup>9</sup>

Now, this distinction between a pretending and a characterizing use of fiction-involving sentences is typical of any *good* creationist position on fictional entities. Indeed, a creationist who limits herself to acknowledging the pretending and the hypostatizing use of a fiction-involving sentence will owe us an extra-justification for why we have to accept fictional entities in our general inventory of what there is.<sup>10</sup> In other terms, if we had to admit *ficta* only because we have an hypostatizing use of certain sentences, we'd better buy an antirealist paraphrase of that use.<sup>11</sup> Thus, my original claim may be refined as follows. If one endorses a good creationism on fictional entities, one that allows for a distinction between a pretending and a characterizing use of fiction-involving sentences, one can ground Kripke's reason for accepting (FV), hence can ground (FV) as well. Since Kripke (1973) seems to endorse both creationism and that distinction, this should be for him a welcome result.

---

<sup>8</sup> By "story" I intentionally mean something different from what I meant by "tale" before. I distinguish between fictional *tales*, which are made of pretendingly used sentences, and fictional *stories*, which are made of the propositions expressed by characterizingly used sentences. For more on this distinction, see my Voltolini (2006, 2009).

<sup>9</sup> Incidentally, there would no improvement if an antirealist accepted that there is a characterizing use of sentences like (1) but adopted an intensionalist account of that use, by saying that in such a use a sentence like (1) is elliptical for a sentence of the form "in the Doyle's stories, Holmes is a detective" on its *de dicto* reading. For even in that case, a sentence containing an intensional operator would be true iff its embedded sentence were true in an unreal world. Thus in the case of a *de dicto* reading of the former sentence, the sentence on that reading would be true either descriptivistically, i.e. iff the denotation of the embedded singular term in the unreal world had in that world the property expressed by the embedded predicate—see Lewis (1979)—or non-descriptivistically, i.e. iff the embedded sentence, as uttered within a fictional context, hence when pretendingly used, were true in the world of that context—see Walton (1990). As a result, in either case the idea that a sentence in a characterizing use is true *stipulatively* would simply vanish.

<sup>10</sup> I am unsure as to who, among the traditional creationists, can be ranked within the bad ones. Possibly inspired on this concern by Schiffer (1996, 2003), Thomasson (2003a, b) seems to give up her previous idea (1999) that in what I take to be characterizingly used sentences, names like "Holmes" refer to *ficta*.

<sup>11</sup> I stressed this point in Voltolini (2006, 2009).



## 4 A Metaphysico-Ontological Comment

As we have seen, (FV) can be ultimately justified by appealing to (good) creationism about fictional entities. That is, sticking to (FV) requires (good) creationism about *ficta*. This may also be seen as an argument in favour of a metaphysically moderate version of ontological realism about *ficta*, as (good) creationism is.

Let me adhere to a distinction between *metaphysics*, as the doctrine which studies the nature of given kinds of things, provided that there are any, and *ontology*, as the doctrine which studies whether in the general inventory of what there is, there really are things of a certain kind.<sup>12</sup> Armed with this distinction, on the one hand I may say that from the ontological point of view we need to be realists if we want to stick to (FV). For limiting oneself to accepting the pretending use of fiction-involving sentences, as antirealists often do,<sup>13</sup> leaves (FV) as an arbitrary principle, insofar as that use is unable to justify it<sup>14</sup>. Yet on the other hand, from the metaphysical point of view, we don't need a radical realism, according to which *ficta* are mind-independent things, for instance some kind of Platonistic entities (sets or properties, Platonic attributes)<sup>15</sup>, or inhabitants of worlds different from the actual one.<sup>16</sup> The reason for this is that a radical realist does not acknowledge the pretending use of fiction-involving sentences, but instead assimilates this use to the characterizing use. By so doing, however, she simply fails to see the problem affecting (FV) as a genuine problem. This way of putting things is too dogmatic. Now, by conceiving *ficta* as mind-dependent entities, (good) creationism is a metaphysically moderate form of realism. Since its avoidance of the Scylla of antirealism and the Charybdis of radical realism also allows (good) creationism to treat the problem that faces (FV), the metaphysically moderate form

---

<sup>12</sup> For this distinction, see, for instance, Thomasson (1999).

<sup>13</sup> Or even reducing the hypostatizing use to a form of pretending use, as inveterate antirealists do. Cf. Walton (1990).

<sup>14</sup> This would also be the case if the antirealist accepted that there is a characterizing use of fiction-involving sentences but claimed that such a use can be accounted for in intensionalist terms. See fn. 9 above.

<sup>15</sup> See Parsons (1980) for the first and Zalta (1983) for the second alternative.

<sup>16</sup> See, again, Lewis (1978) for a "variable-domain", and Priest (2005) for a "fixed-domain" conception of this variant of radical realism.

of ontological realism that (good) creationism constitutes seems to be the best position about fictional entities.<sup>17</sup>

---

<sup>17</sup> Previous versions of the paper have been presented at the Roundtable on *Creationism vs. Fictionalism on Fictional Entities*, at the XXII World Congress of Philosophy, Seoul National University, July 30-August 5, 2008, Seoul, and at the LOGOS *Workshop on Fiction*, University of Barcelona, December 1-2, 2009, Barcelona. I thank all the participants for their stimulating comments.

# Creating Non-Existents

GRAHAM PRIEST

**Abstract.** *Towards Non-Being* gives a noneist account of the reference of words which do not refer to existent objects—in the context, in particular, of intentional states. The account is a realist one, in the sense that the domain of objects is the same at each world, and so does not depend on the behaviour of objects which exist there. In this paper, I discuss an anti-realist version of the theory. What non-existent objects are available at a world supervenes on the behaviour of the existent—and, particularly, sentient—beings at that world. An appropriate formal semantics is given; and its philosophical ramifications—notably, with respect to the naming of non-existent objects—are explored.

## 1 Supervenience and the Non-Existent

Some objects do not exist: purely fictional objects, like Holmes and Anna Karenina; objects of various intentional states, such as worshipping (God—any of the ones you don't believe in); failed objects of scientific postulation, such as the planet Vulcan. One may certainly contest this noneist claim; but I have defended it in *Towards Non-Being*,<sup>1</sup> and I shall assume it without further argument in what follows.

What concerns me here is this. Many people feel that non-existent objects depend, in some sense, on existent objects. Without Leo Tolstoy, no Anna Karenina; without Arthur Conan Doyle, no Holmes. The non-existent supervenes on the existent, and particularly the story-telling ac-

---

<sup>1</sup> Priest (2005); hereafter, TNB.

tivities of cognitive agents (“story telling” being understood in a suitably general sense).

Now, there is a quite trivial sense in which the non-existent supervenes on the existent. Sherlock Holmes does not live in Baker St, and never did. (No camping out in Baker St would ever have spotted him.) Yet, in the stories of Doyle, Holmes lived in Baker St. *That* is a property of Holmes. And that property does indeed depend on Doyle’s actions. Had Doyle not written his stories, Holmes would not have had the property of living in Baker St in them.

But many have felt that the dependence of the non-existent on the existent is deeper. It’s not just that Doyle is responsible for the appropriate properties of Holmes; in some sense, Doyle *created* Holmes. Though this is not a feeling by which I am particularly moved,<sup>2</sup> I want to explore the extent to which it can be accommodated in a noneist picture.

Of course, the one thing one can’t say concerning creation is that Doyle brought Holmes into existence. Fictional objects and their like do not exist at all. *A fortiori*, they cannot have been brought into existence. Yet this could still be the case: The domain of quantification contains existent objects plus some non-existent ones. What non-existent ones are in the domain does indeed supervene on the actions of the existent ones there; in particular, the story-telling ones. Thus:

**(H)** had Doyle not written his stories, nothing would have been Holmes.

In this sense, the non-existent objects depend on the existent ones.

This is clearly a species of anti-realism about non-existent objects. Their availability depends, in some sense, on the cognitive states of existent objects. In what follows, I want to consider this view, its coherence, and its implications. I will do this in two parts. First, I will provide a semantics which accommodates the view, and makes it possible to formulate and satisfy the appropriate supervenience constraint. Then, armed with the semantics, we will look at some philosophical implications of the view.

---

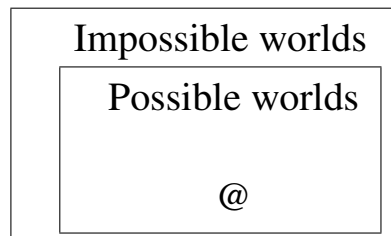
<sup>2</sup> It is quite natural, it seems to me, to hear this as a metaphor, which may be cashed out in various ways. One might, for example, understand it as some kind of stipulation, as does Deutsch (1991).

## 2 Formal Semantics

### 2.1 The Semantics of TNB

TNB gives a noneist world-semantics for intentional predicates, such as “pity” (as in “I pity Karenina”), and intentional operators, such as “fears that” (as in “The reader of *The Final Problem* fears that Holmes has fallen to his death over the Reichenbach Falls”), or “in the novel it is the case that” (as in “In *The Final Problem* Holmes does indeed fall over the edge of the falls, but survives”). All the worlds have the same domain—and so the semantics are not compatible with the thought that the constitution of the domain of a world supervenes on what exists there (which may vary from world to world). In this section I will describe these semantics. In the next, I will describe the modifications which give rise to a variable-domain version of the semantics, in which supervenience can be satisfied.

First, the worlds. Some of these are possible; the actual world, @, is one of them. The others are not. Impossible worlds are required since we can have intentional states directed towards impossibilities. For example, I can wish to square the circle, or dream that my father is my mother. Impossible worlds are worlds that realise the contents of such intentional states.<sup>3</sup> So we have the following picture.



Because we are dealing with worlds some of which are impossible, negation cannot behave classically everywhere. In particular, there must be inconsistent and incomplete worlds. So statements may be true or false at a world; but these states of affairs are independent: both or neither may obtain. Hence we need to give separate truth and falsity conditions. Write  $w \Vdash^+ A$  to mean that  $A$  is true at world  $w$ , and  $w \Vdash^- A$  to mean that  $A$  is false at  $w$ . The truth/falsity conditions of atomic sentences are natural:

<sup>3</sup> TNB further distinguishes, within the category of impossible worlds, between open worlds and closed worlds; but this matter need not concern us here.

$$w \Vdash^+ P c_1 \dots c_n \text{ iff } \langle \delta(c_1), \dots, \delta(c_n) \rangle \in P_w^+$$

$$w \Vdash^- P c_1 \dots c_n \text{ iff } \langle \delta(c_1), \dots, \delta(c_n) \rangle \in P_w^-$$

where  $\delta(c)$  is the denotation of the constant  $c$ ,  $P_w^+$  is the extension of the  $n$ -place predicate  $P$  at  $w$  (the set of things of which it is true there), and  $P_w^-$  is the anti-extension (the set of things of which it is false). The truth/falsity conditions of the connectives are much as one would expect. In particular:

$$w \Vdash^+ \neg A \text{ iff } w \Vdash^- A$$

$$w \Vdash^- \neg A \text{ iff } w \Vdash^+ A$$

Every world comes with a (the same) domain of objects,  $D$ . At each world, an object may or may not exist. Thus, there is a monadic existence predicate,  $E$ , whose extension at world  $w$  is the set of thing that exist there. Quantifiers,  $\mathfrak{A}$  (all) and  $\mathfrak{S}$  (some), work in the standard way. Let us assume that the language is augmented by a set of constants  $\{k_d : d \in D\}$ , such that  $\delta(k_d) = d$ .<sup>4</sup>  $A_x(c)$  is  $A$  with all free occurrences of  $x$  replaced by  $c$ . Then:

$$w \Vdash^+ \mathfrak{S}x A \text{ iff for some } d \in D, w \Vdash^+ A_x(k_d)$$

$$w \Vdash^- \mathfrak{S}x A \text{ iff for all } d \in D, w \Vdash^- A_x(k_d)$$

Dually for  $\mathfrak{A}$ . In other words,  $\mathfrak{S}x A$  holds at  $w$  just if *something* in  $D$  satisfies  $A$  at  $w$ , and  $\mathfrak{A}x A$  holds at  $w$  just if *everything* in  $D$  satisfies  $A$  at  $w$ . Note that the particular quantifier,  $\mathfrak{S}x$ , should be read “some  $x$  is such that”. (And the universal quantifier  $\mathfrak{A}x$  as “all  $x$  are such that”.) It is *not* to be read as “there exists an  $x$  such that”, or even as “there is an  $x$  such that”. If one wants to say such things, one has to use the existence predicate explicitly, thus:  $\mathfrak{S}x(Ex \wedge A)$ . An object that does not exist at a world does not have some lesser grade of being there. If it does not exist (at a world) it simply *is not* (there).

Given this set-up, the semantics of intentional predicates are simple. Intentional predicates work in exactly the same way as any other predicate. The extension of “kicks” (at a world) is just the set of pairs such that the first kicks the second (there); the extension of “fears” at a world is just the set of pairs such that the first fears the second (there). Thus, when John

---

<sup>4</sup> In TNB I treated quantifiers differently, defining satisfaction. The present method is equivalent and simpler.

fears something, this is a relationship between John and the object of his fear. John has immediate phenomenological acquaintance with the object; but the object itself may or may not exist.

The semantics for intentional operators are those which are standard in the world-semantics of epistemic and doxastic logic. For every intentional operator,  $\Psi$ , in the language, there is a corresponding binary accessibility relation,  $R_\Psi$ , between worlds.  $wR_\Psi w'$  means something like: at  $w'$ , things are as, at  $w$ , they are  $\Psi$ d to be. So if  $\Psi$  is “John fears that”,  $@R_\Psi w'$  iff  $w'$  realises all the things that John actually fears.<sup>5</sup> We then have:

$w \Vdash^+ \Psi A$  iff for all  $w'$  such that  $wR_\Psi w'$ ,  $w' \Vdash^+ A$

$w \Vdash^- \Psi A$  iff for some  $w'$  such that  $wR_\Psi w'$ ,  $w' \Vdash^- A$

Other details of the semantics in TNB need not concern us here, except for a mention of Characterisation. If one characterises an object in a certain way (say, as a Victorian detective of acute powers of observation and deduction, etc.), one has no guarantee that the object in question really does have those properties (at the actual world). It does have those properties at some worlds, however; namely, those that realise the situation about the object envisaged (e.g., the ones described in Doyle’s Holmes stories).

Formally, one can represent characterisation with an indefinite description operator,  $\varepsilon$  (so that  $\varepsilon xA$  is read as “a thing,  $x$ , such that  $A$ ”). The denotation of  $\varepsilon xA$  is then determined as follows. There are worlds where  $\mathfrak{S}xA$  holds. (Given impossible worlds, every condition holds at some world.) Let  $w$  be @ if the condition holds there; otherwise, choose some other world. Let  $d$  be any of the objects satisfying  $A$  at  $w$ . (These two steps can be accomplished, formally, by applying suitable choice functions. Informally, they are realised by an intentional act. More of this anon.) Then  $\delta(\varepsilon xA) = d$ . This ensures that if  $@ \Vdash^+ \mathfrak{S}xA$  then  $@ \Vdash^+ A_x(\varepsilon xA)$ . Moreover, let  $\Psi$  be an intentional operator with an accessibility relation satisfying the following condition:  $wR_\Psi w'$  iff  $w' \Vdash^+ A_x(d)$ . One might think of  $\Psi$  as “so and so is thinking that  $d$  satisfies  $A$ ”. Then  $@ \Vdash^+ \Psi A_x(\varepsilon xA)$ .<sup>6</sup>

<sup>5</sup> In the cause of perspicuity, I simplify here—though not in TNB itself—incorporating the agent into the intentional operator.

<sup>6</sup> This is slightly different from the account given in TNB, but cleaner, and has the same effect. Again, I simplify in certain ways irrelevant here.

## 2.2 Variable-Domains and Supervenience

Reformulating the semantics to allow for variable domains requires only minor modifications.

$D$  is still the domain of all objects (in which all constants find their denotation). But now each world,  $w$ , has its own domain,  $D_w$ , a subset of  $D$ . Intuitively,  $D_w$  comprises those objects that exist at  $w$  plus those non-existent objects that get in in virtue of the actions of those that do exist. The extension and anti-extension of a predicate are subsets, not of ( $n$ -tuples of)  $D$ , but of  $D_w$ . All truth and falsity conditions are the same, except that in those for the quantifiers, “ $D$ ” is replaced by “ $D_w$ ”.

If  $c$  does not denote an object in  $D_w$  then, according to these semantics, any atomic sentence involving it (even  $Ec$ ) is neither true nor false there.<sup>7</sup> However, another strategy is to let such sentences simply be false. This requires changing the falsity conditions for atomic sentences to:

$$w \Vdash^- Pc_1\dots c_n \text{ iff } \langle \delta(c_1), \dots, \delta(c_n) \rangle \in P_w^-, \\ \text{or for some } 1 \leq i \leq n, \delta(c_i) \notin D_w.$$

There are, moreover, reasons to prefer this strategy. For example, we may wish to consider the consequences of an author not creating a character; and in worlds where this is the case, the object had better not be in the domain of quantification there. Bearing this in mind, return to the conditional  $H$ . This is of the form:

$$A > \neg \mathfrak{S}xx = h$$

where  $>$  is a suitable counterfactual conditional. To evaluate it (at @), we go the nearest worlds where  $A$  is true (so where Doyle did not write his stories). In realist, constant-domain, semantics  $\mathfrak{S}xx = h$  is true at such worlds, so  $H$  is false—which seems right. In the variable-domain semantics, in worlds where  $A$  is true, according to the original truth/falsity conditions,  $\mathfrak{S}xx = h$ , and so its negation, is neither true nor false, and so the conditional is still not true. With the alternative truth conditions,  $\mathfrak{S}xx = h$  is false, and so its negation is true, as, then, is  $H$ , which seems right.<sup>8</sup>

<sup>7</sup> So, unlike TNB, there will be truth value gaps even at possible worlds.

<sup>8</sup> This particular example could be accommodated by just changing the falsity conditions of the identity predicate. But this would seem not to be enough. Take for  $A$ , instead: Doyle



Given these semantics, we can now formulate the supervenience condition precisely. Given an interpretation, say that two worlds,  $w_1$  and  $w_2$ , are *identical with respect to a set of objects,  $X$* , if the extensions and anti-extensions of all predicates at the two worlds are the same with respect to all members of  $X$ .<sup>9</sup> The supervenience of the non-existent on the existent can now be formulated as follows. (Recall that  $E$  is the existence predicate.)

**(Supervenience)** If  $E_{w_1}^+ = E_{w_2}^+ = X$ , and  $w_1$  and  $w_2$  are identical with respect to  $X$ , then  $D_{w_1} = D_{w_2} = Y$ , and  $w_1$  and  $w_2$  are identical with respect to  $Y$ .

Imposing this constraint on interpretations completes the picture.<sup>10</sup>

### 3 The Anti-Realist Picture

#### 3.1 Creation and Naming

The semantics of the previous section suffice to show that the anti-realist supervenience picture is technically coherent. This does not show that it is philosophically coherent, of course, or that it is philosophically preferable. In what follows, I will explore these matters.

Let us start with what seems to me to be the most significant philosophical difference between the realist and the anti-realist accounts. This concerns the naming of non-existent objects.

Assume something like the causal theory of names (TNB, 7.5). As in standard theories of this kind, an object is baptised with a name by some agent. The referent of the name is then picked up by any person with whom they interact causally in a certain way, any person with whom *they* interact, and so on. The object for baptism may be singled out by a

---

did not write about Holmes. With the old falsity conditions, in a world where this is true, Holmes is in the domain of quantification, and so  $H$  is again false. With the revised conditions, it comes out true, since Holmes does not have to be in the domain of quantification to make  $A$  true.

<sup>9</sup> And if, as in TNB, the accessibility relations are also indexed by members of  $X$ , these need to be the same too.

<sup>10</sup> One might, of course, insist that the supervenience base comprise a *particular class* of predicates, such as those concerning the content of narratives. Supervenience can obviously be reformulated in this way.

definite description, or by pointing. What the story about intentionality of TNB adds to this picture is that the pointing may not only be physical—non-existent objects cannot, of course, be pointed to physically—it may be mental. That, after all, is what intentionality is all about. So far the realist and the anti-realist pictures are no different.

It is when we consider the mechanism of the baptism that differences emerge. In the realist case, when Doyle coined the name “Holmes” he gave it to a non-existent object, picked out as an object which was a detective with acute powers of observation and inference, etc., in the worlds that realised the story he wished to tell. This was achieved with an act of mental pointing; and, realistically conceived, the object was available to be pointed at. But how does the pointing work? How does the act pick out one of the enormous number of non-existent objects? In many worlds there are objects—different objects—which are detectives with acute powers of observation and inference, etc. How does the act pick out one of these? The supposed incomprehensibility of this has, in fact, been one of the major objections raised to the account of TNB by commentators, such as Bob Hale.<sup>11</sup> I do not, myself, find a problem with a notion of mental pointing that can do this, any more than I find a problem with a notion of physical pointing that selects an object at random.<sup>12</sup> (Close your eyes and point to someone in a crowd.)

But on the anti-realist account, matters are different. On this account, the domain of objects at a world is not fixed once and for all, but depends on the actions of the objects that exist there. Thus, Sherlock Holmes would not have been in the domain of objects at the actual world had it not been for the story-telling activities of Doyle. Holmes’ baptism, on this account, is quite different. Doyle created Holmes. He did not, of course, bring him into existence; Holmes does not exist. But it was in virtue of Doyle’s story-telling activities that Holmes came to inhabit the domain of objects of the actual world. Doyle’s cognitive activities resulted in the expansion of the domain of objects that were available for reference and quantification. Since Doyle’s phenomenology did not select Holmes, but created him, the question of how the selection was possible does not arise.

---

<sup>11</sup> Hale (2007).

<sup>12</sup> See Priest (201+).

It might be thought that this makes an anti-realist account of non-existent objects more plausible than a realist account. If it made redundant an appeal to the need of intentionally pointing to one of a number of linguistically indiscriminable objects, perhaps it would. Unfortunately or otherwise, this appears not to be the case. Here's a story. *There were once two men who lived in New York. They were both exactly 6' tall, liked Mozart, and enjoyed going to the movies together.* Let's call them Tweedledum and Tweedlee. I might now ask you questions about them. Was Tweedledum Australian? Answer, we don't know: in some worlds that realise the story, he was; in some he wasn't. To answer the question, you must have given the name "Tweedledum" to one of them. On the anti-realist picture, I created the pair of objects at exactly the same time; but there is still no linguistic information that discriminates the two.<sup>13</sup>

### 3.2 Identity

The realist and the anti-realist accounts also have different implications concerning identity. The criterion of identity given by TNB, 4.4, is a version of Leibnizian indiscernibility: two objects (whether existent or not) are the same just if, at every world, any property of the one is a property of the other.<sup>14</sup> The two accounts can both accept this criterion; but its application may give different results in the two cases.

Suppose that you and I, independently, tell stories about some non-existent character. Suppose that we both call her "Ricki". By happenstance, we say exactly the same about Ricki: the stories are identical. Are we talking about the same object? We are if the two Rickis satisfy the Leibniz condition, but do they?

If one is a realist, there is no general answer to the question of whether or not the two are the same.<sup>15</sup> I selected an object to christen "Ricki"; so did you. In both cases they are objects of which the story is true in certain worlds. We may have chosen the same object (and the same worlds), we

---

<sup>13</sup> There is also the strange case of the two roots of  $-1$  (TNB, 4.4). There is nothing to distinguish between these; but at some stage, some mathematician (or committee of mathematicians) must have decided that one of these was  $+i$  and the other was  $-i$ .

<sup>14</sup> Actually, it is at every closed world. But this subtlety is not relevant here.

<sup>15</sup> If the telling of the stories is not independent, the matter may well differ. If you hear me talking about Ricki, and then tell your own story about her, it is the same Ricki. You have picked up the reference of the name from me, as the causal theory of reference has it.

may not; it just depends. Nor may we ever know. Similarly, if you and I point to people in different historical photographs (say 60 years apart), we may never know whether we have pointed to the same person. Of course, if we were to know everything about them, we would know whether they were the same person. But if we were to know everything about your Ricki and mine, we would also know whether they were the same person. In both cases, the knowledge may be denied us. Such is realism.

The anti-realist situation is different. Ricki and her properties supervene on the activities of her creator. But which exactly? There are a couple of ways one may go here. A natural answer is that she supervenes on the content of the story told. Thus, if the stories are the same, as they are in this case, the Rickis are the same. Alternatively, she may supervene, not just on the content of the story, but on the fact that it was told by the particular teller as well.<sup>16</sup> In this case, the Rickis are different, since the tellers are different. Which of these views is the better, we need not go into here. The important point to note is that *both* give determinate (though different) answers to the question of the identity of the two Rickis, unlike the realist account, which leaves the matter undetermined.

Finally, note what supervenience does *not* mean. It does not mean that a non-existent object, characterised in a certain way, has only those properties it is explicitly characterised as having—or those that follow from these—at the worlds that realise that characterisation. Thus, if Ricki is characterised as being either left-handed or right-handed, then in any world that realises the characterisation, she will be either the one or the other. Supervenience means only that if the two Rickis have the same supervenience base, any world in which the one is left-handed, so is the other.<sup>17</sup>

### 3.3 Other Matters

Let me finish with four more miscellaneous matters.

---

<sup>16</sup> One might subscribe to such a view on the ground of some variety of the doctrine of the essence of origins.

<sup>17</sup> As TNB, 6.4, explains, objects may have properties beyond those they are characterised as having—or that follow from these—in the worlds that realise the characterisation. The realist/anti-realist distinction does not affect this matter.

First, if one can create objects, then, presumably, one can create groups of objects. Thus, in writing *Julius Caesar*, Shakespeare created not only Mark Anthony, but the howling crowd,  $C$ , which he addresses after Caesar's assassination. Did he create all the members of the crowd? Well, crowds certainly have members; and in every world that realises Shakespeare's story  $C$  exists and has members. In different worlds,  $C$  may have different members, and the members of  $C$  must be things which exist at that world, and so members of the domain there. But it does not follow that Shakespeare created each one of them.<sup>18</sup> It is up to the director—so to speak—at each world, to decide who it is that goes into the crowd.

Next, it might well be thought that just as non-existent objects can be created, so they can be destroyed—or at least lapse into oblivion. Let us suppose, for example, that a community has a belief about an object which does not exist, perhaps a god of some kind. Let us suppose also that gradually, over generations, the belief lapses, and all reference to the object is forgotten (no written records, no living memories, etc.). Then at least arguably, the community has lost the ability to refer to that object, and it is no longer in the domain of quantification.

Third, what consequences does this view have for worlds themselves? That depends. If we take worlds to be existent objects of some kind, then none whatsoever. However, TNB, 7.3, argues that worlds other than the actual are non-existent objects. Hence, if we assume an anti-realist approach to non-existent objects, we must apply it to these also.<sup>19</sup> What worlds occur in the domain of quantification of a world itself supervenes on the activities of the cognitive agents at that world. (Clearly, worlds can be in the domain of quantification of a world: we quantify over them in the actual world.) In particular, the worlds in  $D_{@}$  supervene on the activities of those who theorise about them—us.

Finally, let us consider a possible objection to the anti-realist semantics. It might be suggested that the semantics still has a residual realism. The domain  $D$  comprises a bunch of objects; some of these may not exist (anywhere) and there is no reason to suppose that they are mind-dependent

---

<sup>18</sup> Though one might argue that he created those for which there are speaking parts.

<sup>19</sup> In the same chapter it is also argued that abstract objects, and in particular mathematical objects, are non-existent objects. If this is so, then the debate about realism and anti-realism (constructivism) in the philosophy of mathematics can be thought of as a special case of the debate about the nature of non-existent objects.

in any sense. To avoid this residual realism, one may take  $D$  to be  $D@$ . Again, the semantics is our construction, and the objects in  $D$ , if they do not exist, are objects we have created.

That is:

(\*) All non-existent objects in  $D$  are created by existent ones.

What of the status of this claim? Arguably, is it a necessary truth, and so true in all possible worlds. It does not follow, however, that in *every* world each non-existent object is created by some object that exists at that world. One can characterise an object as non-existent but non-created; so at some worlds this characterisation is satisfied. What follows is that worlds where (\*) fails are impossible worlds.

## 4 The Choice

So much for the two pictures: the realist one, with its constant-domain semantics, and anti-realist one, with its variable-domain semantics and the supervenience condition. I have not tried to decide between the two. All I have done is to put the anti-realist picture on the table, and contrast it with the realist picture of TNB. Some of the considerations relevant to a choice between the two have been briefly traversed here. Many others, I am sure, have not. But at least the space for discussion is now open.<sup>20</sup>

---

<sup>20</sup> Versions of this paper were given to the Melbourne Logic Group, September 2008, and to the conference *Reference and Non-Existence*, held by the LOGOS group at the University of Barcelona, June 2009. Many thanks go to the members of these audiences for helpful discussions, and especially to Mark Sainsbury and Bob Stalnaker. Thanks also go to Franz Berto and Shahid Rahman for helpful discussions.

# Sweet Nothings: The Semantics, Pragmatics, and Ontology of Fiction

FRED ADAMS

## 1 Introduction

Suppose one thinks, as I do (Adams et al., 1992–2007)<sup>1</sup> that on the occasion of use in fictional or non-fictional assertions, names contribute their bearers (and only their bearers) to the semantic content of what they express. So, in the non-fictional “Obama is a pretty sharp cookie” (meaning that he is smart), “Obama” contributes the very man named to the content of what is expressed by the sentence on the occasion of use. It is not uncommon to say that “Obama” contributes Obama to the proposition  $\langle \textit{Obama, being smart} \rangle$ .

The content expressed by “Obama is a pretty sharp cookie” is true if it matches the facts, as it surely does. Similarly this is so for “Obama is young”, “Obama is energetic”, or “Obama has a lovely family”. All are true. And for negative utterances such as “Obama is not tall”, “Obama is not president”, or “Obama is not smart”, all are false because their contents

---

<sup>1</sup> The view I will express here has been developed over many years (see references), and I owe special thanks to all my former co-authors who have worked with me and influenced my current thinking about these matters. This is especially true in the case of Gary Fuller and Robert Stecker.

do not match reality. A negative existential sentence such as “Obama does not exist” too is clearly false, for its content conflicts with reality.

Now suppose this “direct reference” theory of names is true. What happens when there is a name in common use, but there is no bearer? Scientists once speculated that there was a planet named “Vulcan” between Mercury and the sun and that Vulcan was the cause of perturbations in Mercury’s orbit around the sun. Later it was discovered that there was no planet and that the perturbations in Mercury’s orbit were caused by space-time curvature in proximity to the massive sun. What of sentences such as: “Vulcan is a planet”, or “Vulcan is influencing Mercury’s orbit”, or especially the negative existential “Vulcan doesn’t exist”. The first two may seem to say something false and the last one may seem to say something true. But how can this be the case if the direct reference view of names is true?

And consider sentences of fiction. Suppose one utters “Sherlock Holmes was a pretty sharp cookie”, or “Sherlock Holmes lived at 221B Baker Street in London”, or the negative existential “Sherlock Holmes does not really exist”. Some of these seem false and some seem true. But how can they actually *be* as they seem, if the direct reference theory of names is correct?

If some of these sentences, fictional or not, seem true or seem false, either the seemings are mistaken, or perhaps there is something else going on. Perhaps there are actual entities created by assertions of scientists or by fictional assertions of authors, and perhaps these utterances of sentences employing names without presumed types of real-world bearers actually have a different kind of bearer. Perhaps sentences of fiction create a type of entity called a “fictional character”, and perhaps sentences of science similarly create an artefact of scientific theory—call it a “theoretical entity”—and perhaps sentences of fiction and science are *about* these artefacts of human invention. Were that the case, the names would not really be “empty”, though they would have referents that were not actual physical objects. Still, the names could be used to say things that are true and false (of fictional characters or of theoretical entities) and this could explain why such sentences as those above (where some names lack physical objects as referents) can seem to be true or false.

In what follows, I will offer an account that rejects this latter alternative of suggesting that names in science or fiction have referents, just not



physical objects but fictional characters or theoretical entities instead. I will suggest that, on the correct theory of names, names (for things) have physical objects as their referents, if they have referents. And if they have no referents, they have no meaning. Surprisingly, this will entail that the sentences above that employ “empty” names, are neither true nor false. Surprisingly, “Vulcan does not exist” turns out not to express a truth on this view. Nor does “Sherlock Holmes does not exist” express something true or false. In fact, sentences of fiction or non-fiction that are not about names but simply use names, cannot be used to say things true or false, on this view. I will account for the seeming’s (that some of these sentence types can seem true or false) by appealing to the pragmatic features of language and utterance and cognitive features of how the mind works.

## 2 The Semantics of Fiction

Let’s begin with the semantics—the logical form of expressions containing empty names. When “*a*” is an empty name, what is the logical form of an expression of the type “*Fa*” or “ $\neg Fa$ ” or “ $\neg(\exists x)[x = a]$ ”? As I said in opening remarks, the meaning of a name is its bearer on the occasion of use, whether the name occurs in a sentence of non-fiction or in a sentence of fiction.

So let’s take non-fiction first. “Vulcan is a planet” expresses the partially empty  $\langle \text{---}, \textit{property of being a planet} \rangle$ . “Vulcan” attempts to reach out and grab an object and contribute that object to the content of the utterance of the sentence. But there is no object. Of course, the scientists who first proposed the idea that there was a planet so-called, did not know this. So they did not know that they were expressing an incomplete content.<sup>2</sup> They believed they were expressing a complete content with the very planet named as the subject. They happened to be wrong about the planet and the content. There is more to be said about why the scientists believed they were expressing a complete thought (even though they were not), and we will take this up in the next section of this paper.

---

<sup>2</sup> The notion used here of an incomplete or gappy proposition owes much to the work of David Braun (1993), who sent me his paper on this just as Fuller and Stecker and I were developing our criticism of Devitt’s semantics of names. Devitt challenged us (in conversation) to come up with a direct reference account that handles empty names (implying that it couldn’t be done). We accepted the challenge.

We have now seen the logical form of positive sentences employing empty names. What about negative sentences and especially negative existentials? A negative sentence of the form “ $\neg Fa$ ” would express the negation of  $\langle \_, \textit{property of being a planet} \rangle$ . It would do so using the empty name “Vulcan”. This is a crucial fact. There are other empty names. They too can be used to express the same partially incomplete content. The fact that “Vulcan” was used to do it this time is an important etiologi- cal feature of the utterance and its content. This will become particularly important when we turn our attention to fiction, but we can see the import here as well. Suppose scientists came up with another idea for a planet to explain oddities in Pluto’s orbit. Suppose they call it “Ioh”. Suppose they say “Ioh is not bright”, because they cannot see it with their telescopes. Then they express the negation of  $\langle \_, \textit{property of being bright} \rangle$ . But this time they do it using the name “Ioh”. And surely their belief that there is such a planet named is a very large part of why they believe they have asserted something with a very different content. Actually, they would be wrong about that too. They would have asserted the same incomplete content, though they would have intended to assert a different, complete content. We can still say they did something different because they used a different name (“Ioh” not “Vulcan”) and they associated different descrip- tions with the different names. (More about this in the next section.)

Perhaps the most surprising result is when we look at negative exis- tentials, for they surely seem true (a powerful seeming, even to me). Sen- tences of the form “*a* does not exist” cannot express truths, for there is no object being discussed (though one may not know that). The empty name “*a*” cannot reach out and contribute an object for comment to enter the content of the utterance. So “Vulcan doesn’t exist” (as might be uttered by a contemporary astronomer) actually has the logical form:  $\neg(\exists x)[x = \_]$ . We cannot enter the entity named by “*a*” into the empty slot, for recall that in the logical form only the content of the name can be entered. So this position is reserved for the bearer of the name “*a*”, but there is no bearer. Hence, the slot is unfilled. Thus, the expression again yields an incomplete content or proposition (regardless of whether the speaker or author of the sentence knows this or not).

One may retort that the scientists were not thinking of nothing. Yet, in a sense they were, in so far as they used the name “Vulcan”. In a sense, they were not in so far as they used the term “planet” or “is not bright”.

These terms are not empty. They pick out properties that exist and enter them into the content of the expression (an incomplete proposition). Naturally, the scientists had a theory about there being a planet in a certain location in our solar system. They had an idea and perhaps even a visual image or a drawing of such a thing (maybe even mathematical equations depicting how it would influence Mercury's orbit, if it existed). However, surely the term "Vulcan" did not refer to these things in the sentence "Vulcan does not exist", For these things all exist. And just as surely it was not these things that were supposed to make "Vulcan is a planet" true. For none of these things would make that sentence true (nor do they). So, surely the scientists were not using the term to pick out any of these things. They were intending to pick out a planet for comment.

On the view represented here, the semantics of fiction is exactly the same as the semantics of empty terms in non-fiction. Let's stick with names. Since the names are fictional (i.e., empty) they lack referents. (I will have nothing here to say about terms that are not empty but appear in fiction. They get their semantics in the way that any filled name does inside or outside of fiction.) The meaning of any name is its bearer on the occasion of use. No name—No bearer. The semantics for "Sherlock Holmes" parallels that of "Vulcan". If a name lacks a bearer in non-fiction, it does not acquire a bearer just because it is employed for purposes of fiction (or so I shall maintain).

"Sherlock Holmes smokes a pipe" expresses the incomplete content  $\langle \_, \textit{property of smoking a pipe} \rangle$  and is neither true nor false, since it attributes this to no one. There is no actual person named. Indeed, that is part of what made Doyle's Sherlock Holmes novels fictional. They were not intended to be about an actual person named "Holmes".

"Sherlock Holmes does not live at 221B Baker Street" may seem to say something false, but it does not. It expresses the negation of the incomplete content  $\langle \_, \textit{living at 221B Baker Street} \rangle$  (which is, of course, neither true nor false). As with the case of "Ioh" above, "Watson does not live at 221B Baker Street" may also seem to say something false, but does not because it also expresses the negation of  $\langle \_, \textit{living at 221B Baker Street} \rangle$ . Since the latter sentence uses the name "Watson" it would be used with the intention to say something different (within fiction or about fiction) because one would associate different descriptions with the name

“Holmes” than with the name “Watson”. (We shall return to these matters in the section below.)

The negative existential sentence “Sherlock Holmes does not exist” expresses the incomplete content  $\neg(\exists x)[x = \_\_]$ . As with “Vulcan” the object that is supposedly named cannot be supplied, because the name has no bearer. (We shall consider whether there is a *fictional character* named, an ontological referent that could make the negative existential true, in Section 4 below.) Since no object is named, none can enter the content expressed by the negative existential sentence. This is so even though the sentence may seem to express a truth to some (or a falsehood to others).

What about sentences about fiction, rather than sentences within fiction? Consider the sentence: “In the Doyle novels, it is fictionally asserted that Holmes is a detective”. Is this true or false? True, if understood properly. This sentence asserts a content of the form: Using the name “Holmes”, in the Doyle novels it is fictionally asserted:  $\langle \_\_, \textit{being a detective} \rangle$ . This assertion is not directly about the object purportedly named. It is a level up. It is about something Doyle does in creating the fictional work. It is about a fictional assertion of the author or of the work. In so far as it is about this, and semantically interpreted accordingly, the sentence that expresses this content is true. Still, notice that the content of the fictional assertion embedded within the non-fictional sentence, does not itself express a complete content that can be either true or false.

False sentences would be handled similarly. Consider the false: “In the Doyle novels, it is fictionally asserted that Watson is a grocer”. This has the content that, using the name “Watson”, in the Doyle novels it is fictionally asserted that:  $\langle \_\_, \textit{being a grocer} \rangle$ . Since this is not true, the sentence about the work is false (failing to correspond with the actual fictional assertions or implications of the text).

Or consider that in some passages of the novel, it may fictionally be asserted: “Moriarty is now dead and does not exist”. The content expressed would be that using the name “Moriarty”, in the Doyle novels it is fictionally asserted that:  $\langle \_\_, \textit{being dead} \rangle$  and  $\neg(\exists x)[x = \_\_]$ . This sentence about fiction would then be true (if it matched the text in actual wording or implication of the work).

There can also be sentences about fiction that compare and contrast individuals fictionally depicted in one work with those fictionally depicted

in others. I've dealt with these elsewhere (Adams, Fuller & Stecker, 1997) and won't repeat the account here, as it can be seen to iterate the pattern above.

### 3 The Pragmatics of Fiction

Let's begin again with non-fiction. In the case of "Obama is a pretty sharp cookie", we say that "Obama" contributes Obama (the man) to the content expressed. Of course there are many descriptions that people store in their informational and cognitive files about Obama: "The first actually black president", "The former junior senator from Illinois", "The husband of Michelle", "The only president who grew up in Hawaii", and so on. When one asserts that "Obama is a pretty sharp cookie", one may impart information to the effect that some or all of these things are true of him as well, e.g., that the former junior senator from Illinois is a pretty sharp cookie.

Notice that none of these associated descriptions of Obama, though doubtlessly true of him, are part of the meaning of "Obama". As Kripke (1980) taught us, Obama would still be Obama (necessarily), even if none of these other things were true of him. Still, since these other things are (contingently) true of him, we can use them to help keep track of him. We associate these and other descriptions of him in order to characterize him in our minds, in our knowledge of him, and maybe in our interactions with him.

When I was young, I walked into class one November day to hear "Kennedy has been shot". What was expressed was  $\langle \textit{Kennedy, being shot} \rangle$ . What was imparted was "the president has been shot", "the husband of Jackie and father of Caroline and John-John has been shot". More was pragmatically conveyed than was semantically expressed. This is because these associated descriptions were firmly fixed to my (and that of millions of others) informational storage files for John F. Kennedy. Pragmatic conveyance of information is economical and efficient and a great benefit to communication, generally. One need not utter all that one can convey with a single utterance. This is a good thing.

As with filled names, there are files for empty names too and in uttering empty or fictional names, there are pragmatic features at work. When

scientists discovered that there was no planet between Mercury and the sun, they no doubt uttered “Vulcan does not exist”. Using the name “Vulcan”, they semantically expressed only the incomplete  $\neg(\exists x)[x = \_\_\_]$ , which is neither true nor false, but they conveyed much more. They conveyed that there is no planet between Mercury and the sun (true). They conveyed that there is no tenth planet (true). They conveyed that something else must account for the perturbations in Mercury’s orbit (true). They conveyed much that was true but unspoken. These truths conveyed can contribute to the intuitive sense that the negative existential sentence says something true (though, if I’m right, it does not).

Another thing that can give one the sense that the negative existential is true is that (as David Braun has reminded me in correspondence) the sentence is sitting right there in the belief box (of most people). Normally, when a sentence enters one’s belief box one has the feeling that it expresses a truth. If one lacked that feeling, one would not allow the sentence entry into one’s set of “believed sentences”. So it surely seems to express a truth. Perhaps nearly everyone (but a few philosophers of the semantics of fiction) would profess belief in this negative existential. I confess that even I find the “seeming to be true” urge within myself when I utter the sentence or think it to myself. But my theory tells me to resist the urge because the sentence does not express a truth (or falsehood). There are other such cases. Euclid’s parallel postulate seems true to me (of space). Sometimes sentences about space and time (interpreted as Newton would, not Einstein) seem true to me—though I know they are not. Seemings can be wrong.

Sentences employing empty or fictional names can (do) produce the same seeming’s. It seems true that Sherlock Holmes does not exist (the flesh and blood man). As is clear by now, the account I am telling has this expressing only the incomplete negative existential logical form. No true or false proposition is expressed. Why does it seem true? Well the sentence can be sitting in one’s belief box where the cognitive system’s presumption is that only true things get to sit. And, there are true pragmatic things conveyed. There is no super sleuth by the name of “Sherlock Holmes” (true enough). There is not and never was a detective by that name living at 221B Baker Street (true enough). There were no two people (named “Holmes” and “Moriarty”) locked in a constant battle of wits

of good versus evil. There are many such truths pragmatically imparted but not uttered in the sentence “Holmes does not exist”.

## 4 The Ontology of Fiction

Why accept the view above? It will be said that a view that posits theoretical entities or fictional characters accords better with intuition or common usage.<sup>3</sup> And it does. That I do not dispute. But the view I am offering is one based upon a theory of names that I’ve developed over many years (Adams et al., 1992–2007) of thinking about how names work in language and in thought, and in the explanation of intentional action. It is one arrived at in considering not only issues in the philosophy of language, but in philosophy of mind in the dispute of broad versus narrow content. It is developed against a background of a theory of the explanation of purposive behavior and intentional action involving names. Although it has some unintuitive features, it is offered as part of a unified account of how names work in thought, talk, and deed. As we know, sometimes our true theories can conflict with common usage or intuition, as is the case in many areas of science and logic or mathematics, for instance. So while unintuitive features there may be, in the end the benefits of the unified theory of names (that they have the same semantics on every occasion of use) outweigh the conflict with common usage.

Thomasson (1999, 2003a, 2003b) has developed a view that countenances the existence of fictional characters (as ontological entities). On her view fictional characters are “made up” by authors. Fictional characters are “like artefacts in being created, contingent members of the actual world” (Thomasson, 2003a, p. 138). Though fictional characters are not concrete objects, they are said to “appear” in literary works of various kinds. They are able to “survive” the death of their creator, to be compared to other characters (real or fictional), and “enter into” other contents of thoughts or utterances, on this view.

---

<sup>3</sup> Thomasson and Martinich & Stroll say this in their works referenced below. I discuss their views here. But many others working in this area also want negative existentials to be true and want there to be ways to explain reference to fictional characters. I’m limiting my target here only to the ideas of Thomasson, and Martinich & Stroll. Consideration of other views will have to wait for another occasion.

On Thomasson's view, fictional characters are not Meinongian (non-existents that exist in a special way). They are not possible objects in possible but non-actual worlds. They are not mere abstracta or person-types (or kinds).<sup>4</sup> That is what they are not. To repeat, what they are is "an abstract cultural artefact created at a certain time by the act of an author writing a work of fiction" (2003a, p. 139). A character is a contingent member of the actual world. So, Doyle created Holmes and Holmes (Doyle's fictional creation) exists.

What does such a view have going for it? And does it have more going for it than the theory I've articulated above? One thing is that it agrees with common usage. As I've admitted already, it does. But what does this common usage amount to? To her credit, Thomasson does much to attempt to take the sting out of countenancing the existence of fictional characters. Indeed, so much so that she believes that "once we see what it takes [...] for there to be fictional characters, it becomes evident that it makes little sense to deny them" (2003a, p. 143).

For Thomasson, it is "sufficient" for a fictional character to be created that an author "write a work of fiction involving names *not* referring back to extant people or characters of other stories, and apparently describing the exploits of individuals named (or, if you like, pretending to refer to and assert things about a person, as part of an understood tradition of story-telling pretence)" (Thomasson, 2003a, p. 148).

Thomasson thinks anyone denying that these are sufficient conditions for the creation of fictional characters is "taking talk about fiction too seriously" (Thomasson, 2003a, p. 149). She thinks the person is mistakenly thinking that fictional characters are more than the product of literary practices (maybe thinking characters have to be real persons matching their descriptions, to be things into which we can enter causal relations, or other mistaken ideas about the ontology of fictional characters). Thomasson insists that there are no other necessary conditions for the creation of fictional characters than what authors do, "since the literary practices that set up the reference of terms like "fictional character" are *definitive* of the existence conditions for members of the kind" (Thomasson, 2003a, p. 149). She adds: "[...] *nothing additional is needed* to enable us to refer to a fictional character than for it to be true that, for example, a novel in

---

<sup>4</sup> So her view contrasts with views like those of Zalta (1983) or Parsons (1980).



a certain tradition makes certain claims—indeed these rules of use ensure that for most sentences of the form “*Pa*” in a work of literature (where “*a*” is a name and is not being used to refer to an extant person) we are entitled to infer that there is some fictional character, *a*, such that, according to the story, *Pa*” (Thomasson, 2003a, p. 150).

Thomasson concludes the minimalist construal of her view this way: “denying fictional characters while accepting the existence of the relevant fictionalizing practices only involves twisting the ordinary use of terms such as “fictional character”, severing the ordinarily permitted inferential connections between talk about fictional stories and what they say, and talk about the fictional characters about which things are said in the relevant stories” (Thomasson, 2003, p. 150). Thomasson reiterates that once one understands how extremely minimal are these conditions for the existence of fictional characters, it becomes “difficult” and “unnecessary” to deny that there are fictional characters, “so understood” (Thomasson, 2003a, p. 151).

Perhaps the first things to be said against Thomasson’s account, is that it has become *so minimalist* that one begins to wonder what the difference is between her account (which says that there are fictional characters) and my account (which says that there are not). On my view, there are fictional assertions in the contents of the literary works of fiction. Thomasson agrees. On my view, there are no physical entities in the physical world, in the head, or in possible worlds that constitute fictional characters. Thomasson agrees. On my view, there really *is nothing* else than the physical fictional assertions that one makes in storytelling or in literary practices of which fictional characters could be comprised. Thomasson agrees. Then what, exactly, is it that makes her (but not I) a *realist* about fictional characters? As far as I can tell, not much—except that she says *they exist* and I say *they do not*.

One begins to wonder if this is merely a terminological difference, but I think it is not and will try to convince you otherwise. So the second thing to be said about the difference in views is that on her view “Holmes exists” expresses a truth. On my view, it does not. What truth does it express, on her view? It expresses the truth that the Holmes character exists. Now what does this amount to? If it amounts only to the fact that there are fictional assertions using the name “Holmes” in the Doyle

literary works, then I agree about that, but don't see that this justifies the assertion "Holmes exists" is true.

Next, consider the very first sentence that Doyle ever pens using the name "Holmes". Suppose it is the sentence of the logical form "*Pa*" (where "*a*" = "Holmes"), "Holmes lives at 221B Baker Street London". Does that instance of the use of "Holmes" refer to the Holmes character? Hardly. How could it? The Holmes character does not yet exist. Even by Thomasson's own standards, it is the literary practices that determine what is true here. Until Doyle fictionally asserts enough about the Holmes character to put some "meat on his bones" so to say, there is no fully articulated character there to be referred to. Maybe there is one (though I deny it) after a suitably long set of such fictional assertions about the fictional person to flesh out a significant set of properties about him (he is *F*, *G*, *H*, *I*, *J*, *K*, etc.). But surely this is not so on the first occasion of use. And "Holmes" cannot just refer to what is in the head (mind) of Doyle, for then "Holmes" would not mean the fictional character but actual objects in the physical world (events in the head of Doyle). And these could not "appear" in other works or be discussed in works about the Doyle fictional works and so on. They have the wrong identity conditions to be the referent of "Holmes" on Thomasson's theory. So, even on Thomasson's own terms, the content of the first occasion of use of "Holmes" by Doyle cannot be the same as the content of the use by the end of the first novel (or of the last).

None of this is lost on Thomasson. She develops (Thomasson, 2003b, p. 212ff.) an elaborate account of the difference between *de re* and *de dicto* uses of empty names and how the former refer back to previous uses of the names in developing the character in the work, while in the early uses of the empty name in the fictional work, the fictional names indeed do not yet refer to fictional characters. Only later in the works or in subsequent works can reference to characters take place. (Although she does not clearly endorse it, she acknowledges (Thomasson, 2003b, p. 214) that one could say fictional names do not refer to fictional characters in fictionalizing discourse, although their use in fiction enables these names to later refer to characters in sentences about fiction.) If the early uses of empty names do not refer to fictional characters, then what is *their* semantic content and can they be used to express truths or falsehoods in positive or negative existential sentences? Thomasson does not exactly

say, but it is fairly clear that they cannot. In my view, this is a weakness of her account.

Next, consider that there surely seems to be a meta-level semantic claim being made by Thomasson in her account. This too does not escape her notice. If “Holmes” refers to the Holmes character, not an actual person or entity in the physical world, then the name can only refer, if it does refer, via the literary use of the name by Doyle. The works have to exist for the name “Holmes” to refer to them. The use of the name in the creation of the character cannot refer to its own creation. No other names depend for their semantics on the literary use of authors or literary works for their semantics or their reference. So, at the very least, to adopt Thomasson’s view of the meaning of fictional names requires that fictional names get their meanings differently (and mean different things) than the way non-fictional names get their meaning (or than what they can mean). Surely, an account that says names have the same meaning on every occasion of use (i.e., their bearer, if they have one) is a more unified and economical theory of names overall, than a theory like Thomasson’s. To my mind, this is a significant check in the plus column of my theory (if one is keeping score).

Lastly, what do we lose if, rather than being realists about fictional characters, we take the view that they do not exist? All there is are the literary works and the fictional assertions and the uses of the empty names—all connected causally to the author of the work. Thomasson thinks the fictional characters exist and are “made” are “artefacts” of the literary practices of authors. What do we lose if we say that everything Thomasson puts in the supervenience base is indeed there, but there just are no ontological entities called “fictional characters” that supervene on the literary bases? I think we lose nothing ontologically. We may even maintain our current ways of speaking—maintain our “usage”. But we need not let the tail wag the dog—that is, we need not let our ways of speaking determine ontology. Instead, when it comes to names, we can let our unified semantics be our guide to what there is (inside of fiction and outside).

In a recent book, Martinich & Stroll (2007) (hereafter M&S) also claim that there are fictional characters and that there are “fictional facts” that are a type of “institutional fact” that make sentences about fictional characters true or false (p. 15). Their claims come in the context of a book that argues for the complete overthrow of the sort of referential and causal theory of

meaning and reference upon which the theory I am advocating here is based. Limited space will not allow me to reveal their full attack here. Nor will I be able to respond completely to their full argument, as I would with more time and space. Instead, I will limit my remarks to their claims about institutional facts and that these are created by fiction and support the view that there are fictional characters.

M&S spend the early portions of their book denying what they call the “axiom of existence”, viz. that one cannot refer to what does not exist. Instead they offer an account based upon “linguistic practice” (p. 25) that shares much with the account given by Thomasson. They all maintain that linguistic practices give names of fiction their sense by the practices of telling tales and myths, writing stories and novels and the production of other types of discourse. For M&S, as for Thomasson, this involves concatenating names and descriptions, as in the Doyle novels, saying Sherlock Holmes (“was a detective”, “had Dr. Watson as a friend”, “lived at 221B Baker Street”). M&S suggest that by the linguistics practices of Doyle, for instance, institutional facts come to be and with them fictional truths—truths that can be supported with evidence by “quoting appropriate parts of Conan Doyle’s short stories”.

Now if we set aside the full attack on causal and referential theories of meaning, the basis for their acceptance that there are fictional characters (and truths about them) is the very same basis as offered by Thomasson—the concatenated sentences of fiction. And, like Thomasson, these linguistic practices (that yield M&S’s “institutional facts”) are the means by which it is possible to refer to fictional characters. M&S hold these to be the basis for truths in fiction and truths about fiction (for example, concerning Sherlock Holmes). They add that “the reason that one can refer to Holmes is that statements about him are accepted by those who participate in the practice of fiction” (p. 35).

All of my replies to Thomasson above apply equally to the views of M&S. So I won’t repeat them. I do think those replies to Thomasson present significant objections to the view of M&S. However, there are some new replies due to the uniqueness of their appeal to institutional facts as providing the basis for their claims that fictional names refer to fictional characters. And their non-referential view presents them with a puzzle Thomasson does not have—viz. distinguishing fiction from non-fiction.

First, let's consider their appeal to institutional facts. They say things like the rules of the game of baseball create institutional facts. There are things like "outs" that only exist because of the institution. Fair enough. But for something to be an "out" it must *instantiate* the rules. Someone has to tag the runner or the baseman must touch the bag or plate (ball in hand) prior to the runner reaching the bag or plate.

No one instantiates the "Holmes role". I can refer to actual events (the last out of the 2009 World Series game) that constitute outs in part because of the rules of baseball. I cannot refer to an actual event of Holmes entering 221B Baker Street.

Second, for the so-called institutional or fictional facts to be established, the elements in the ascriptions have to have meaning. On my view, the names and predicates get their meanings from their causal relations to the things named (if names) or the properties designated (if predicates). Since M&S reject such an account, they are advocating a "meaning as use" analysis of meaning. Still the predicates must have a meaning. If Doyle's pen slipped and into the story went "Holmes took the framus into his hand", even though there was a concatenation of name and predicates, it would have no meaning. "Framus" has no content. I just made it up. To have a content, it must have an established use. And it must be used to say true things to have an established use. One cannot maintain that first comes the use and then comes the ability to say true things. There have to be things to which the term "framus" correctly applies for the term to acquire meaning and use.

On the view of M&S, "Holmes did so-and-so" is true because of established use, because of established linguistic practice. But that would be like saying that framuses exist if I write a nonsense poem (like *Jabberwocky*) about them. Then we could say true things about them and refer to them. It gets the cart before the horse. It says that truth and meaning of terms can come prior to there being uses for these terms to say true things about the world. To me this seems to be special pleading to make the semantics fit the way people talk (remember that I do admit people talk as though there are characters and as though negative existential sentences employing empty names express truths or falsehoods). M&S (like Thomasson) use the way people talk to support their views (M&S, 2007, p. 8).

Lastly, on the account I have given above, part (a large part) of what makes something a fiction is that the terms of the work do not refer (and therefore do not express complete propositions capable of being true or false). On the view of M&S this *cannot be* what distinguishes fiction from non-fiction. Indeed, M&S (p. 16) take great care to say (over and over) that “there is no sharp line between fiction and nonfiction”. What makes something a fiction or an institutional fact, on their view, is purely subjective. M&S (p. 32) say: “These kinds of facts exist only because people agree to accept them as facts”. Of course, this view has untoward consequences. For political reasons, Sherlock Holmes could come into or go out of existence. For psychological reasons the same could be true. Imagine some mass hypnosis. Or more imaginatively possible, imagine a horrific act that wipes out most of civilization. Only a few countries of people survive. One day the survivors happen upon the Doyle novels and think that these are stories about an actual flesh and blood detective. That is, they are treated as nonfiction. Then what? Does “Holmes lives at 221B Baker Street” express the same content as it does for us now? M&S would have to say it is true for us now. Would it be the same truth for the survivors who take it as nonfiction? How could it be? Fiction is not nonfiction. Or suppose things went differently after the catastrophic event. Suppose the Doyle novels were discovered, but people who read them now rejected them as either fiction or non-fiction. That is, suppose they were not accepted as institutional facts. Did Holmes then go out of existence? Did once true sentences about him come to lack truth value? If so, the resultant view that M&S would be forced to accept (the Doyle sentences now lack truth values because no one accepts them as expressing institutional facts) is the view that I am proposing. They would just have gotten to that view via an admittedly weird, imagined circumstance that their view seems to have to permit. Their view is unstable. It yields my view in some imagined circumstances, but not others. My view is stable. It remains constant across both the actual circumstances and the imagined ones.

## 5 Conclusion

Here I have briefly reviewed the semantics, pragmatics, and ontology of fiction. The view defended grows out of the direct reference theory of names and the causal and referential accounts of meaning. We have only scratched the surface of what could be said. Much more can be found in the references, but I hope to have given enough here to allow one to understand what is at stake. The account I defend allows for names to mean the same thing everywhere they are used—within fiction, about fiction, comparing fiction, and outside of fiction. I would say the same thing about predicates, but the story gets longer to tell it completely. I do think it is the major strength of my account that it offers a unified account of the meaning of names (fictional or nonfictional). The other accounts here, for reasons I discuss, have to allow variation and instability in their accounts.

I have also discussed two of the recent accounts of the ontology of fiction that countenance the existence of fictional characters. I've presented the ways in which this type of view differs from the one defended here. I've also given reasons why I think my account is right and the others are wrong. Again, there is much more to be said than space permits,<sup>5</sup> but I hope the reader has here enough to see the conceptual layout of the views, how they differ, and the types of support each side musters in defense.

As is usually the case, there are many other worthy views of the semantics and ontology of fiction that have not even come up for discussion in this short paper. I have not attempted to deal with them because their approaches are so far from the one I am offering. The two alternatives I discuss here by Thomasson and Martinich & Stroll are at least in the same ballpark, even though we disagree.<sup>6</sup>

---

<sup>5</sup> In Adams et al. (1997) there is discussion of Currie, van Inwagen, Walton, and others. In Adams & Dietrich (2004) there is discussion of Taylor and Schiffer, among others.

<sup>6</sup> I want to thank the editor of this volume, Gary Fuller, and Amy Thomasson for useful correspondence and conversation about the topics discussed here.





# Fiction and Acceptance-Relative Truth, Belief and Assertion

R. M. SAINSBURY

A speaker may presuppose what is untrue to facilitate communication, as when an anthropologist adopts the presuppositions of his informants in questioning him. Most innocent of all are cases of fiction and pretending: speaker and audience may conspire together in presupposing things untrue. (Stalnaker, 1970, pp. 39–40)

Stalnaker mostly used the notion of presupposition in explaining conversational dynamics. But, as this quotation illustrates, he envisaged applying the notion to fiction. The present paper is largely an elaboration of his idea.

The quotation from Stalnaker shows that he is using “presupposition” to denote something speakers do (as opposed to a relation between propositions). In later work, he elaborates a propositional attitude, *acceptance*, which he contrasts with belief. It’s an attitude one should take to anything one presupposes, and is independent of the dynamics of conversation. Using the later terminology, we could rephrase what, in the quotation, Stalnaker says anthropologists do: they accept, though they do not believe, many of the things their informants believe. It’s this attitude of acceptance that is the target of the present paper. My aim is to illustrate

the way in which acceptance contrasts with belief, and to use the contrast to illuminate some intuitions about fiction.

Many people take sentences like the following to be in some sense true:

- (1) Holmes lived in Baker Street.
- (2) Pegasus is a horse.
- (3) Anna Karenina is more intelligent than Emma Bovary.
- (4) Even Dr Watson is cleverer than George Bush.
- (5) The Greeks worshipped Zeus.
- (6) John thought about Pegasus.

These sentences appear to be simple, formed just from a predicative expression plus the appropriate number of noun phrases. For such sentences, it is tempting to suppose that their truth requires the referring expressions to have referents, and the temptation seems irresistible for the first four sentences. Yet everyone should be reluctant to see what has been said so far as a “proof” that there really are fictional characters—that these characters belong to our reality, as opposed to the unreal world of fiction. (Even those who like the conclusion shouldn’t think it can be reached quite so easily.) The plan for this paper is to show how an irrealist—one who denies that our reality contains any fictional entities—can accommodate what has been said so far by drawing on a distinction between the propositional attitudes of acceptance and belief.

## **1 A Problematic Argument**

- I. “Holmes lived in Baker Street” is true.
- II. It is a simple sentence.
- III. “Holmes” is a referring expression.
- IV. If a true simple sentence contains a referring expression, there exists something the expression refers to.

V. There exists such a thing as Holmes.

Here are some options for those who regard the argument as unsatisfactory (I hope that includes everyone!):

- (A) Reject (I): the sentence isn't really true, but is only true in the fiction, or is merely faithful to the stories.
- (B) Reject (II): the sentence is not really simple, but is implicitly complex, dominated by an operator that can make a truth from a non-truth (e.g. "According to the *Holmes* stories").
- (C) Reject (III): "Holmes" is not really a referring expression. E.g. it's meaningless, or we only pretend that it has a bearer. This option requires also rejecting (1).
- (D) Reject (IV): Sentences built from "intensional transitive verbs" are counterexamples, e.g. the truth of "Ponce de León looked for the fountain of youth" does not entail that there exists such a thing as the fountain of youth.

The response I propose is closest to (A): rejecting (I). However, I don't think one can merely assert that "Holmes lived in Baker Street" is not true. One has to explain the perfectly genuine intuitions that make us wish to say that it is true. The essence of the present idea is that we properly take one of two propositional attitudes to the sentence: we accept it; but we do not believe it. Acceptance brings with it a full cargo of acceptance relative notions: truth, assertion and so on. Our judgment that the sentence is true is implicitly relative to accepting the stories. That explains why it is sincerely assertible, even by someone under no delusions about the fictionality of Holmes. Yet acceptance can be withdrawn or bracketed, and when we do that we have to regard the sentence as not true; we do not believe it.

The first step is to spell out the distinction between acceptance and belief, showing that it is available independently of providing some understanding of fiction. Many authors have engaged in this project, but I mostly won't stop to highlight how my version of this distinction differs from others.

## 2 Exemplifying the Distinction in Non-Fictional Cases

Sales assistants need to accept the following, even if they do not believe it:

- (7) The customer is always right.

There's no chance of someone with ordinary common sense believing this. We all know there are customers who lie, who try to defraud, who suffer from buyer's remorse, and so on. But to do their job properly, in certain circumstances sales assistants have to act as if they believed it. That means: listening carefully to what the customer says, not challenging it, making remedial proposals that take for granted the customer's story; and so on. Although acceptance does not involve belief, it does not involve disbelief either. Sometimes an accepted account is believed, sometimes not. The acceptance is required only in delimited contexts, in this case, when interacting with a customer (or a representative of one). After hours, it's quite all right for sales assistants to tell one another about how wrong many of their customers were.

The example gives an initial fix on the distinction. In the remainder of this section, I'll give a range of further examples. In §3 I'll apply the distinction to fiction. Finally, in §4, I'll offer an abstract presentation of the essential marks of the distinction between acceptance and belief.

The distinction has been said to be needed in order to enable us to give adequate descriptions of such forward looking activities as planning. Here's an example from Bratman (1992):

You are planning to build a house. You need to see if you can afford it before you begin. You sum the highest estimates from the various trades, reaching a total of \$200k. You don't *believe* your house will cost this much, but you *accept* that it will: this is your working assumption.

Although intended to be on the side of the distinction between acceptance and belief, the example seems to me hostile to it. There is an obvious way in which, instead of using the distinction between propositional attitudes, we can describe the case in terms of a single attitude, but to different

contents, both of which you believe: that the cost will not exceed \$200k, and that the cost will be less than \$200k. The contrasting descriptions can be displayed thus:

<b>two attitudes, one content</b>	<b>one attitude, two contents</b>
1. You accept that the cost will be \$200k.	1. You believe that the cost will not exceed \$200k.
2. You do not believe that the cost will be \$200k.	2. You believe that the cost will not amount to \$200k.

In my opinion, the one attitude plus two content view gives a more plausible description of the case. The beliefs are perfectly consistent, could readily be acknowledged by the subject, and will be fully adequate to explain the behavior.

This tells us one thing to look out for in examining cases designed to illustrate the distinction between acceptance and belief: check whether the case could as well or better be described in terms simply of the attitude of belief, though to different contents. We'll raise that question for the examples that follow.

*Robbery* (adapted from Lewis, 1979): We are planning a bank robbery, though not in order to commit a robbery but to enable the bank to improve its security. The plan currently under discussion is one we both agree is likely to fail, and we are rehearsing it to see exactly where it goes wrong.

*I say:* "We'll get an accomplice to do a classic stick-up. While everyone's busy with that ..."

*You interrupt:* "We'll go in through the underground ductwork."

*I agree:* "Exactly."

You're sincere when you say that we'll go in through the ductwork. I have a positive attitude to what you say ("Exactly"). But you don't really believe it. Nor do I. At that moment, we *accept* that we'll go in through the ductwork. It's the right thing to say and think. The natural description of the case is in terms of two attitudes to a single content: it is *accepted* but not *believed*.

Could we describe the case using a single attitude but two contents? There are two questions. One is: what semantic content should we ascribe to the utterances? Another is: how should we best describe the underlying states of mind of the participants?<sup>1</sup> There's only one answer to the first question: when you utter the words "We'll go in through the underground ductwork" you say that we'll go in through the underground ductwork; and you say nothing else.

That is strictly consistent with not including this content in an explanatorily optimal description of your mental state.<sup>2</sup> However, given that you've uttered these words in a sincere fashion, one would expect the content to show up in a full description of your mental state. You don't *believe* we'll go in through the underground ductwork, or ever embark on a robbery. So if the content is to feature in your mental states, it needs to be related to an attitude other than belief. Acceptance is the proposed alternative.

On a two content view, your relevant states are to be described only in terms of your beliefs, so that *we'll go in through the ductwork* will not be included. What should we put in its place? It would have to be some kind of conditional. It can't be: *if we rob the bank, we'll go in through the ductwork*. We mutually know that this is the worst plan, so if we were to rob the bank, we wouldn't do it this way. Maybe it's something like: *if we were to put this plan into effect, we would go in through the ductwork*. In the example, the idea of going in through the ductwork was meant to have struck us with something like the force of an insight. It wasn't an element in an already determined plan. Maybe all that defined "this plan" was that we'd use a conventional stick-up as a distracter, and our prior low ranking of any plan of this kind was based on the thought that the resulting police presence would thwart anything else we might try to do. In the reported conversation, we are searching for possible alternatives. This suggests that perhaps some other conditional would match our mental state better, e.g. one in which "would" is replaced by "could" or "might": *if we were to put this plan into effect, we could/might go in through the ductwork*. It would

---

<sup>1</sup> Thanks to Stacie Friend for clarifying this distinction.

<sup>2</sup> An example of this consistency: In *Problems of Philosophy*, Russell suggested that proper names have only their bearer as public content. But to give an illuminating account of the mind of a user of the name, you need to mention not the public content, but the description the user associated with the name on that occasion.

be hard to choose among similar candidates to be believed contents. The effort seems to get us no closer to a good account of the minds of the participants, and to involve speculation that goes beyond what's given in the story. By contrast, the account in terms of acceptance seems accurate, precise and illuminating.

*Lawyer:* A lawyer and his assistant are poring over a draft contract. Frowning, the lawyer says:

If interest rates rise more than 1.7%, we'll owe the other side a million dollars a month!

The lawyer is fully sincere: he *accepts* what he says. The upshot might be that the contract needs to be revised before he could recommend his client to sign. The lawyer does not *believe* what he says: he knows it's not true now, and he will act so as to ensure that it remains false.

As before, there's no room for doubt about the semantic content of the lawyer's utterance, so the only question is whether this content should feature in an optimal description of his mental states. The fact that he has said this gives a *prima facie* reason for answering that it should. But then it can't feature as the content of a belief.

A suitably related content the lawyer might believe may well, as in the case of *Robbery*, be conditional, perhaps: *if this contract were in force, then, if interest rates rise more than 1.7%, we'll owe the other side a million dollars a month!* But can that be quite right? The lawyer might know that if the contract were in force, his client would have fired him by the time the debt to the other side arose, so that the "we" in the consequent would be inappropriate. And can we count on everyone being able to operate mental contents with double conditional embeddings? At the very least, there are niggly doubts about whether we can extract an appropriate content for belief. By contrast, the appropriate content for acceptance stares us straight in the face.

*Violinist:* A disturbed patient is recounting his (entirely fictitious) early history to his therapist:

When I was young, I played the violin. I performed Beethoven's sonata in E flat at the Wigmore Hall.

The therapist knows this is false, but decides it's best to roll with her patient's delusions and says:

Did you play an encore?

The question presupposes that the violinist performed at the Wigmore Hall, though the therapist does not believe this presupposition. Even so, the content *my patient played at the Wigmore Hall* can quite properly be among the contents of the therapist's mental states, accepted but not believed. In this case, it's hard to see how one could explain the therapist's mental states without introducing this content. That's because it's presupposed by the question, and the therapist must accordingly be at least implicitly aware that this is so.

*Anthropologists:* Here are two sentences one might find in a description of Latin American culture a few hundred years ago:

- (8) When he was not among the clouds, Chac could be found near falling waters.
- (9) The moon goddess was the wife of the sun, but her true name has not passed down to us.

The anthropologists who affirm these sentences may combine full sincerity with firm and clear-sighted atheism; yet the truth of the sentences presupposes the existence of Chac and the moon goddess. The proposed explanation is that the anthropologists accept but do not believe what they are saying, along with the attendant presuppositions.

As in every case, an account of what the anthropologists *say* cannot be other than homophonic, and, as always, it would be surprising if the content of what a person says does not figure as the object of a propositional attitude that person holds. But if, for some reason, we are determined to admit only believed contents, we would have to consider possibilities like:

- (10) According to the Maya, when he was not among the clouds, Chac could be found near falling waters.
- (11) According to the Maya, the moon goddess was the wife of the sun, but her true name has not passed down to us.



One problem with the first suggestion is that the anthropologists in the conversation, while accepting the utterance about Chac, may disagree about what the appropriate operator prefix is. One of them might indeed think that a Mayan belief is being reported, but the other might have some skeptical doubts about the cohesiveness and integrity of the so-called Mayan empire, doubts that make the very use of the term “Maya” suspect in his eyes. Intuitively, this disagreement does not prevent them agreeing about Chac; but the operator-prefix suggestion does not guarantee a common content for them to agree on.

For the second utterance, the prefixing suggestion faces a distinct problem: the Maya never thought about us; in particular they did not think that their name for the moon goddess had not come down to us. The scope of the prefix must therefore end after “the sun”. It is then hard to explain how the anaphoric pronoun “her” functions. By contrast, there’s no problem at all, if the content of the original utterance is accepted but not believed.

The anthropological example introduces myth, arguably a species of fiction. But the contrast between acceptance and belief is plainly the same contrast as that involved in the other cases, going right back to “The customer is always right”. So we have a contrast that’s robust outside fiction and, as Stalnaker said, is all ready for application to fiction. How, exactly, should that application go?

### **3 Applying the Distinction to Fiction**

People often have rather vacillating intuitions about whether sentences like “Holmes lived on Baker Street” are true. Those who wish to say that it is will cultivate a particular kind of context. For example, they invite you to imagine yourself in a Literature 101 multiple choice exam, where you have to decide whether to check “Holmes lived on Baker Street” or “Holmes lived on Dover Street”. We all know the right answer. But what makes it *right*? It’s natural to say that it’s because the sentence “Holmes lived on Baker Street” is *true*.

Theorists who wish to persuade us that the sentence is not true will try to make other contexts salient. They will ask us about the absence of a suitable Holmes from any of the potentially relevant censuses of England (1871, 1881, 1891), about his absence from any local memoirs of the pe-

riod, and about the apparent nonexistence of any such address as 221B. Imbued with the high seriousness of this context, we are likely to record the sentence as not true. After all, it was just a story.

A proper account needs to explain both of these seemingly conflicting intuitions. The contrast between acceptance and belief does the job to perfection. The first intuitions arise from our accepting the *Holmes* stories, and so accepting that Holmes lived in Baker Street: we accept this when reading, thinking about, or rehearsing the stories. The second intuitions arise from our not believing that Holmes lived in Baker Street: taking the most austere position, there is no such person, and so no question of his living anywhere. It is perfectly consistent to both accept that  $p$  and not believe that  $p$ . Both intuitions are vindicated as entirely reasonable, and this is a desirable result. I'm not aware of any other account which gives such a simple and fitting account of these facts.

The account can also explain what, from many other viewpoints, is a surprising asymmetry. Let's suppose that the *Holmes* stories recount an incident in which Holmes met Gladstone. Most people think that "Holmes met Gladstone" is a lot better, in some truth-like way, than "Gladstone met Holmes". Yet the sentences are necessarily equivalent. Here's the explanation that a belief-acceptance theorist can offer: it's often the case that the first word in a sentence sets up a framework that needs to be accepted in the course of interpretation. When the first word is "Holmes", the framework is typically set to the *Holmes* stories: we are primed to accept what they say. If they say that Holmes met Gladstone, we accept this. When the first word is "Gladstone", the framework is typically set to the real world. We know that Gladstone can't really have met Holmes. So we have no inclination to accept the sentence. This account is supported by the fact that replacing "Gladstone" by a little-known name, or by a name that is both a name for a real person and for a fictional character, undermines the contrast.

As Stalnaker (2002) stresses, participants in a conversation can get along fine even if they have different *beliefs* about matters their conversation presupposes, provided they coincide in their *acceptance* of the relevant presuppositions. An example outside fiction: atheists can debate with theists about the best version of the doctrine of the Trinity. The conversation presupposes that there is a God; one party believes this, the other does not. However, the conversation could not take place at all (in the con-

structive spirit I envisage) unless both parties accepted that there is a God. Acceptance is local: as with “The customer is always right”, acceptance in a context carries no commitment to acceptance in any other context. Belief, by contrast, like truth, is seamless.

Can an irrealist about fictional characters count on the acceptance-belief contrast to rid her of all ontological anxieties? Certainly not. Contrast the following three sentences:

(12) Holmes lived on Baker Street.

(13) Ponce de León looked for the fountain of youth.

(14) The Greeks worshipped Zeus.

We have seen that the contrast does a good job of explaining our attitudes to (12). It has nothing to offer with respect to (13): this is a sentence we ought not merely to accept, but to believe absolutely and from the most austere perspective. Unlike (12)’s presupposition of Holmes, the truth of (13) does not presuppose that there exists a fountain of youth. We should reject, or at least restrict, the semantic principle previously labeled IV:

If a true simple sentence contains a referring expression, there exists something the expression refers to.

The restriction should exclude sentences built from intensional verbs like “looks for”. Formulating it properly is no doubt a difficult matter. For present purposes, it’s enough that “lived” is definitely not an intensional transitive, and “looked for” definitely is. There is a straightforward way to show that the contrast between acceptance and belief will not help us understand (13). Contrast:

(15) There’s no Sherlock Holmes, but he lived in Baker Street (all the same).

(16) There’s no fountain of youth, but Ponce de León looked for it (all the same).

(15) is not acceptable. The first phrase sets the context to reality, and so to a situation in which we do not accept the stories, and so do not accept the

second phrase.<sup>3</sup> By contrast, (16) is both acceptable and believable: it's true from the most austere perspective.

With this test to hand, we can now consider (14). To some, this has seemed true, to others false; that already suggests the acceptance-belief contrast may do some useful work. Let's see whether (14) patterns with (15) or with (16):

(17) There's no such god as Zeus, but the Greeks worshipped him (all the same).

This strikes me as, at a minimum, infelicitous. This suggests that even though we can accept that the Greeks worshipped Zeus, since we can accept that there is such a god, for example in the service of giving simple accounts of the Greeks' behavior, we should not believe that this is so. Once the acceptance of the god is explicitly precluded, by our acceptance of there being no such god, we cannot happily go on to believe that the Greeks worshipped him. This also accounts, in the now familiar way, for the two kinds of intuitions about (14).

Let's return to our original slate of sentences that might pose problems for irrealists, repeated here:

- (1) Holmes lived in Baker Street.
- (2) Pegasus is a horse.
- (3) Anna Karenina is more intelligent than Emma Bovary.
- (4) Even Dr Watson is cleverer than George Bush.
- (5) The Greeks worshipped Zeus.
- (6) John thought about Pegasus.

---

<sup>3</sup> What must be accepted can naturally shift even midway through a sentence (though it does not do so in (15)): "Sherlock Holmes lived on Baker Street, or so the story goes". The sentence as a whole does not require us to accept anything we don't believe, but the first clause does. A problematic question is whether we have to accept the existence of something in believing it does not exist. If not, the following would also be an example of mid-sentence shift of what must be accepted: "Holmes was a detective—but of course he doesn't really exist".

We've seen how the present suggestion deals with (1) and (5). (2) and (3) will be treated just like (1). The interest of (3) is that cross-fictional comparisons are sometimes supposed to raise difficulties for irrealists who place their trust in operator approaches, since Tolstoy's novel did not speak about Emma, and Flaubert's novel did not speak about Anna.

(6) is normally taken to pattern with (13) ("Ponce de León looked for the fountain of youth"): built from an intensional transitive, it can be true (absolutely) even if there is no Pegasus. If we apply the earlier test, there's some room for doubt:

(18) There's no such thing as Pegasus, but John thought about him (all the same).

Those who think this is unacceptable may prefer to classify it with (5) rather than, as is more customary, with (13).

(4) asks us to accept that Dr Watson (of the *Holmes* stories) and Bush can be compared, which in turn requires accepting that there are such people as Dr Watson and George Bush. Is that the same as accepting that there are such *real* people as Dr Watson and George Bush? For this example, it may not matter how one answers. But consider the following familiar sentence:

(19) Holmes is more famous than any real detective.

If accepting this involves accepting that Holmes is a real detective, we seem to be committed to accepting that Holmes is more famous than himself. The more cautious formulation of the previous paragraph seems more appropriate: in accepting (19) we accept that there is such a detective as Holmes, and we are committed to accepting that he is not real.

The upshot is that an irrealist should not in every case appeal to the acceptance-belief contrast to explain away apparently pro-realist intuitions. But they should do this in some cases. Panaceas should not be trusted.

## **4 Nine Marks of Acceptance as Opposed to Belief**

This section aims to get a more theoretical fix on the contrast, which up to this point has mostly been explained by examples.

*One*: “To accept a proposition is to treat it as a true proposition in one way or another” (Stalnaker, 1984, p. 79). One may treat something as true without believing that it is true. Cohen also takes something like this idea as a starting point. To accept that  $p$  is:

To have or adopt a policy of deeming, positing, or postulating that  $p$ —i.e. of including that proposition or rule among one’s premisses for deciding what to do or think in a particular context, whether or not one feels it to be true that  $p$ . (Cohen 1992, p. 4)

What it is to *treat as true* will vary from case to case. To treat it as true that the customer is always right is to behave in a certain way towards customers; one does not have to persist in this behavior out of hours, or if one’s role is reversed and one is the customer oneself. Treating the Holmes sentence as true involves not challenging it, in the context of retelling or discussing the stories, and using it as a basis for inference and further imaginative exploration of the circumstances described in the story. It does not require one to take seriously a search for a Holmes in the residential records of the Borough of Marylebone.

*Two*: Acceptance does not entail belief. That’s the essential part of the story. I follow most accounts in using “acceptance” so that belief entails acceptance. One treats as true what one believes to be true, but one may treat as true something one does not believe to be true.

*Three*: Acceptance is voluntary. Many take this as a contrast with belief, which has been argued not to be voluntary. Given the ruling in *Two*, that belief is a species of acceptance, this mark is to be understood as applying only to acceptance that is not belief.

To illustrate the voluntary character: a sales assistant whose concern to keep his job is not overriding may decline to accept that the customer is always right, thereby putting his employment at risk. The therapist might have declined to accept that her patient played at the Wigmore Hall, thereby plunging into a different style of therapy.

The alleged involuntary character of belief does not sit easily with the assumption that belief is the key locus for evaluating an agent’s rationality, praising agents who believe rationally, blaming those who do not. Involuntary behavior is often said to be exempt from praise or blame. As Cohen

(1992) hints, it might be that acceptance is a better locus for such evaluations.

*Four:* Sometimes (though not always) what is accepted is what is *presupposed*, in the way that “The present King of France is bald” presupposes that there exists a present King of France. Here I’m thinking of presupposition as a semantic relation between propositions. One should accept all the presuppositions of any proposition one accepts. Speech acts that are not sayings, but, for example, questions, may have presuppositions (the question “Did you play an encore?” presupposes that you performed at the venue in question). That’s consistent with presupposition being a relation between propositions; questions that have presuppositions can be regarded as involving propositions (in mickey mouse form: Is the proposition that *you played an encore* true or false?).

*Five:* In conversation, participants may need to share contents as “common ground” (in the sense of Stalnaker, 2002). They do not need to believe these contents (as with atheist and believer discussing the Trinity): what one believes, another may accept without believing.

*Six:* Acceptance is not essentially conversational. It can play a similar role in silent soliloquy. An atheist philosopher of religion, preparing an academic paper alone in his study, may need to accept God’s existence in order to work out the best version of the doctrine of the Trinity.

*Seven:* Acceptance is local but belief is not: accepting something in one context involves no commitment to accepting it in another. A shift of context can properly involve rejecting something previously accepted. After hours, the experienced sales assistant can straighten out the inexperienced one: “Don’t believe what they say about the customer being always right”. This means that there’s nothing necessarily irrational about accepting that *p* and also (in another context) accepting that not-*p*. Belief is different in this respect, being what I earlier called seamless: setting aside change over time, if in some context you believe that *p*, there’s no context in which it’s correct to believe that not-*p*. Belief “aims at truth”, and truth meets the analogous condition: if *p* is true, then not-*p* is not.

*Eight:* Acceptance carries commitments (and in this respect resembles belief): normally, to accept something (within a context) involves a commitment to accept obvious relevant consequences of it (within that context). Once you’ve accepted that Holmes lived in Baker Street you are committed to accepting that he lived in London.

*Nine:* There are acceptance-relativized notions of *sincerity*, *correctness*, *truth* etc. Given appropriate acceptances, it is true, “really” true if you like, that Holmes lived in Baker Street. Sincerity demands that this is what you assert, if the situation is fitting.

Consider the view that knowledge is the norm of assertion: it’s right to assert that  $p$  only if you know that  $p$ . It’s not obvious that this view has to convict someone of wrongdoing if she asserts that  $p$  in a situation in which she accepts that  $p$ , though doesn’t (globally) believe that  $p$ , and so doesn’t (globally) know that  $p$ . The examiner in Literature 101 asks: Where did Holmes live? Your answer, “Baker Street”, sounds like an assertion that he lived in Baker Street. Although you have not mistaken the fiction for factual narrative, you are fully sincere. Relative to what is being accepted in this context, you know Holmes lived in Baker Street. So it’s far from obvious that the full legitimacy of such fictional assertions conflicts with knowledge as a norm of assertion. Once acceptance-relative truth is admitted, acceptance-relative knowledge follows, as does the propriety of evaluating assertions relative to what is accepted in the context.<sup>4</sup>

---

<sup>4</sup> A version of this paper was presented at BW6: the *Sixth Barcelona Workshop on Issues in the Theory of Reference*, organized by the LOGOS group. The conference theme was “Reference and Non-existence”. My thanks to the organizers for inviting me to participate, and to other participants for their helpful comments.



# Fictional Realism and Its Discontents

ROBERT HOWELL

The term “fictional object” will be used here to label any purely imaginary entity (or putative entity) that occurs as a character in fiction, for example Anna Karenina. (Many real things also occur in fiction, for example Moscow in *War and Peace*; but they are not our present concern.) Two strands of thought have, in recent years, dominated the discussion of such entities: the realist views about fictional objects that have been defended by Peter van Inwagen, Nathan Salmon, Amie Thomasson, myself, and others; and the pretense or make-believe accounts that derive principally from the work of Kendall Walton and that have been developed by Gregory Currie, Frederick Kroon, Stephen Yablo, and others.<sup>1</sup> I have thought for many years that some form of realism stands the best chance of giving an adequate account of fictional objects. But recent criticisms point out new and fundamental problems. In this essay, I will explore the prospects

---

<sup>1</sup> For realism, see van Inwagen (1977, 1983, 2003); Salmon (1998); Thomasson (1999); Howell (1979, 1983, 1996, 1998, 2005); Scott Soames (2002); Stephen Schiffer (2003); Alberto Voltolini (2006). Saul Kripke appears to have developed a realist theory in the early 1970s, but it has never been published. See Salmon (1998). For make-believe views, see Walton (1990, 1993, 1997, 2000); Currie (1990); Kroon (2000, 2005); Yablo (2001, 2005); Crimmins (1998). I lack space to discuss other important work on fiction, including Nicholas Wolterstorff (1980) and the gappy-proposition views of David Braun (1993, 2005) and of Frederick Adams, Gary Fuller, and Robert Stecker (1997). See also the bibliography. For surveys of recent work, see Howell (1979, 1998, 2005), Thomasson (1999), and Anthony Everett (2005). A companion piece to the present essay, Howell (2010), will appear in a volume edited by John Woods. Anyone interested in the discussion below may want to consult that sequel.

for a realist answer to these problems or at least an answer that preserves something of the realist spirit.

To this end, and after a brief review of fictional realism in §1 and of make-believe theories in §2, in §3 I describe the problems. §4 notes some inadequate solutions to them. In §5 and §6, I then indicate an avenue for future investigation. This avenue retains aspects of realism while also abandoning some key realist claims. I will not develop these new views in full detail, and they raise questions of their own. I hope, however, that they may provide a promising direction for future work on these issues.

## **1 Realism**

The details of realist theories differ, but realists unite in arguing that many statements involving the term “Anna Karenina” are, when properly understood, true. And they argue that the best way to account for the truth of these statements (and for the truth also of general statements about the characters of fiction) is to suppose that there are indeed genuine entities such as Anna Karenina. The statements in question include apparent references to and descriptions of Anna Karenina herself (“Anna Karenina has a son”), comparisons of her with real objects and with other fictional objects (“Anna Karenina and Tolstoy were both Russian”, “both Anna Karenina and Emma Bovary die unhappily”), and statements expressing the fact that a character can be the common object of different psychological acts and can occur in different stories (“Jane and Samuel both are thinking about Anna Karenina”; “Huck Finn occurs in three novels by Mark Twain”). In the case of each of these sentences, realists argue that we should take the term “Anna Karenina” to function as a genuine singular term that designates an object, the character Anna Karenina. And they take that object in a certain way to satisfy the predicates in the statements and so to have the properties that those predicates express.

In arguing for such results, realists also especially stress the ontological commitment to characters of fiction that is carried by our acceptance of the truth of general statements such as “There are more characters in Tolstoy’s major novels than in Flaubert’s”, “Some characters occur in several different novels”, and “some characters in nineteenth-century novels are described in more detail than any characters in eighteenth-century

novels”. And they note that this commitment is reinforced by the most straightforward way of understanding the inference relations that appear to hold among such truths.<sup>2</sup>

Different realists spell out these ontological commitment in different ways. Ignoring Meinongian proposals and some other suggestions, I focus here on the form of realism that is, to my mind, especially plausible.<sup>3</sup> That form of realism embraces actualism and so takes the character Anna Karenina, like every object to whose being we are committed, to exist in the actual world. By “realism” I will hereafter mean actualist realism.

Although realists commit themselves to the actual-world existence of the character Anna Karenina, such realists agree that in an important sense it is true that Anna Karenina does not really exist—that there exists no woman, identical to that character, who has the properties that the novel ascribes to Anna. Assuming, as they do, that the term “Anna Karenina”, in our relevant actual-world use, designates the character, they then need to reconcile the nonexistence of the person, Anna Karenina, with the existence of that character and with the view that, as we have seen, sentences such as “Anna Karenina is a woman” are, when properly understood, true. They do so by distinguishing the actual-world existence of the *character* from the actual-world existence of the particular, concrete *woman* who bears the name “Anna Karenina” in the novel.

They take the character to exist, but they deny that there exists any entity that is both that character and that has all the properties that Tolstoy’s novel ascribes to Anna Karenina. Rather, they take the character, as it exists in the actual world, to be an abstract entity that lacks any specific space-time location or other marks of the concrete. Thus the character has, in the actual world, such properties as *being created by Tolstoy* and *being the main character of Anna Karenina*, as well as such logico-metaphysical properties as self-identity. But, in the actual world, the character does not have such properties as *being a woman* and *having two children*. Instead, it has only such properties as *being an entity such that it is true, according to the novel Anna Karenina, that that entity is a woman (and has two children, and so on)*.<sup>4</sup> Only “in the world of the story”, as we say, does that

<sup>2</sup> Van Inwagen (1977, 1983, 2003) emphasizes these last points about characters.

<sup>3</sup> For Meinongian views, see Terence Parsons (1980).

<sup>4</sup> Realists elaborate the present distinction in different ways. I think that the way suggested below is the most satisfactory. (It also brings out the parallel between claims about

entity actually have the properties of *being a woman, having two children, and so on*.<sup>5</sup>

We can understand this distinction between the kinds of properties by introducing the operator “it is *Anna Karenina*-fictional that” (or “according to *Anna Karenina*, it is true that”). Using this operator, we can represent *de dicto* claims made by the novel: “it is *Anna Karenina*-fictional that there are people walking in the street”. We can also represent the novel’s *de re* claims about actual-world objects—for example, “Moscow is an object such that it is *Anna Karenina*-fictional, of that object, that that object is a large city”. And we can represent the *de re* claims that, as realists suppose, the novel makes in introducing specific fictional characters and then ascribing properties to those characters—for instance, the claim “There is an object  $x$  such that it is *Anna Karenina*-fictional that ( $x$  is named “Anna Karenina” and  $x$  is a woman)”. The object whose actual-world existence is responsible for the truth of this last claim is, in the actual world, the object that satisfies the open sentence after the quantifier. That is, it is the object that has the property of *being an entity  $x$  such that it is Anna Karenina-fictional that ( $x$  is named “Anna Karenina” and  $x$  is a woman)*; and that object will be the abstract entity that is the realist’s character.<sup>6</sup>

As many realist discussions have shown, by appealing to formulations along the present lines and using standard logical techniques, we can represent the sorts of sentences that we noted above. By looking at the truth conditions of these representations, we can then see how and why the truth of such sentences commits us to the existence of characters of fiction considered as actual-world, abstract entities of the sort noted above.<sup>7</sup> We also can accommodate such familiar phenomena as the indeterminacies

---

fiction and those about psychological attitudes such as belief.) Everett (2005, p. 644) objects to such distinctions as being obscure. But I do not think that there is any obscurity in the distinction between *being  $F$*  and *being said by the story to be  $F$* .

<sup>5</sup> The notion of the world of a story for much of this essay can be taken as primitive, for it is common ground between the present sort of realists and those critics, such as Everett, who are considered below. It is natural to represent such a world as the set of propositions that are, according to the story, true (that is, are fictionally true given the story). As noted in §6, my talk there of assuming that there is a unique world of the story can be understood in this way; but it can also be understood in other ways. See also fn. 65 below.

<sup>6</sup> In the relevant cases (as for example here) the existential quantifier can be replaced with a uniqueness quantifier. But for the most part I will not make that point explicit below.

<sup>7</sup> Thus e.g. we have “ $(\exists x)[[it\ is\ Anna\ Karenina\text{-fictional\ that}\ (x = Anna\ Karenina \wedge x\ is\ named\ \text{“Anna\ Karenina”} \wedge x\ is\ a\ woman)] \wedge Jane\ thinks\ about\ x\ and\ Samuel\ thinks\ about\ x]$ ” as a representation of the Jane-Samuel claim above. And we have, say, “ $(\exists x)(\exists f)[x\ is\ a$

and inconsistencies that often belong to fictional characters.<sup>8</sup> In addition, we have a framework in terms of which we can consider various specific accounts of how the ostensible proper name “Anna Karenina” is to be understood semantically.<sup>9</sup> Moreover, realism allows us to do these things without forcing us into a Meinongian commitment to purely nonexistent fictional objects.

## 2 Make-Believe Theories

The principal competitor with such realism is the kind of pretense or make-believe theory that Kendall Walton and authors influenced by him have developed. Walton’s account of fiction offers a forceful explanation of why we make claims such as “Anna Karenina is a woman” even while we take the term “Anna Karenina” not to designate any actual woman who is as the novel describes.<sup>10</sup> On Walton’s account, that term is empty. Characters of fiction do not actually exist, and the ostensible names of purely

---

character  $\wedge$  it is  $f$ -fictional that ( $x$  is a woman)]” as a representation of the claim that there is a character in some fiction who is a woman.

<sup>8</sup> See e.g. Howell (1979, pp. 134–5), and Howell (2005), as well as many other authors. The basic idea is that, for any actual object  $o$  and any property  $P$ , we have excluded middle, so that it is true at the actual world that either  $o$  has  $P$  or  $o$  does not have  $P$ . However, a story  $s$  can leave it open whether a character has  $P$  or does not have  $P$ . Or a story can take it (either by accident or quite deliberately) that  $o$  has and does not have  $P$ . These facts can be formulated in a way that leaves the actual-world—and the presence of  $s$  in the actual world—perfectly determinate and consistent. Thus, assuming that  $s$  is a usual sort of realistic novel, it will be true, in the actual world, that it is  $s$ -fictional that (either  $a$  has  $P$  or  $a$  does not have  $P$ ). But it may not be true either that it is  $s$ -fictional that ( $a$  has  $P$ ) or that it is  $s$ -fictional that ( $a$  does not have  $P$ ). If  $s$  tells an inconsistent story about  $a$ , then it will be true, in the actual world, that, say, it is  $s$ -fictional that ( $a$  has  $P \wedge a$  does not have  $P$ ). The general strategy here can of course also be adopted by pretense and related theorists who reject the actual existence of characters of fiction.

<sup>9</sup> For example, realists can take the actual-world term “Anna Karenina” to be a genuine proper name of (and, on the rigid-designator view of proper names that I accept here, a rigid designator of) that abstract entity. In the world of the novel, “Anna Karenina” certainly functions also as a genuine proper name (and rigid designator) of the woman Anna Karenina. But, in the actual world, that term functions only as a fictional name of such a woman—it is only an *Anna Karenina*-fiction that that term rigidly designates that woman.

<sup>10</sup> I don’t distinguish here between make-believe, pretense, and “simulation” versions of Walton’s views. The present comments give only a cursory sketch of a rich, deep theory. For further comments, see Howell (1996) and Howell (2005).

fictional entities that occur in stories are only fictional names, not genuine proper names of existent objects.<sup>11</sup>

Walton arrives at these and related results by focusing on games of make-believe. He supposes that works of fiction function as props in such games, games whose rules mandate us to imagine roughly that the sentences of the fiction are true. In doing so, we imagine that the term “Anna Karenina” in the text is a genuine proper name of a woman whom Tolstoy’s sentences characterize. Yet, Walton holds, there actually are no such objects; it is all a pretense.

If I say, for example, simply that “Anna Karenina is a woman”, then this sentence may be uttered simply as *part of* the pretense that I carry out when I read the novel and pretend that I am reading a true description of a genuine world. And I then use this sentence (and not, say, the explicit, non-pretending claim that “according to *Anna Karenina*, Anna is a woman”) because I am copying the actual sort of behavior—of verbally describing a real person—that I am pretending to perform.<sup>12</sup> Because there is no Anna Karenina, there is no proposition here that is such that I pretend, *of* it, that *it* is true. Rather, and at most, within the game it is fictional *that there is* a true proposition that my sentence expresses. Suppose, however, that I use the sentence “Anna Karenina is a woman” *outside* that pretense in the real world in order actually to assert something about the story. In that case, I express a true proposition. But I do not express any proposition about Anna Karenina (there is no such entity). Rather, I express the proposition that to assert such a thing (namely, something exemplified by the claim “Anna Karenina is a woman”) *in* the appropriate game of make-believe is, in that game, to speak truly.<sup>13</sup>

<sup>11</sup> Walton does not, as far as I know, explicitly discuss actualism. But his rejection of Meinongianism and his treatment of terms such as “Anna Karenina” commit him to actualism, as far as I can see.

<sup>12</sup> Walton (1990, pp. 219, 221, 223, 391–2). In correspondence, Walton notes that, in some cases, I may not have in mind any actual such behavior that I am trying to mimic or copy. (See also, for example, Walton, 1993, §§II and III.) But, according to his view, I am still proceeding in some pretense- or simulation-involving way, a way that (in examples like the Karenina one) is not a case of the actual-world describing of a person.

<sup>13</sup> For more details, see Walton (1990, p. 400 and, generally, pp. 396–415). Strictly speaking, according to Walton I express the proposition that the novel *Anna Karenina* is such that anyone who engages in the “Anna Karenina is a woman” kind of pretense in a make-believe game authorized for that novel (a game, roughly, of making believe that its sentences are true) makes it fictional of himself in that game that he speaks truly. This formulation brings

Imagine, now, that I compare Anna Karenina to Theodor Fontane's character Effi Briest by claiming that "Anna Karenina has more experience of the world than does Effi Briest". Then, says Walton, I simply play a further, unofficial game that combines the games individually authorized for the novels *Anna Karenina* and *Effi Briest*. It is then fictional *in* that game that there is a true proposition expressed by my claim. My claim may, however, be uttered outside that unofficial game as a real-world assertion about the worldly backgrounds of the two characters. Then what it expresses is the true proposition that to assert such a thing (namely, something exemplified by the Anna-Effi claim) *in* that game is to speak truly. The idea of a single unofficial game evidently also suggests a way of treating the case in which two people talk about the character Anna Karenina (or the case in which I form a belief about that particular character). In addition, that treatment can be applied to the case in which Anna Karenina appears in two different stories.<sup>14</sup>

Walton uses this apparatus to deal with many other statements ostensibly about fictional objects, including negative existential claims such as "Anna Karenina does not exist". He takes such claims, when they are actual assertions (and not simply parts of a make-believe), to be disavowals of the pretenses to refer that are exhibited in our use, in them, of ostensible names of fictional objects.<sup>15</sup> He offers related analyses of claims

---

in a participatory game with respect to that novel. See fn. 14. These details are not important here, however.

<sup>14</sup> Walton offers a further explanation for our impression that, in reading the text, we are brought into *de re* contact with a particular entity, Anna Karenina. The explanation is that besides playing games in which we make-believe that the sentences of the fiction are true, we also play participatory games in which we make-believe that our own experiences of reading the text are experiences of reading about real objects named by terms such as "Anna Karenina". (See Walton, 1990, pp. 130–7, especially pp. 135–6; Howell, 1996, p. 416. Walton, chap. 9, notes complications about narration; they can be ignored here.) The effect is, for Walton, that I come to make-believe that there is an object *x* such that my experience of reading the sentence "Anna Karenina is a woman" is an experience of reading *about x*. So, inside this overall *de dicto* make-believe (that there is such an object *x*) is embedded a *de re* claim that I make-believe is true (namely, the claim that my experience is one of reading *about x*). I ignore this complication below.

<sup>15</sup> Roughly, in asserting the claim just mentioned, one pretends to use "Anna Karenina" to refer; and then one disavows the kind of attempted reference exemplified by that pretense. The claim "Anna Karenina does not exist" then asserts a truth—namely, that to attempt to refer in the "Anna Karenina"-way will not succeed. (I ignore various complications.) For criticisms, see e.g. Stanley (2001, pp. 60–4). Walton (2000) responds and extends his analysis further.

like “Anna Karenina is a fictional character”. In making that claim, he suggests, we may be engaged in a further, unofficial game in which there are things called “fictional characters”. By playing this further game and, within it, taking “Anna Karenina” to refer to a character, we betray our original pretense (in our authorized game with *Anna Karenina*) that “Anna Karenina” refers to a real person. Walton suggests that our talk of characters may itself be viewed as being part of a metaphysical make-believe to the effect that characters exist and appear in realms that are associated with the relevant fictional works.<sup>16</sup>

My goal in describing Walton’s views here is not to launch a full-scale evaluation of them, although in fact I do not accept Walton’s treatment of characters as creatures merely of make-believe. Rather, I have presented the make-believe view in order to provide a foil for the account of fiction that I sketch in §5 and §6. I should note immediately, however, that actualist realists of course vehemently reject the anti-character aspects of make-believe theories. They reply to Walton and his followers that our talk of characters is indeed literal (and true), not part of a pretense. They also fault such theories for being unable, as they see it, properly to represent the truth of claims that appear to commit us to the existence of characters. And they take such theories not properly to account for the truth of claims that seem to imply that there is a single fictional object that occurs in different stories or that is the common object of several different psychological acts. In addition, realists argue that nothing stops them from incorporating into their own views the insights of the make-believe theory. Thus realists can recognize the make-believe that a claim such as “Anna Karenina is a woman” may involve; and they can do so while also taking that make-believe to commit us to a genuine entity, the character Anna Karenina.<sup>17</sup> So it has seemed to realists that their position can stand as a defensible account of fiction and its objects, even given the value of make-believe treatments of these matters.

---

<sup>16</sup> Walton (1990, pp. 423–4); also Everett (2005, p. 640ff.) The realist criticisms mentioned in the next paragraph, and others, can be found in, for instance, van Inwagen (1983, 2003); Howell (1979, 1996, 1998, 2005); Thomasson (1999, 2003b). For replies, see, for example, Walton (2000, 2003), and Yablo (2001).

<sup>17</sup> See e.g. Howell (1996, 1998, 2005); van Inwagen (2003). The idea would be that the relevant make-believe introduces or is *of* the character Anna Karenina just as (on realism) it is fictional *of* that character, according to the story, that that character is the woman Anna Karenina.



### 3 Problems

However, there are problems. Those that I have in mind are not the questions that bedevil all accounts of fiction, non-realist as well as realist (for example, the issue of how to understand negative existentials such as “Anna Karenina does not exist”). Nor are they the sorts of difficulties for whose solution one can immediately see a plausible realist strategy.<sup>18</sup> Rather, they are fundamental problems that appear to show that the present form of realism is untenable. These problems were, as far as I know, first noted by Anthony Everett (2005). He divides them into two groups: those turning on the ontic indeterminacy of characters, and those turning on logical incoherence. My discussion of these problems appeals to the apparatus that I have developed above, rather than to Everett’s own presentation. But I do not think that the force of his points is thereby lost.<sup>19</sup>

#### Ontic Indeterminacy: (a) In Identity.

It is possible that a realistic novel  $f$  (and not just some sort of logical fantasy) should introduce fictional objects  $o$  and  $n$  but leave it open whether or not  $o$  is identical to  $n$ .<sup>20</sup> Thus imagine a highly realistic Dr. Jekyll-and-Mr.-Hyde detective story that clearly raises the question of whether Jekyll

---

<sup>18</sup> For example, make-believe theorists ask how a realist can explain why we say “Anna Karenina is a woman” instead of speaking the literal truth, namely, that there is an abstract object that has the property of being such that it is *Anna Karenina*-fictional that that object is a woman. Adopting the view noted in fn. 17, realists can answer, however, that we say simply that “Anna is a woman” because, as devotees of fiction, we are interested in the contents of our make-believe that the story is true. So we say something that we intend to be evaluated at the world of the story.

<sup>19</sup> Everett (2005). He argues by appeal to two principles: (P1) “If the world of a story concerns a creature  $a$ , and if  $a$  is not a real thing, then  $a$  is a fictional character”; and (P2) “If a story concerns  $a$  and  $b$ , and if  $a$  and  $b$  are not real things, then  $a$  and  $b$  are identical in the world of the story iff the fictional character of  $a$  is identical to the fictional character of  $b$ ” (p. 627). Using the problem cases that I note below, Everett offers still further criticisms of realism. But I ignore those criticisms here, for I believe that realists can manage them if they can deal with the cases below.

<sup>20</sup> By a “realistic novel” I mean a novel, such as *Washington Square* or *Anna Karenina*, that, roughly, describes objects and situations (and a whole course of history) that could exist in the actual world without contravening physical or psychological laws or the usual probabilities. I don’t, of course, mean a novel that espouses the theory of fictional realism. The Jekyll-Hyde case is mine, not Everett’s. (He tells a related story about the characters Frick and Frack.)

and Hyde are the same person but that (unlike Robert Louis Stevenson's novel) does not settle that matter. Such a novelistic effect could be created deliberately—it might be crucial to the story's impact on the reader—or simply through an authorial slip. Because such a novel presents a realistic fiction about the actual world, excluded middle holds in the world of this novel. So (using obvious abbreviations) it is true at the actual world that (i) it is *f*-fictional that (either Jekyll = Hyde or Jekyll ≠ Hyde). (By "Jekyll ≠ Hyde" I mean " $\neg(\text{Jekyll} = \text{Hyde})$ ".) But (ii) it is also true at the actual world that it is not *f*-fictional that (Jekyll = Hyde); and (iii) it also is true at the actual world that it is not *f*-fictional that (Jekyll ≠ Hyde).

However, the present form of realism further agrees that, because Jekyll and Hyde are particular fictional objects, (iv) Jekyll and Hyde are characters that have an existence in the actual world (as abstract entities; but that fact is not relevant here). Moreover, because Jekyll and Hyde are purely fictional objects, it seems clear that whether they are or are not the same character in the actual world depends entirely on whether it is fictional, according to the novel, that they are identical to each other or it is fictional, according to the novel, that they are not thus identical.<sup>21</sup> So it is true at the actual world that (v) Jekyll = Hyde iff it is *f*-fictional that (Jekyll = Hyde). And it is also true at the actual world that (vi) Jekyll ≠ Hyde iff it is *f*-fictional that (Jekyll ≠ Hyde).<sup>22</sup>

<sup>21</sup> Of course what counts as fictional according to the novel is a complex matter that depends not just on the actual words in the text but also on how a properly informed reader will understand the novel, and so on. See especially Walton (1990, chap. 4). But that fact is consistent with the points made here.

<sup>22</sup> From claims of the form (v) and (vi) one can derive results such as (using further abbreviations) " $\neg f(a = b) \rightarrow f(a \neq b)$ " and " $\neg f(a \neq b) \rightarrow f(a = b)$ " for fictions *f* that admit realist characters *a* and *b*. These results are incorrect, in themselves, in the case of stories that leave open the identity or not of *a* and *b*. However, it is realists who are committed to (v) and (vi) and hence to these results. They are so committed given that—as they should—realists agree that the identity (or distinctness) of *a* and *b* as actual-world entities depends entirely on the identity (or distinctness) that is ascribed to *a* and *b* by the story. Anyone who rejects realism is free to reject these results. (See also the final paragraph of §6 below.) Note also that, given (ii) and (iii), (vii) below can be derived from just the left-to-right implications in (v) and (vi). Benjamin Schnieder and Tatjana von Solodkoff (2009, p. 143) point out complications that arise in connection with principles about character identity when characters occur in stories from which they do not themselves originate. (For example, Rosenkrantz and Guildenstern occur not only in *Hamlet* but also in Tom Stoppard's play.) I ignore these complications in this essay. They do not affect the specific points that I make below using (v) and (vi), and they do not weaken the force of the Everett problems, as I am considering

Given claims (i) to (vi), it follows, however, that it is true at the actual world that (vii)  $Jekyll \neq Hyde$  and  $Jekyll = Hyde$ . But this result is catastrophic for any form of actualist realism that rejects the holding of genuine inconsistencies in the actual world. And all the forms of such realism that I know—certainly all the forms that I find plausible—will reject any such holding.<sup>23</sup> Realism thus appears inconsistent, given the possibility of indeterminate character identities in fiction (for purely fictional objects).

One might try to escape this contradiction by introducing a third truth value, “undetermined”, in such a way as to make the actual-world truth value of both “ $Jekyll = Hyde$ ” and “ $Jekyll \neq Hyde$ ” undetermined. (This assignment of truth value could be motivated by the fact that  $f$  does not settle the identity or nonidentity of Jekyll and Hyde.) If we introduce this third truth value, then the left-to-right implications in (v) and (vi) will not be true in the actual world. (A conditional with an undetermined antecedent and false consequent is, at best, undetermined.)<sup>24</sup> So the preceding actual-world contradiction cannot be derived. However, this point does not resolve our present difficulty. The disjunction “either  $Jekyll = Hyde$  or  $Jekyll \neq Hyde$ ” will now itself be undetermined in the actual world, given that each of its disjuncts is undetermined. So the identity or nonidentity of the genuine, existent objects, Jekyll and Hyde, will be indeterminate at that world. Yet, as Everett urges, it is surely not plausible to suppose that

---

those problems here. It also would not be difficult to restate my views (including the new position that I develop in §6) in a way that takes such complications into account explicitly.

<sup>23</sup> I assume that realists (at least of the sorts here under discussion) will not want to appeal, at this point, to some ontically applicable paraconsistent logic in order to allow for the actual-world truth of the inconsistent (vii) or of the other inconsistent claims noted below. Defenders of such logics are welcome to apply them here. But I take it that current actualist realists (at least the ones I know) will not want to adopt such logics. I sometimes wonder why those realists who respond to Everett’s problems by accepting ontic indeterminacies of identity or existence (see below) draw the line there. Why don’t they simply move on to some full-scale paraconsistent logic? That would certainly give them a clear, straightforward way of accepting all of Everett’s results. However, I won’t pursue this question here. I don’t see good independent reasons (beyond the need to escape the present sorts of contradictions) to suppose that the actual world exhibits ontic inconsistencies. And I suppose that current actualist realists concur.

<sup>24</sup> These are the truth values that, intuitively, we might want to assign to the various claims here, given the value “undetermined” and the fact that  $f$  does not settle the identity or nonidentity of Jekyll and Hyde. They correspond to the truth values assigned by Kleene’s strong three-valued logic.

actual reality is thus indeterminate.<sup>25</sup> And it seems that actualist realists should agree.

### **Ontic Indeterminacy: (b) In Existence.**

Everett notes that in Tatiana Tolstoya's novel *The Slynx* it is left open whether there really is a Slynx.<sup>26</sup> From the perspective of the world of that novel, the Slynx may just be a figment of various characters' false beliefs. Or perhaps there really is, in that world, such a creature. But then (and to put the matter in my present realist terminology) it will be indeterminate whether there exists, in our actual world, a character (an abstract object), the Slynx. Yet the actual world is not indeterminate in this way about whether an object exists in it or not.

Given the present form of realism, Everett's formulation of the Slynx problem may need revision. On the one hand, if the Slynx is no more than a character in the delusions of various inhabitants of the *Slynx* world, then, according to realism, this character exists as an abstract entity in that world. And then (on at least one way of developing a realist account of Tolstoya's novel) that abstract entity itself exists in our actual world as an abstract entity *a*. (We, in the actual world, recognize, as a character, the merely delusional Slynx character that is accepted by various inhabitants of the *Slynx* world.) On the other hand, the Slynx may actually exist, as a concrete creature, in the *Slynx* world. In that case (and given the present form of realism), the Slynx exists as an abstract object *b* in our actual world.

On this way of reading the situation, and contrary to Everett's own discussion, the problem is not that it is indeterminate in the actual world whether the Slynx exists in that world. The problem is that it is indeterminate which of the two evidently distinct abstract objects exists in the actual world, *a* or *b*. However, and whether we interpret the Slynx case in Everett's way or in this last way, that case still has a devastating bite. Surely (one imagines that present realists will say) the actual world is fully de-

---

<sup>25</sup> Everett (2005, pp. 628–30). Everett rejects various ways of escaping this indeterminacy, including attempts to discredit Gareth Evans' well-known argument that it cannot be indeterminate whether genuine entities are identical (Evans, 1978). Everett does not derive the above contradiction but restricts himself to arguing (using his (P1) and (P2)) for the actual-world indeterminacy of what is in effect Jekyll's identity with Hyde.

<sup>26</sup> Everett (2005, pp. 630–2).

terminate as regards which specific entities exist in it. Yet the Slynx case shows that, as Everett puts it, realism is committed to the presence, in the actual world, of a “pernicious case of ontic indeterminacy”. Realism is not committed merely to the presence of a benign case of indeterminacy that arises from imprecision in the concepts that we use to describe the world.<sup>27</sup>

### Logical Incoherence.

Here Everett develops several cases. The first makes the problem very clear, and I focus just on it.<sup>28</sup> In Everett’s story “Dialethialand”, Jules both is and isn’t identical to Jim. Realists should thus agree that in the actual world we have a character *ju* (an abstract object), of which it is a “Dialethialand”-fiction that that character is Jules. And we have a character *ji* (an abstract object), of which it is a “Dialethialand”-fiction that that character is Jim. But, according to the story, Jules = Jim and Jules  $\neq$  Jim. And from principles like (v) and (vi) in the Jekyll-Hyde case, it follows that it is true at the actual world that both  $ju = ji$  and  $ju \neq ji$ .

In this case, actualist realism appears to prove itself inconsistent. In the two previous cases, we have either ontic indeterminacy or outright ontic inconsistency. It seems that, on the face of it, these results should be unacceptable to the realists whom we are discussing; and many readers will find them unacceptable on any plausible view of the actual world. Despite the attractive features of realism noted in §1, it follows that we must abandon realism unless we can find a way around Everett’s problems. Everett himself holds that there is no such way. He rejects the actual existence of fictional characters in favor of a pretense view that allows for indeterminacies and inconsistencies in characters—but only within the relevant fictional worlds. (And certainly both realists and pretense theorists can agree that, unlike our actual world, fictional worlds may exhibit such phenomena.)

---

<sup>27</sup> See Everett (2005, p. 628), and (2005, pp. 628–32) generally. Everett argues against various other responses to this case. Thomasson (2010) disagrees with my treatment of the *a, b* situation. For a bit more on this point, see Howell (2010, §IV).

<sup>28</sup> Everett (2005, pp. 633–8). The second case concerns a story in which identity is asymmetric: Cicero is Tully, but Tully is not Cicero.

## 4 Can Realism Escape?

One might try to respond to the preceding problem cases by arguing that we cannot really suppose there to be any genuine characters in the above stories, for we cannot imagine indeterminate or logically inconsistent states of affairs in ways that let us take those states of affairs to involve genuine characters. Or one might argue that while there are genuine characters in the stories, they are not inconsistent or indeterminate. Rather, they belong only to substrands of the stories that are themselves consistent and determinate (say, a consistent Jules substrand of the dialethic story and a consistent Jim substrand). However, Everett argues convincingly that these and related avenues of escape do not succeed. As he stresses, we can indeed imagine the relevant states of affairs. (We can at least imagine them adequately enough to arrive at the fictional worlds whose characters generate the above problems.) And it is the whole point of stories such as the Jules-and-Jim one to present an inconsistent world. The strategy of factoring such stories into consistent substrands will not work.<sup>29</sup>

Moreover, and as Everett stresses, we can follow the doings of Jules and Jim or of Jekyll and Hyde as we read the stories. We can engage imaginatively with these characters, and we can enter sufficiently into the worlds of the stories to speculate on the fates of the characters as the stories develop. In this respect, reading about such characters is no different from reading about determinate, consistent characters such as Hamlet, Anna Karenina, and Loki and Baldur. Why should realists accept the existence of characters of this latter sort and yet reject the existence of charac-

---

<sup>29</sup> For the present points, see Everett (2005, pp. 634–8).

ters of the sorts that occur in Everett's problem cases?<sup>30</sup> No satisfactory, non-question-begging grounds have been given for doing so.

Nor do various other realist attempts to defuse the above problems work any better. As thought shows, the Jekyll-Hyde problem will disappear if we suppose that the story  $f$  really concerns two pairs of object between which the story itself does not distinguish: one a pair of in-fact-identical objects, and the other a pair of in-fact-distinct objects. (Or, again, the problem will vanish if we hold that there really are two distinct stories here,  $f^*$  and  $f^{**}$ . These stories will coincide in the actual sentences that they contain, but while  $f^*$  concerns a pair of identical objects,  $f^{**}$  will concern a pair of distinct objects.) The difficulty with such ideas is not, I think, that there could not be such *recherché* fictions. (Perhaps, with ingenuity, we could describe, in a detailed, coherent way, how such stories might run.) The difficulty is that the sort of realist novel that we are here imagining is surely a single story about a single pair of objects about whose identity or nonidentity the story simply does not inform us. So the

---

<sup>30</sup> For the point about imaginative engagement, see Everett (2005, pp. 634–5). I think that there are clearly characters in the Jules-Jim case (assuming some coherent theory of characters) as much as in the case of *Hamlet* and *Anna Karenina*. (I also think that it is clearly possible to have a fiction, such as the Jules-Jim one, that is best interpreted as taking an explicit inconsistency to hold true, within the world of the story, in a central fashion that cannot be interpreted away as merely apparent or accidental.) But the implications of Jekyll-Hyde cases are especially compelling. Imagine that an author begins a story with characters  $a$  and  $b$  without yet deciding whether or not  $a$  is identical to  $b$ . In chap. 30, the author eventually indicates that they are distinct. Then no one will deny that there really exist two characters here. But now imagine that in chap. 30 the author announces, instead, that although  $a$  and  $b$  are clearly *either* identical *or* distinct, the story will not reveal which. Do we then really want to conclude that there are no characters  $a$  and  $b$  here at all?

present way out of the Jekyll-Hyde case will not work.<sup>31</sup> And attempts to apply analogues of it to the other problem cases above also will not work.

In addition, it is implausible to suppose that there is indeed a fact of the matter about whether or not Jekyll is identical to Hyde. Were that supposition correct, then exactly one of the disjuncts in the claim “either Jekyll = Hyde or Jekyll  $\neq$  Hyde” would be true at the actual world. For example, and even though the content of *f* does not tell us the following, there might somehow be, in the world of the novel, a truthmaker that makes Jekyll’s identity with Hyde the case. And hence, although we readers could never know this fact, it might be true in the actual world that Jekyll is identical to Hyde. However, and again, the sort of realist detective story that we are here imagining is not like this. It is not that there is a fact of the matter about Jekyll’s identity to Hyde, a fact that already exists (where? in some never-never land?) but about which the story does not tell us. Our Jekyll-Hyde story is not like a history of France that breaks off before we learn whether Charlemagne was or was not the same person as Charles Martel. Readers of our story are not epistemically deprived in the

---

<sup>31</sup> I have a similar reaction to an ingenious way out of this case that correspondence with van Inwagen has suggested. (The present formulation of this way out is my own, and van Inwagen is of course not responsible for the conclusions that I’ve drawn from it.) Why not suppose that there is a fictional set of persons—itsself a single, determinate, realist character of the novel—that is such that the novel ascribes to that set the property of *having exactly one or two members* but does not ascribe to that set either the property of *having exactly one member* or the property of *having exactly two members*? I think that there could certainly be such a novel (presumably, a logically or philosophically fairly sophisticated one). But (a) the present Jekyll-Hyde novel is not a story about such a single set—or pair—of persons. (It is not a story whose main character simply is such a set.) It is a story with the definite main characters Jekyll and Hyde, a story that simply does not tell us whether those characters are identical or not. (In this respect, it is no different from the Charlemagne/Charles-Martel history of France that is mentioned below.) To reinforce this observation, ask yourself the following: would you want to say that, when you start reading the story imagined in fn. 30, you suppose that there are indeed the characters *a* and *b*; but when you are told in chap. 30 that the identity or nonidentity of *a* and *b* will not be revealed, you then suddenly decide that there is just one character in this story after all, the set (or pair) of *a* and *b*? (b) In any case (and as this last question helps to bring out), the above description of the set that is supposed to be the main character of the Jekyll-Hyde fiction is incomplete. If that description is to come even close to capturing the content of that fiction, it must include the fact that, according to the story, Jekyll and Hyde—each of whom the story may describe, individually, in great detail—are both members of that set. But now the problem for realism returns. Jekyll and Hyde appear to be characters of this fiction in their own right, as well as characters in virtue of their membership in that set. So what do we now say about their identity or distinctness as abstract entities in the actual world?



way that readers of that history text would be. Nor, as she writes, does a novelist spend her time trying to discern the murky details of some far-off, independently given fictional world, sometimes getting those details right, sometimes getting them wrong.

Again, realists cannot plausibly deal with the Jules-Jim incoherence by taking there to be, in the actual world, just one consistent character. That character would be, in the story world, identical to Jim and also identical to Jules, Jules and Jim then being both identical and distinct in that world. This suggestion distorts the point of this deliberately dialethic fantasy, which is that we have characters who are both identical and distinct. Similarly, the point of the Jules-Jim story is distorted if we take there to be two characters in the actual world, each of which is identical to Jules and to Jim, who then both are and aren't identical to each other in the story world.

We also distort our understanding of fictions and their characters if we seek to avoid the Jekyll-Hyde case and Everett's other examples by modifying claims (v) and (vi) of §3 above (and Everett's own principles (P1) and (P2)—see fn. 19). (a) Thus (and using obvious abbreviations), a realist might suppose that instead of accepting (v) and (vi), we should adopt the principle (P\*): Jekyll = Hyde (in the actual world) if it is *f*-fictional that (Jekyll = Hyde); Jekyll ≠ Hyde if it is *f*-fictional that (Jekyll ≠ Hyde); and otherwise it is indeterminate (in the actual world) whether Jekyll = Hyde or Jekyll ≠ Hyde. When it is applied to the Jekyll-Hyde case, (P\*) will certainly let us avoid the §3 contradiction. But appealing to (P\*) also plunges us into an actual-world indeterminate identity and so opens up the further problems that I consider below. So I do not think that adopting (P\*) provides a way out of the difficulties.

(b) Again, a realist might hold that (P\*\*) Jekyll = Hyde if it is *f*-fictional that (Jekyll = Hyde); Jekyll ≠ Hyde if it is *f*-fictional that (Jekyll ≠ Hyde); and otherwise Jekyll ≠ Hyde (in the actual world). (P\*\*) thus alters the straightforward connections in (v) and (vi) between the actual-world character identities and the fictional-world character identities.<sup>32</sup>

(c) Or, indeed, a realist could introduce an idea of character that severs entirely the connection between actual-world character identities (and their determinacy, or not) and fictional-world character identities (and their de-

<sup>32</sup> Compare Schnieder & von Solodkoff (2009, p. 143). Thomasson has also noted principles like (P\*) and (P\*\*) in correspondence and in Thomasson (2010).

terminacy, or not). (A realist might hold, for instance, that to each occurrence of a fictional name or referring term, within the fiction, there corresponds a new, distinct, actual-world character—so that there would be at least as many actual-world Anna Karenina characters as there are occurrences of the term “Anna Karenina” in the text.)

Proposals such as (b) and (c) will indeed avoid the Everett problems. To my mind, however, these proposals make radical modifications in our usual concept of character, a concept whose content any plausible form of realism ought to respect. In our usual dealings with characters, both as readers and as critics, we treat characters as, fundamentally, things that are *identified with* the persons (salient objects, machines, animals, etc.) that figure *in the stories*. We talk, after all (as critics of fictional realism often note) about the character who *is* the woman, Anna Karenina (and not about the abstract object that has the property of being said by the story to be a woman). So our concept of a character, even on the realist view, ought to maintain, for characters as abstract objects existing in the actual world, the identity or distinctness relations that hold or don't hold among the characters as persons in the fictional worlds. We wouldn't (in our ordinary talk of characters) say that while the identity of Jekyll and Hyde is left open in the story, nevertheless the actual-world identity of the *characters* Jekyll and Hyde is settled, so that, for example, in the actual world Jekyll  $\neq$  Hyde (as (P\*\*) has it) or in the actual world Jekyll = Hyde (as a further (P\*\*\*) might have it). What we say about characters such as Jekyll and Hyde tracks what we say about the persons in the fictional worlds because the characters, understood in the way characters are described in everyday life and in literary criticism, *are* those persons.<sup>33</sup>

Probably the best strategy for the realist, at least in regard to the Jekyll-Hyde case, is to argue that the actual world does indeed admit of some cases of indeterminacy. In particular, the actual world allows that it is neither determinately the case that Jekyll is identical to Hyde nor determinately the case that Jekyll is distinct from Hyde, even though Jekyll and Hyde are actually existing entities.<sup>34</sup>

<sup>33</sup> On characters, see also fn. 68 and 79 below. I thank Thomasson (who disagrees with this view of characters) for comments that stimulated some of these present remarks.

<sup>34</sup> One might try analogous moves for the Slynx case, and my comments below will apply to that case, also. Everett argues at pp. 628–33 against various responses to the Slynx and the Jekyll-Hyde cases that rely on indeterminacy. My present observations are based on a dis-

In support of this idea, we might appeal to the existence of actual-world cases of vagueness. Thus most people will agree that at some point in removing grains of sand from a heap, we no longer have the original heap (or, indeed, any heap at all). But it is not clear at what point; in the actual world, the heap/non-heap boundary is vague. Or, again, it may be clear that we have a range of mountains, with several clearly distinct, highest points. But where exactly Mount Alpha gives out and Mount Beta begins may nevertheless not be settled, it seems. And, depending on what the borders of Alpha and Beta are taken to be, it might even not be settled whether Alpha (which we've identified by looking at the mountain range from vantage-point *A*) is identical to Beta (which we've identified by looking at the mountain range from a different vantage-point *B*). Or, again, it might not be determined whether Alpha is identical to Gamma, where Gamma amounts to the physical extent included by Alpha, taken together with one additional stone.<sup>35</sup>

If we agree that such examples of vagueness already exist in the actual world, then a realist might argue that the existence of Jekyll-Hyde (or Slynx) cases poses no irresolvable objection to fictional realism.<sup>36</sup> Yes,

---

cussion with Amie Thomasson and comments by John Woods. I am not trying to report their own views in any detail. Thomasson, who defends fictional realism, discusses vagueness in Thomasson (2007), chap. 5. She concludes that there is no serious ontic vagueness in the world (although she also suggests a way around Evans' argument against indeterminacy of identity). She doesn't there consider the sorts of specific problems raised above about realist treatments of fictional objects. But I believe that she is sympathetic to something like the strategy suggested here. (Thomasson, 2010, which I saw after completing the main body of this essay, also seems to support such a strategy.) I believe that realists such as van Inwagen also will find some such strategy correct. (See fn. 42.) However, it presumably will not appeal to Salmon, who rejects indeterminate identity. (See, e.g., Salmon, 1981, 243–6, and Salmon, 2002.) I don't know how Salmon will want to handle the present problems.

<sup>35</sup> Compare Timothy Williamson (2003, pp. 705–12). I don't commit realists or myself to his sort of epistemicism about vagueness, however. But I do accept his conclusion that supervaluational treatments of vagueness give no grounds for concluding that objects are themselves, ontically, vague (Williamson, 2003, p. 706). So I won't consider such treatments in my discussion, below, of the strategy for defending realism that is indicated in the next paragraph of the main text and in fn. 36.

<sup>36</sup> As various of my points below imply, I don't myself suppose that fictional indeterminacy (in which a story *s* leaves it open whether or not object *a* has property *P* or is identical to object *b* or even exists) is a case of vagueness of the sorts noted here and below—of the heap or of the Mounts Alpha, Beta, Gamma, Tripod, and Quadruped sorts. (Thus, for example, and using the non-vague language of Euclidean geometry, I might write a story about geometric figures in which it is left open whether or not object *o* is a triangle. Here there is fictional indeterminacy but no vagueness. Similarly, were it somehow not to be determined,

reality is such that although the abstract Jekyll and Hyde entities exist, it is indeterminate whether or not they are identical. But, similarly, reality is such that although mountain ranges clearly exist, it is sometimes indeterminate, in the vagueness sense, whether one mountain is or is not identical to another. If we accept the actual-world existence of the latter sort of cases, then we should also accept the actual-world existence of Jekyll-Hyde or Slynx cases.

There is a huge literature on vagueness and indeterminacy, and there are different views of the phenomena, not all of which will support the above realist escape from the Jekyll-Hyde problem.<sup>37</sup> It is not clear that we should agree that there really are any genuine ontic indeterminacies of the vagueness sort in reality itself, as against indeterminacies that arise because of vagueness in the concepts that we apply to reality. I do not want to commit realists (or myself) to the existence of such ontic indeterminacies. However, in order to cast doubt on the present way out of the Jekyll-Hyde problems, it is enough to note the following. Even if ontic indeterminacy of the vagueness sort turns out to infect reality in some cases, the heap and mountain examples do not support the actual-world existence of the specific, Jekyll-Hyde kind of ontic indeterminacy. The story *f* that we have imagined is not a logical or philosophical fantasy in which people (or the particular people Jekyll and Hyde) have vague boundaries either physically or psychologically. That story is a realistic fiction, not a story according to which it is not clear where Jekyll (perhaps described from one vantage point) ceases and Hyde (perhaps described

---

in the actual world, whether object *o* is an exact Euclidean triangle or has exactly 100,000-foot-long hairs on its head, then—assuming for argument's sake that “foot-long hair on its head” is not vague—there might be no vagueness involved.) But I am not trying to settle the relations between indeterminacies of various sorts and cases of vagueness. Instead, I simply note a possible realist strategy for answering Everett: accept the actual-world existence of cases of vagueness; and then argue that, given that such cases exist, it will not be intellectually repugnant to accept, also, the existence of some actual-world, ontic indeterminacies of the above sorts. This strategy can be employed without deciding the relation of such indeterminacies to cases of vagueness. I thank Brad Armour-Garb for pressing me to clarify these matters.

<sup>37</sup> For example, epistemicists hold that there is a fact of the matter about at what point the heap gives out, but this fact is necessarily unknowable by us. On the most straightforward application of such ideas to the Jekyll-Hyde case, there is no genuine ontic indeterminacy in that case. Rather, one of the disjuncts in the claim “either Jekyll = Hyde or Jekyll  $\neq$  Hyde” is true at the actual world, even though we cannot know which. However, and as noted above, this position is implausible in the case of story *f* and its characters Jekyll and Hyde.

from another vantage point) sets in. Nor is *f* a philosophical fantasy that asserts that it is positively indeterminate whether or not Jekyll is Hyde.<sup>38</sup> In the world of the realistic Jekyll-Hyde story, each of Jekyll and Hyde is a determinate person who is sharply distinguished from all other people in the story. If this story were made actual, then if we met Jekyll or Hyde, we wouldn't meet a person who was like the indeterminate (if such they are) Mounts Alpha-Beta and Alpha-Gamma. We would meet people who were clearly and sharply identical or distinct.<sup>39</sup> But then it seems that the abstract-character objects Jekyll and Hyde should also be determinate in their identity or distinctness.

To reinforce this point, recall that, as we saw above, the identity and nonidentity of realist characters in the actual world will reflect—and will in fact be wholly determined by—the ontic identity and nonidentity of the characters as they occur in the relevant fictional world. Surely, then, the determinacy or indeterminacy of realist characters in the actual world should likewise reflect and be wholly determined by the ontic determinacy or indeterminacy of those characters as they occur in the fictional world. Thus suppose that a story is ontically vague about where Mount Tripod ceases and Mount Quadruped begins. Or imagine that some philosophical fantasy asserts that it is positively and ontically indeterminate whether or not Mount Jumble is identical to Mount Clutter. Then if the Tripod-Quadruped vagueness is really an ontic matter inside the fictional world, realist principles should require that it is also vague, in the actual world, whether the Tripod character is or is not identical to the Quadruped character. Or, again, given that the positive Jumble-Clutter identity indeterminacy within the fictional world is indeed ontic, realist principles should require that it is ontically indeterminate, in the actual world, whether the Mount Jumble character is identical to the Mount Clutter character. (Of course anti-realists, such as Everett, who reject ontic indeterminacy will then use these points to infer that none of the four present characters is an actual-world entity.)

---

<sup>38</sup> In such a case of positive indeterminacy, it will not just be true both that it is *not f*-fictional that (Jekyll = Hyde) and that it is *not f*-fictional that (Jekyll ≠ Hyde). It will also be true that it *is f*-fictional that [it is not determinate that (Jekyll = Hyde) ∧ it is not determinate that (Jekyll ≠ Hyde)].

<sup>39</sup> Or if these people involved any indeterminacies, it would not be because of the fact that the story itself didn't say whether they were identical or distinct.

However, even if we were to grant, at least for the sake of argument, that these two cases involve allowable sorts of ontic indeterminacy, the Jekyll-Hyde case is different from them both. Jekyll and Hyde are determinately identical or distinct in the *f* world. So the Jekyll and Hyde characters should be determinately identical or distinct in the actual world, too. As noted above, and unlike what happens with Mounts Tripod and Quadruped, there is no vagueness, in the fictional world, about where Jekyll gives out, spatially, and Hyde sets in. Moreover, even if the Jumble-Clutter case should lead realists to some sort of actual-world ontic indeterminacy (which I don't myself suppose), the Jekyll-Hyde case does not involve a Jumble-Clutter sort of positive, ontic indeterminacy within the story world. It is just a confusion to run the Jekyll-Hyde case together with the Jumble-Clutter case (or, for that matter, with the Tripod-Quadruped case) and then to take the problems posed by both cases to admit of a single, non-question-begging realist account that proceeds in terms of actual-world, ontic indeterminacy.<sup>40</sup> Realists thus cannot plausibly escape the Jekyll-Hyde problem by arguing, in the way just noted, for the actual-world indeterminacy of Jekyll's identity or non-identity with Hyde.<sup>41</sup>

Perhaps with more work realists could construct a stronger indeterminacy response to the Jekyll-Hyde problem than I have given above. At the least, a realist might argue from general principles in ways that would require us to say more in order to settle the question.<sup>42</sup> I am skeptical about

---

<sup>40</sup> Suppose that we read about Charlemagne and Charles Martel in the history text imagined above. Then even if that text is our sole source of information about these rulers, we have no inclination to suppose that because the text doesn't tell us whether or not they were identical, in actual reality their identity or distinctness is somehow indeterminate. A similar point ought then to hold for the realist's abstract Jekyll and Hyde characters, as far as I can see. See also Everett (2005)'s rejection of fuzzy-logic solutions to cases like the Jekyll-Hyde one (p. 630).

<sup>41</sup> This point (and the preceding discussion) also shows that the attempt to resolve the Jekyll-Hyde problem by introducing the truth value "undetermined" is unsatisfactory.

<sup>42</sup> Thus van Inwagen (1990, §§18–19), defends indeterminate identity and existence. In correspondence, he has noted the possibility of accepting such identities in order to deal with the Jekyll-Hyde case. I haven't examined his defense of indeterminate identities here (or the further things that other realists, such as Thomasson, might urge in this connection). So it is conceivable that realists who accept indeterminate identity and existence might be able to escape the Jekyll-Hyde problem and others. However, I remain doubtful. For one thing, van Inwagen (1990)'s defense involves sets that, in the actual world, don't have a definite number of members. Or else it seems to involve actual-world metaphysical individuals that are not capable of being members of sets as sets are characterized in standard set theory (p. 260). I find such ideas very puzzling. (Here note, further, Nicholas J. J. Smith, 2008. He argues that

the realist's prospects here, however. And, in any case, I don't see any particularly plausible way to extend the present indeterminacy considerations so as to defuse the Jules-Jim problem (and related problems) about inconsistent character identities.<sup>43</sup> We thus have found no satisfactory realist way out of the problems noted in §1. I have not tried to consider every

---

we cannot make clear sense of vague identity, as that idea is defended by van Inwagen and others, for we cannot model such identity set-theoretically in any way that makes sense. See also Williamson, 2002.) For another thing, the sorts of cases that van Inwagen appeals to in motivating his defense are those in which the identity or nonidentity of the relevant entities is supposed to be vague. (See his pp. 241–2 example in which an infernal machine disrupts Alpha's life or physical constitution in such a way as to yield Omega, who then is supposed not to be determinately identical to, or distinct from, Alpha.) But, as I have stressed above, the Jekyll-Hyde case is not one of ontic (or other) vagueness. Even if ontic cases of vague identity actually exist (as I haven't agreed), the Jekyll-Hyde case isn't such a case. The existence of actual-world ontic vagueness thus does not immediately support the existence (or the intellectual respectability) of any actual-world Jekyll-Hyde indeterminacy. In this connection, it is worth noting that, in any case, the Jekyll-Hyde problem arises out of the fact that intentional (and intensional) phenomena—such as its being fictional or believed that so-and-so is the case—may characterize their objects in incomplete or inconsistent ways. That fact points to a non-ontic, representational source for the Jekyll-Hyde indeterminacy. It does not point to a kind of ontic indeterminacy that belongs to the objects, Jekyll and Hyde, in any actual-world existence that they might turn out to have. For this sort of reason, I also think that it is wrong to locate the source of the Jekyll-Hyde problem and of similar problems in any sort of imprecision in our usual concept of a character of fiction. Such problems arise out of what is allowed by the intentionality (and intensionality) of “according to the story, it is true that ...” and similar locutions. They do not arise out of any failure of precision in the meaning of the term “character” or in the concept of a character, a failure that calls for remedy by our somehow “precisifying” that concept (for example, by replacing (v) and (vi) by (P\*) or (P\*\*), in the way discussed above). The concept of character is already, in the relevant respects, precise enough. (It is the concept of those central, distinctive figures that stories are, *de re*, about—here see fn. 68 and 79.) The Jekyll-Hyde problem no more shows a need to precisify the concept of character than the existence of a story that leaves open the exact nature of the triangular shape of an object (equilateral, isosceles, or scalene) shows a need to precisify the concept of a Euclidean triangle. On these matters, see, further, Howell (2010), §IV.

<sup>43</sup> Thus in the world of Everett's dialethic story it's not indeterminate that Jules = Jim and indeterminate that Jules ≠ Jim. In that world, the identity and the nonidentity of these characters are sharply determinate and crystal clear. And then (as argued in the Jekyll-Hyde case) that determinacy should carry over to the Jules-Jim character(s) in the actual world. So, in the actual world, it will also be clearly and determinately the case that Jules = Jim and Jules ≠ Jim. I might note again Schnieder & Solodkoff (2009)'s response to Everett. They make interesting points, but I think that they underestimate the force of Everett's examples. Thus they in effect argue that, given that the story leaves open the identity or distinctness of Jekyll and Hyde, the identity principles that they call Identity and Identity\* demonstrate the clear, actual-world *distinctness* of Jekyll and Hyde. So, they suggest, there is no problem here for a fictional realist (p. 143). However, and as my §3 discussion of the (i)–(vii) argument

escape route that one might devise. But I don't know of others that look more promising than those above, and I lack the space to carry the present search further. It seems that we should consider alternatives to realism.

## 5 A New Approach: (I) Reading History

Given the attractions of realism, this last conclusion is daunting. If we do not recognize actually existent characters, then what are we to make of the various claims that we noted in §1, including for example the general claim that there exist such-and-such characters in a story (or that there exist more characters in one story than in another)? Or how understand our claims that Jane and Samuel both think about the same character or that a single character occurs in different fictions? As indicated in §2, make-believe theorists offer pretense accounts of these claims, accounts which realists reject. But, as many philosophers have stressed, such claims do provide a strong *prima facie* case for realism.<sup>44</sup> I will say a little more about such claims below, but I will postpone any detailed discussion of them until a further essay.<sup>45</sup> Here I will simply take it that, given the discussion in §§3 and 4 above, the Everett problems show that fictional realism must be rejected. And I will focus on developing an alternative to realism that continues to respect what seems to me most plausible in the overall realist position.

---

shows, a further principle (vi) about distinctness demonstrates the actual-world *identity* of Jekyll and Hyde, given that the story leaves open Jekyll's identity or nonidentity with Hyde. So Schnieder & von Solodkoff have not escaped the bite of Everett-style cases. They have simply given one half of an argument that leads realism into an actual-world inconsistency (claim (vii) of §3). As implied by my comments above on (P\*) and (P\*\*), I also think that their new identity principles distort central parts of the concept of a character.

<sup>44</sup> Such claims are "external" claims about the existence and nature of characters and of our psychological relations to them, not "internal" claims about what happens in the world of the story. They are examples of what Everett (2005, p. 625) calls "fictional-object sentences". (Note also Walton's 1990 distinction between ordinary and nonordinary statements, pp. 396–411.) Everett suggests that the truth of various such claims yields the strongest positive arguments for realism; and Thomasson (from whom I have borrowed this terminology) has made a similar point in correspondence about her artefactual, actualist fictional realism. See also my arguments in Howell (1979). Van Inwagen's defense of realism by appeal to general sentences that quantify over characters (see §1) focuses on a special set of such claims.

<sup>45</sup> See §6 below and especially fn. 67 and 68 and the main paragraph that they tag. In Howell (2010), I discuss such external claims further.



The part of the realist view of our claims about fiction that I find most satisfactory is this. Realists say that these claims—especially the general claims and the same-character claims of the sorts just noted—express literal, not make-believe, truths, truths that are in some sense about characters of fiction. And those attracted to realism usually also accept the related position that, as we standardly make them, straightforward claims such as “Anna Karenina is a woman” or “Huck Finn attended his own funeral” in some way express literally true assertions about the characters that they concern. At least to me, the general view here seems very hard to resist. On their face, all these claims provide examples of statements that in some way function as such assertions and not as pieces of make-believe.<sup>46</sup> In the remainder of this essay, I thus abandon actualist realism in order to explore the prospects for defending something like this bare-bones view. I will suggest a new approach that allows us to accept such a view without granting the actual-world existence of characters and so falling foul of the problems in §3. In the course of doing so, I will also try to say enough about how the new approach would handle the claims in §1 to make it clear why I think it is promising.

The approach that I suggest focuses on similarities and differences between reading history texts and reading texts of fiction.<sup>47</sup> It is a shining fact that, in both cases, we get into cognitive contact with the objects (or putative objects) about which we read by means of our perceptual and cognitive experience of the referring terms and other linguistic units that occur in the texts. Probably few people will deny this point, but I think that it is too often simply rushed past. It is, I think, a fact particularly worth focusing on if we accept the rigid-designator view of proper names that Kripke and others have defended (and if we then suppose that names in fiction function, within the usual story worlds, as such sorts of referring devices).<sup>48</sup> And I will accept such views here and below.

---

<sup>46</sup> As van Inwagen (2003, p. 137, n. 4), notes, it seems just as literally true to say that there are more characters in novel *A* than in novel *B* as it is to say that there are more chapters in novel *A* than in novel *B*.

<sup>47</sup> Of course we also hear stories as well as read them; and blind people touch Braille texts. The account sketched below can be broadened so as to incorporate such facts.

<sup>48</sup> Thus the idea is that (i) in reading the actual-world text of a story, we experience (apparent) names in the text as having the same sort of semantic function that names have in the actual-world texts of newspapers, history books, and other documents that are meant to describe the actual world. And then (ii), given that (as I suppose), in the actual world, names

In particular, we should focus on the usual object-language, actual-world claim that “Anna Karenina is a woman”, this claim being taken by us so as to express, in some way, a truth. This claim is, syntactically, a singular claim, and the term “Anna Karenina” in it purports to function as a proper name of an entity, the character. If there really is, in the actual world, a character (in the realist style) for that term to designate, then this claim will express the structured, singular proposition that we can write as  $\langle \text{Anna Karenina, being a woman} \rangle$ . The object that is the designatum of the term “Anna Karenina” occurs itself in the first slot of this proposition, and the property attributed to that object by the predicate in the claim appears in the second slot. And then we can give that claim its natural, literal meaning: it will express that proposition, even though that proposition is not true at the actual world. Similarly, the natural, literal reading of the sentence “Eleanor Roosevelt is a woman”, as that sentence occurs in a history book, will take that sentence to express the singular proposition  $\langle \text{Eleanor Roosevelt, being a woman} \rangle$ .

However, if we now deny that characters of fiction exist in the actual world, then we cannot take the sentence “Anna Karenina is a woman” to

---

function, in the latter sort of texts, as rigid designators, (apparent) names will also function as such referential devices within our reading experience of the actual-world texts of the stories in which those (apparent) names occur. (Of course there could be very nonstandard stories in which, say, the term “Anna Karenina” functions as a nonce verb, not as an apparent name. But we are not concerned with such odd texts here, or with any nonstandard history texts in which, say, “Eleanor Roosevelt” functions similarly.) Note that the present point is about how we, in reading the actual-world text of the story, experience terms such as “Anna Karenina” as functioning. It is not about how such a term, or a homonymous term, might function “within the world of the story”—how the inhabitants of Karenina’s Petersburg might take “Anna Karenina” to function in her husband’s mouth, for example. (In the usual realistic fictions, such a term, or a homonym, also functions, within the world of the story itself, as a directly-referential, rigid designator. One can imagine nonstandard, philosophical fantasies in which such terms do not function in that way within the world of the story. But as long as the terms function as rigid designators in our experience of reading them in the actual-world texts of those fantasies, my present account will apply.) There are other theories of names on offer, including descriptive theories and theories that combine rigid designation with certain descriptive functions. The kind of view that I sketch below can, I think, be maintained as long as names are taken to be rigid designators. If they are not rigid designators (and if characters cannot be introduced by some means that has the same effect as my use, here, of names as rigid designators), then the present view will not succeed. But standard descriptive views have in part been motivated by the difficulties of dealing with empty names on the directly referential, rigid-designator approach. So if my view seems plausible, we will have less reason to be sympathetic with standard descriptive accounts than we might otherwise have.

be a literal, singular claim that expresses the proposition just noted. Moreover, if we, as actualists, deny that there are any characters of fiction, then we cannot take the actual-world sentence “Anna Karenina is a woman” to be a literal, singular claim that expresses a truth not about an actual-world object but about an object that exists in the world of the novel *Anna Karenina*. (After all, and given actualism, at the world of the novel there is again no entity for our term “Anna Karenina” rigidly to designate. And so no proposition is expressed by that sentence that is true at that world.)

Realists find these results implausible. As noted above, they suppose that such claims (and also generalizations such as “there exists one main character in *Anna Karenina*, namely the woman Anna Karenina”) can be read as literal claims that, properly interpreted, express truths about characters.<sup>49</sup> If we reject the realist idea that characters exist, then it may now seem that we have no alternative but to adopt the make-believe view. So, it may seem, we should accept the idea that such claims are not really literal claims expressing truths of any such sort. Instead, and as indicated in §2, these claims are themselves parts of make-believe and express no truths. Or else, if they express truths, they are (or in some way they imply or they are tantamount to) assertions not about characters but about what, given the rules of the appropriate games of make-believe, it will be fictionally true to assert within those games. Yet I think that realists are right to resist these ways of reading such claims.

We thus appear to be caught in an unhappy dilemma. Either we accept realism and maintain the literal reading of these claims. And then we commit ourselves to the actual-world existence of characters and so, it seems, to the disastrous ontological consequences that §3 has noted. Or else we accept a make-believe view and avoid those consequences by rejecting the actual-world existence of characters. But then we commit ourselves to implausible, make-believe readings of what on their faces are literal claims.

The alternative view of fictional objects that I now sketch separates the question of the actual-world existence of characters from the question of giving literal readings to claims of the above sorts. According to this view, we should agree, with make-believe theorists, that characters do not exist in the actual world. But we should also agree, with realists, that the

---

<sup>49</sup> Of course, as noted in §2, realists can agree that such claims can be used in make-believe. But, so used, these claims will continue to express the relevant propositions.

relevant claims are to be given literal readings; and, on those readings, those claims are, in a certain way, true. Thus we go between the horns of the preceding dilemma by offering a view that is neither a traditional form of realism nor a form of the make-believe theory. Instead, we have an heir of both sorts of theories.

In order to sketch this view, let us return to syntactically singular claims such as “Anna Karenina is a woman” and “Eleanor Roosevelt is a woman”. We will consider these sentences as (let us imagine) they occur in the actual-world text of Tolstoy’s novel and in a text about twentieth-century US history. As noted above, when we read the history text, we get into cognitive contact with the actual woman, Eleanor Roosevelt, by means of our experience of reading that text. And we get into contact, in some sense, with Anna Karenina by means of our experience of reading the novel, even though in the actual world no such person exists. Moreover, in both cases we do so by understanding the relevant sentences, as they occur in the texts. And that fact implies that, by means of psychological processes that involve our perceptually registering and cognitively grasping the syntax and semantics of the words on the page, we arrive at the object-language thought or claim that we put in the form of the relevant syntactic singular sentence. In that thought, we simply read that Eleanor Roosevelt is a woman; or we simply read that Anna Karenina is a woman.<sup>50</sup>

Moreover, and up to the point at which we consciously think (and read) in this last, object-language way, much of this psychological process is non-conscious or at least not explicitly before our attention. Of course, as we read, we sometimes mistake a word or suddenly are confused by an odd bit of syntax; and then we become much more conscious of the reading process than we usually are. (Or, on occasion, we pay explicit attention to the words in an effort to work out the exact claim that is being made. And of course, as we read, we often attend to, and praise or criticize, the author’s style.) The exact details are a matter of psychological fact. It would be helpful to have them specified further, but they are not needed in order to give the gist of the view that I propose.

---

<sup>50</sup> Here and below I ignore pragmatic and other factors that influence our understanding of texts. Introducing such factors will complicate the story that I am telling, but I don’t think that it will essentially alter it.

Whatever exactly goes on as we take in the words on the page, our movement from processing the syntax and semantics of those words to our object-language claims about Eleanor Roosevelt and Anna Karenina can be regarded as involving a kind of disquotations and semantic descent.<sup>51</sup> We begin with a perceptual and cognitive registration of the name “Eleanor Roosevelt”—or of the (apparent) name “Anna Karenina”—and of the predicate “is a woman”, as these bits of language occur in the relevant sentences. We move from that registration, which is usually nonconscious or inexplicit, to the non-metalinguistic, object-language claim that Eleanor Roosevelt (or Anna Karenina) is a woman, a claim that we make seriously and literally and take to be true.

We can model this process of semantic descent in the following way.<sup>52</sup> (I) Let us begin with the Eleanor Roosevelt example. In that case, we confront the public text of the history book, and we perceptually register

---

<sup>51</sup> A number of authors appeal to the ideas of semantic ascent and descent in order to understand existence and nonexistence claims in general, whether these are claims about fictions or claims about supposed real-world phenomena. See, for example, Thomasson (2008) and the earlier work by Hartry Field and Paul Horwich that she cites. Appeal to such ideas can indeed help to illuminate such claims. But here I use the notion of semantic descent, in particular, largely to help us understand our object-language claims (such as “Anna Karenina is a woman”) about the characters of fiction. (In fn. 67, I say a little about negative existential claims about fictional objects, however.) I developed my present use of semantic descent in 2005, independently of these other, more general appeals. While I find them helpful, my use of the idea does not depend on them for its general plausibility.

<sup>52</sup> I here use current views in philosophy of language and metaphysics to provide a sort of model for the process involved. I don’t claim that every stage in the model has psychological reality, and presumably the actual process of semantic descent may involve mental processes that are other than, or are below the level of, the sorts of metalinguistic judgments that I note. (Many different levels of processing may also interact in complex ways.) My claim is simply that something tantamount to the model that I sketch, or something relevantly similar, goes on when we read; and using this model then enables us to throw light on our claims about fictional objects. It would be useful to have the model developed further by psychologically knowledgeable theorists, but that is not necessary in order to give the basic lines of the view that I suggest. It is worth noting, however, phenomena such as the Stroop effect, in which word-meaning/word-color mismatches slow down recognition (for example, the recognition of “blue” printed in red ink). Such phenomena offer indirect evidence for the processing, when we read, of information about word meaning and syntax outside our conscious attention. Recent work also indicates that attention incorporates, at different cognitive levels, a number of different selective processes. (For work on attention, see Harold Pashler, 1998.) For a detailed account of some of the nonconscious neural and other cognitive processes that underlie the activity of reading texts, see Stanislav Dehaene (2009). His points and the work on attention seem consonant with the sort of model I suggest here.

and begin to process the sentence “Eleanor Roosevelt was a woman”.<sup>53</sup> In doing so, we experience the term “Eleanor Roosevelt”, in its occurrence in this sentence, as being an apparent, particular proper name; and we experience the phrase “is a woman” as being a predicate that expresses the property of *being a woman*.<sup>54</sup> As we do so, we make the following (usually nonconscious) metalinguistic judgment:

- (1)  $(\exists x)(x \text{ is an object} \wedge \text{the term “Eleanor Roosevelt” [in its occurrence here in the text] is a particular proper name that rigidly denotes } x \wedge \text{“is a woman” is a predicate expressing the property of } \textit{being a woman} \wedge x \text{ satisfies the predicate “is a woman”})$ .

We also take (1) to be true at the world that we take the text to concern, namely the actual world. (In reading a history text, we do not have to regard such a judgment as true, but I take it that, in the usual case, we do regard the judgment in that way.) We do all these things in what is usually a non-conscious or an inexplicit manner.

Semantic descent can now occur. Given standard logical and semantic principles, the truth of the meta-linguistic judgment (1) is equivalent to the truth of the object-language, syntactically singular judgment that:

- (2) Eleanor Roosevelt is a woman.

So we in effect move from our assumption (1) about the words in the text to an object-language judgment that we take to hold true at the actual world.<sup>55</sup> This object-language judgment expresses the content of what, in the usual case, we consciously judge and take to be true at the actual world

<sup>53</sup> Of course we confront the public text through one or more of its tokens, in the case both of the history text and of the novel.

<sup>54</sup> By our judging that the term “Eleanor Roosevelt” is a particular proper name, I mean that we in effect take that term (in its actual-world occurrence in the text) to be such that it rigidly designates a particular object. Put more exactly, we take that term in this way in the sense that we in effect take the following claim (\*) to be true in regard to the term:  $(\exists t)[t = \text{the term “Eleanor Roosevelt” in its actual-world occurrence in the text} \wedge (\exists x)(t \text{ rigidly designates } x \wedge \text{necessarily } (\forall y)(t \text{ rigidly designates } y \rightarrow y = x))]$ . The relevant part of claim (1) above is to be understood as shorthand for this longer claim (\*). (Given the way that the actual, history-book term “Eleanor Roosevelt” works in claim (\*) through its historical connection with the US civic leader, the unique object that makes (1) true will then turn out to be that civic leader and not, say, the woman who lives down the street and bears the same name.)

<sup>55</sup> Note that the equivalence just noted is simply a more complex case of the equivalence of the truth of the metalinguistic claim  $(\exists x)(\text{“Eleanor Roosevelt” denotes } x \wedge x \text{ is a woman})$

when we read the sentence “Eleanor Roosevelt is a woman” in the history text.

This object-language judgment also commits us to a claim that turns out to have *de re* force. It is now true that we judge that (Eleanor Roosevelt is a woman), and the fact that “Eleanor Roosevelt” in our judgment is taken by us to function as a particular proper name that rigidly designates Eleanor Roosevelt implies that we can also judge that  $(\exists x)(x = \text{Eleanor Roosevelt} \wedge x \text{ is a woman})$ . And, given the actual-world fact that “Eleanor Roosevelt” here functions as a non-empty, rigid designator, the fact that we can make this last judgment implies the explicit *de re* claim that  $(\exists x)[x = \text{Eleanor Roosevelt} \wedge \text{we judge that } (x = \text{Eleanor Roosevelt} \wedge x \text{ is a woman})]$ .

Now of course the mere fact that we begin with the assumption (1) and then descend semantically to (2) does not by itself guarantee that (1) and (2) are true at the actual world and that (2) has the *de re* force that it seems to have. Nor does it guarantee that “Eleanor Roosevelt” actually is, in (1) and (2), the sort of particular proper name and rigid designator that we assume that it is. (To bring these points home, imagine that, totally ignorant of US history, we simply pick up the text, start reading about Eleanor Roosevelt, and then wonder whether we are really reading history or simply some novel that has been got up to look like history.) Instead, these points are guaranteed, in the case of the history text, by the fact that the actual world so-to-speak steps in and discharges our assumption, in judging (1), that (1) is true and that “Eleanor Roosevelt” is the sort of proper name that we are taking it to be. Given the actual-world history of the public text

---

and the truth of the object-language claim “Eleanor Roosevelt is a woman”. The equivalence noted below between (3) and (4) is similar (given the assumption that “Anna Karenina” in the text does indeed denote the relevant object). The equivalence of the truth of (1) and of (2) (or of the two claims just indicated in the present footnote) is of course not an equivalence in meaning—the object-language claim (2) does not mean the same as the metalinguistic claim (1). Rather, the (1)–(2) equivalence holds in the sense that it is necessary that [if standard principles about rigid designation, predication, and properties are true, then (1) holds iff (2) holds]. The equivalence of (3) and (4) in §6 below is then to be understood similarly. Note that, given the way that (1) and (3) are formulated, the present equivalences concern actually-existing terms (specifically, the actual occurrences of those terms in the relevant texts), in the uses that those terms (specifically, their occurrences) actually have in those texts. I am not here discussing singular claims (or the propositions that they express) in general; and I don’t mean to be claiming—wrongly—that the truth of singular propositions about objects implies, in general, either the existence of names for those objects or the truth of metalinguistic claims on the order of, say, (1).

that we are reading, “Eleanor Roosevelt” in its occurrences in that text is indeed a particular proper name for, and rigidly designates, the woman Eleanor Roosevelt, the US civic leader. So the actual-world history of the text (and the actual-world meaning of the predicate “is a woman”, taken together with the actual course of history) guarantees that (1) is true. The truth of (1) then implies, by Modus Ponens, the actual-world truth of (2). The actual world thus guarantees also that there exists, in that world, the singular proposition  $\langle \text{Eleanor Roosevelt, } \textit{being a woman} \rangle$ . And the course of history establishes that proposition as true at the actual world.

## 6 A New Approach: (II) Reading Fiction

Let us turn now to the case of *Anna Karenina*. On the present model, this case exhibits many similarities with, but also some crucial differences from, the case of the history text. As in the history case, we readers confront the public text of a book, in this case the novel *Anna Karenina*. We perceptually register and begin to process the sentence “Anna Karenina is a woman”.<sup>56</sup> In doing so, we experience the term “Anna Karenina” in its occurrence in this sentence as being an apparent, particular proper name; and we experience the phrase “is a woman” as being a predicate that expresses the property of *being a woman*. Unless we are confused, when we experience the term “Anna Karenina” in this way, we do not suppose that that term has that function because someone has produced the text as a record of known fact. Rather, for some other reason we experience this term as having that function. Perhaps we do so simply because we take the term to occur in a text that we know or believe to be a fiction. Or perhaps we do so because, as we read the sentence, we in some bedrock

---

<sup>56</sup> Again we confront the public text through one of its tokens. In this case, that token (and its contained term “Anna Karenina”) traces back to Tolstoy’s initiating act of writing the book. (So that token also traces back to Tolstoy’s act of “creating the character Anna Karenina”, in whatever sense he created that character—I think it is in a sense that can be explicated in terms of the ideas that I am about to develop.) Here and below I ignore complications created by the fact that the original novel is in Russian, not in English.



fashion just do experience it in that way.<sup>57</sup> This experience is usually non-conscious or is at least not explicitly before our attention.

Given that we experience the term in such a way, we make the following (usually non-conscious) metalinguistic judgment:

- (3)  $(\exists x)(x \text{ is an object} \wedge \text{the term "Anna Karenina" [in its occurrence here in the text] is a particular proper name that rigidly denotes } x \wedge \text{"is a woman" is a predicate expressing the property of being a woman} \wedge x \text{ satisfies the predicate "is a woman"})$ .

And we take (3) to be true at a world that we take the text to concern.<sup>58</sup> (In reading the novel, we do not have to regard such a judgment as true, but I take it that, in the usual case, we do regard the judgment in that way. I return below to the question of what sorts of truth and world are at issue.)

Semantic descent can now occur. Given standard logical and semantic principles, the truth of the meta-linguistic judgment (3) is equivalent to the truth of the object-language, syntactically singular judgment that:

- (4) Anna Karenina is a woman

So we in effect move from our (usually nonconscious) assumption (3) about the words in the text (and our assumption that (3) is true) to an object-language judgment that we take to hold true at a world that we

---

<sup>57</sup> It is conceivable that (as Walton's theory of fiction would suggest) we take the sentence and the term to have the function of serving as props in a game of make-believe. If so, then we will be mandated by a certain set of rules of make-believe to take the term to have the function of being a particular proper name. But I don't think that anything like that sort of thing *must* happen when we read fiction. It may be that for some other reason we simply slip into taking the term as appearing to function as a particular proper name, even while we agree, outside this experience, that it does not so function. (Thus we might find ourselves reading a computer-generated text as if its contained apparent names really were proper names of real individuals. Yet that text might not function in any game of make-believe that we are playing with it.)

<sup>58</sup> By our judging that "Anna Karenina" is a particular proper name, I mean that we in effect take that term (in its present occurrence in the text) to be such that it rigidly designates a particular object. Analogously to what we did in the case of the history text, we take that term in this way in the sense that we in effect take the following claim (\*\*) to be true in regard to the term:  $(\exists t)(t = \text{the term "Anna Karenina" in its actual-world occurrence in the text} \wedge (\exists x)(t \text{ rigidly designates } x \wedge \text{necessarily } (\forall y)(t \text{ rigidly designates } y \rightarrow y = x))$ . The relevant part of claim (3) above is to be understood as shorthand for this longer claim (\*\*). (In saying above that we judge that (3) is true, I ignore the possibility of irony, unreliable narrators, and so on, in fiction. I ignore similar phenomena in history texts. Such phenomena will complicate but not undermine the present account.)

take the text to concern.<sup>59</sup> This object-language judgment expresses the content of what, in the usual case, we consciously judge (read) and take to be true at the relevant world when we read the sentence “Anna Karenina is a woman” in the text of the novel.

At this point, however, our treatment of this sentence diverges from the corresponding treatment of the sentence “Eleanor Roosevelt is a woman”. In that case, the actual world—the world at which we took our claim (1) to be true—stepped in to provide a reference for the rigid designator “Eleanor Roosevelt”. And so the actual world guaranteed that the Roosevelt sentence expresses, at the actual world, a genuine singular proposition that involves the sort of *de re* force noted above. But, in the present discussion, we are granting that “Anna Karenina”, in its actual-world occurrence in the text, has no actual-world referent. So the actual world cannot here step in to provide that referent and so to guarantee that our judgment (3) is true and that the sentence “Anna Karenina is a woman” expresses a proposition. Our metalinguistic judgment (3) does in fact express a proposition at the actual world, but it expresses a *false* proposition there. And our syntactically singular judgment (4) expresses no proposition at that world. Taken as a bare sentence standing by itself in the actual world, (4) may represent some attempt to say something literally true. But (4), taken in that way, in fact says nothing at all.

What then explains the conviction, with which I have expressed sympathy above, that (4) does say something that is in some sense literally true? The answer turns on something noted three paragraphs ago: namely, that when we make the judgment in (3), we take this judgment as expressing a truth about a world. We do not have to react to this sentence in this way, but it is a fact about us that, for whatever reason, we do. We read fictions as telling us true stories about a world.<sup>60</sup> When we respond to

---

<sup>59</sup> Actually (3) and (4) will not be equivalent (if we suppose that (4) is without truth value or false if “Anna Karenina” is an empty term) except on the assumption that “Anna Karenina” does indeed denote the relevant object *x*. (Of course a similar assumption is made when we take (1) and (2) to be equivalent. But in that case the actual world discharges the assumption.) I’m ignoring that point here, in order to illustrate how the semantic descent will go. I introduce the needed assumptions about “Anna Karenina” below. For the sort of equivalence here in question, see fn. 55.

<sup>60</sup> Perhaps we react in this way because, as a make-believe theorist might suggest, we are following a mandate that requires us to pretend that such a text expresses a truth. Or perhaps we simply and spontaneously respond to the text in this fashion, even while we would agree, outside this experience, that the text does not really tell a true story about any

the sentence that “Anna Karenina is a woman” by making the judgment in (3) and taking that judgment to be true, we thus in effect assume that the quantifier in that judgment ranges over the objects in the domain of that world. And we assume that the judgment holds true in virtue of the facts about one of those objects in that world.

Supposing, as I will, that we are not here radically confused about the nature of the text that we are reading, that world will be what we call the world of the novel, not the actual world. Now if the world of the novel and its domain of objects existed independently of the story (and of us) as some sort of pre-given entity, then it might seem that that world could now step in and guarantee the reference of “Anna Karenina” in a way analogous to that in which the actual world guarantees the reference of “Eleanor Roosevelt”. So that world could discharge the assumption that we make in (3) and our assumption that (3) holds true in virtue of the facts about one of those objects in that world. If so, our judgment (3) would be true, if that judgment is evaluated at the world of the novel, and the object-language claim (4) would express a genuine singular proposition that is true at that world. So a kind of (non-actualist) realism would hold true in regard to the character Anna Karenina. However, I don’t think that this idea of a pre-given world will do the job of introducing such a character in the way that is here suggested, even if (as many of us do) we accept the general idea of such a world.

In order to explain why this idea will not do that job, and also to see how the job of introducing characters should be done, let me note first that by such a world I mean roughly a domain of objects, with the associated properties and relations that hold of those objects, that exists independently of our claims. That domain then is one with respect to which those claims can be evaluated for truth or falsity. Examples of such worlds

---

world. (Compare fn. 57 above.) My points here about how we experience the text of a fiction are consonant with doubts that I raise in Howell (1996) about Walton’s view of what happens when we “see an object” in a representational picture. According to Walton (1990, p. 293ff.), when such occurs, we always are making-believe, *of* our seeing of the surface of the picture, that there is an object that *that* seeing grasps. But it seems to me that such a make-believe need not enter into picture perception in this way at all. It may simply be that because our experience of seeing the picture is relevantly like our experience of seeing the object, perhaps in some quite abstract way, we spontaneously have an experience that we describe as “seeing the object in the picture”. Like our (usually non-conscious) assumption that the term “Anna Karenina” has a referent, this experience can of course coexist with our awareness, in some other way, that this experience (or that assumption) is not veridical.

would be possible worlds, on one or the other of the usual conceptions of them, or else the sorts of sometimes incomplete and inconsistent worlds that are often suggested as the worlds of a story. Here we will concentrate on the latter such worlds. But, as is often done, we can understand them by appeal to one or another of the accounts that now exist of possible worlds.

Thus, and in line with David Lewis's treatment of possible worlds, we might think of the story world required to make judgment (3) true as being a large concrete object with its own domain of individuals. One of those objects will be the unique object that the term "Anna Karenina", as it occurs in the actual-world text of the novel, rigidly designates. And then in this world that object will satisfy the predicate "is a woman" and will have the property of *being a woman*. Or, again, and in a fashion deriving from Kripke's and others' treatment of possible worlds, we might think of the story world in question as an abstract "way that the actual world might have been". This way would involve the presence of a unique object that the term "Anna Karenina", as it occurs in the actual-world text of the novel, rigidly designates. (This way might then be a possible way that the actual world might be. Or, if the story is inconsistent or incomplete in certain ways, it might be an impossible such way.) Or, finally, and analogously to another treatment of possible worlds, we might consider the story world to be a set of propositions (including singular propositions about the unique, Anna-Karenina object that the term "Anna Karenina" rigidly designates). We might consider the story world to be such a set even if it turns out that the propositions in this set cannot all be jointly true or that this set is incomplete or indeterminate in various ways.

Suppose that there exists such a pre-given story world, distinct from our actual world. The idea now will be that, on any of the above conceptions of such a world, we can introduce a character, in the non-actualist realist way suggested above, as follows. Within this story world (in its domain of objects), there will exist the unique Karenina object, which is the designatum of the term "Anna Karenina" in the actual-world text of Tolstoy's novel and which has the property of being a woman. Hence assumption (3) and so also the object-language claim (4) ("Anna Karenina is a woman") will be true at this world. And the Karenina object, because it does not exist in the actual world itself, will thus be a nonexistent fictional object. Of course this object will be a concrete woman, and so this present idea will not vindicate the sort of actualist realism that we have

been discussing in this essay, which takes characters to be abstract objects existing in the actual world. But we will still have established a kind of non-actualist realism about the characters of fiction.

This idea for establishing such a non-actualist realism will not work, however, no matter which of the above conceptions of a pre-given world we adopt. It is, of course, perfectly conceivable that within the world of the story there occurs a term “Anna Karenina”. Within that world, that term may well then function as a particular proper name that rigidly designates a unique object that is a Russian woman (and so on). If such a thing happens, then, given the relevant understanding of the idea of a particular proper name, there will indeed be a singular proposition, relative to that world, to the effect that the Karenina-object in that world is a woman. So one might think that the above sort of non-actualist realism has thus been demonstrated. The problem here, however, is that there is no unique world of the story which is such that the term “Anna Karenina” occurs in that world and rigidly designates the relevant character object in that world. As Kripke and others have stressed, there will, instead, be many distinct worlds here. Each of these worlds will contain an object—but not the same object for all these worlds—that is rigidly designated by the term “Anna Karenina” as that term occurs in that world. That object will be a woman and will satisfy all the novel’s other descriptions of Anna Karenina. So this present line of thought will not give us a unique Anna-Karenina character of the sort that realism about characters (even the present sort of non-actualist, concrete realism) requires.<sup>61</sup>

Moreover, and from the viewpoint of the position about (3) and (4) that I am here developing, the line of thought in the last paragraph is in any case off-track. According to that position, when we make the assumption (3) and then move to the object-language claim (4), we do not assume (or make-believe or pretend, and so on) that the term “Anna Karenina” itself occurs *within* the relevant world of the story. Instead, when we assume that (3) is true, we make an assumption about *the actual-world term* “Anna Karenina” as that term occurs in (and is read by us as it occurs in) the actual-world text of Tolstoy’s novel. Our assumption is that that single, actual-world term, in its occurrences in the text, rigidly designates a unique object that occurs in the world of the novel, an object that satisfies

---

<sup>61</sup> The point here is part of what in Howell (1979, pp. 139–40), I call the “Kripke-Kaplan-Plantinga problem”. See the next two paragraphs below and Kripke (1980, pp. 157–8).

“is a woman” and so has the property of being a woman. If that assumption were true, then it is clear that we would indeed have arrived at the required unique Anna-Karenina character and the required singular proposition. That character would be a concrete woman existing in some world other than our actual world, and our actual-world term “Anna Karenina” would rigidly designate that woman as she exists in that other world. So we would indeed have vindicated the sort of non-actualist realism (with concrete characters) that was sketched two paragraphs ago.

However, the problem now is that there is no way in which the assumption noted in the last paragraph can be true. How can the actual-world term “Anna Karenina”, as it occurs in the text of the novel, manage to designate an object in some other world (which is not our actual world) that could function as the world of the novel? That term cannot do so through an actual-world dubbing or through any other sort of actual-world historical connection of that term to an object in another world. No such connection runs between these worlds. Nor does that term come to designate any unique object that exists in such another world because, in that other world, that object and no other object satisfies all (or sufficiently many) of the descriptions of Anna Karenina that are given in the novel. After all, there may well be many different objects in each such world (or across these worlds) that satisfy all these descriptions.<sup>62</sup> And it appears that there is no other way (besides this unique-satisfaction way and the dubbing or historical-connection way) in which the actual-world term “Anna Karenina” in the text of the novel could come to designate the required character object in another world. So the idea suggested above fails. We have not arrived at any way in which our assumption of the world of the story can step in and validate any kind of fictional realism.<sup>63</sup>

---

<sup>62</sup> The points here are a form of the Kripke-Kaplan-Plantinga problem mentioned in the last footnote.

<sup>63</sup> It might be objected that, even if this assumption of a pre-given world of the story cannot itself do the job, still the following line of thought may succeed. On the view that I am developing here, readers at least *experience* the actual-world term “Anna Karenina” as being a particular proper name of a unique woman. So, it might be suggested, within each of what we can call the worlds of this experience, there will occur the single term *t* that is identical to that actual-world term. And, in each of these worlds, that term *t* will be a particular proper name of a woman in the sense expressed in claim (\*\*) of fn. 58 above. So in each such world that same term *t* will rigidly designate a unique object. But then that unique object will be the same object for all such worlds, given that that object is rigidly designated by the same term *t* and given that term *t* functions as the same particular proper name in all

Instead of validating fictional realism in such a way, my present view of fictional objects says that the world of the story enters here simply through the assumption that we make as we read the text of the story. This assumption, like our assumption that (3) is true, is usually nonconscious. As my remarks above have implied, it is, in effect, the assumption that the truth of (3) rests on the existence of a unique world or reality that is described truly by the text.<sup>64</sup> We assume that this world involves a certain domain of objects which exist in, and have various properties in, this world. According to our assumption, for each ostensible referring term in the actual-world text that we experience as being a particular proper name, there is a unique object in this world which is such that that term rigidly designates that object. On the present view, this assumption is made (usually nonconsciously) by ordinary, nonphilosophical readers. So it is most natural to interpret this assumption as amounting to the supposition that a certain concrete world with its contained domain of objects exists. Relative to that supposition, the relevant singular propositions that involve various of those objects, such as the Anna-Karenina object, will then also

---

those worlds. So we arrive at a non-actualist, realist-style character after all. However, the answer to this objection is that there is no reason to suppose that that same term *t* functions in all these worlds as the *same* particular proper name. Thus, below, I take us to make the simple assumption that the actual-world term “Anna Karenina” in the text of the novel does indeed rigidly designate one unique object in the world of the story. If we make that assumption, then we do indeed guarantee that the same actual-world term functions in all these worlds as the *same* particular proper name. But, short of making such an assumption, we have no way of guaranteeing, just by appeal to the reader’s experience, that what he or she experiences as a name of one unique object actually is the same particular proper name in all the worlds or circumstances that are compatible with what the reader experiences. Nothing in the reader’s mere experience, considered phenomenologically in an internalist way, guarantees that that term is tied to the same unique object that exists in all those worlds. (Similarly, and without the actual world to guarantee the reference of “Eleanor Roosevelt”, our mere experience of that term as always being *a* proper name of *a* particular woman does not guarantee that there is a unique woman who is such that that term is, in all our experience of it, always the particular proper name of *that* particular woman.)

<sup>64</sup> This assumption, if expanded to cover the entire text of the novel, is equivalent to the assumption that that text (given the sort of syntactic and semantic interpretation noted), is true. Make-believe theories provide one way of understanding how this assumption might be made. But, as indicated earlier, this is not the only way of understanding the assumption. It may simply be a primitive fact about us (from the point of view of what is possible at the level of philosophical description) that, when we read and experience the text, we make such an assumption. And then we maintain that assumption (even when we know that the text is not true of the actual world) by shifting the evaluation of the truth of the text from the actual world to another, fictional world.

exist. And I will understand this assumption in such a way below. However, and as reflection shows, we could also interpret the assumption in terms of the idea of a way that the world might have been or of various sets of propositions. But I will not adopt either of these latter interpretations here.<sup>65</sup>

With this—usually nonconscious—assumption of a world that makes (3) true in place, as well as our original (and usually nonconscious) assumption that (3) is indeed true, we carry out the semantic descent described above. So we arrive at the object-language, syntactic singular judgment (4)—our conscious judgment that “Anna Karenina is a woman”. As we have seen, judgment (4) expresses no proposition at the actual world. However, *relative to the assumptions just noted*, this judgment does express a proposition. That is, *given* those assumptions and the usual semantic principles, it follows that there is a unique object *x* that exists in the world of the novel, and that object is such that it is rigidly designated by the linguistic entity that we have experienced as being the particular proper name “Anna Karenina”. That unique object is the Anna Karenina entity that we can call *ak*.<sup>66</sup>

Given the assumptions just noted, there thus exists the singular proposition  $\langle ak, \textit{being a woman} \rangle$ , and that proposition will be true at the world of the novel. (It will be the semantic content of our actual-world claim that Anna Karenina is a woman.) Moreover, and given that we all in effect make these same assumptions with regard to the same judgment (3), we in the actual world can take that proposition to exist and to be true at the world of the novel. And we can therefore take claims such as “Anna Karenina is a woman” to be literal truths with respect to that world. In addition, and because the assumptions are ones that we all make about the single, actual-world term “Anna Karenina”—a term that occurs in the public text of the novel—we all can be seen to take that claim to express the same proposition. And, given assumptions of this sort about other judgments concerning the names of fictional objects, we can take there to

---

<sup>65</sup> In making the assumption that the world of the novel exists, we thus need not now be assuming, specifically, that a set of propositions made fictional by the story exists, although for theoretical reasons we could proceed in that way if we thought that it was the most plausible way. (Here contrast the initial account of the world of the novel sketched in fn. 5.)

<sup>66</sup> More particularly, given claim (\*\*\*) in fn. 58 and the assumption that (3) holds true at the world of the story, there does indeed exist in that world the unique object just noted, the entity *ak*.



be a single character that occurs in several different novels.<sup>67</sup> In a somewhat similar way, and given that we have made the relevant assumptions, we can also count up the number of fictional objects that the novel *Anna Karenina* introduces and the number of objects that *Madame Bovary* introduces. And we can then assert truly (relatively to these assumptions) that there exist more characters in Tolstoy's than in Flaubert's novel.<sup>68</sup>

The above process of arriving at characters is not vacuous. It would be vacuous if the theory that I am sketching were no more than the view that, after consciously assuming that fictional objects exist and are the referents of various singular terms in a novel, we then immediately turned around and announced that there are indeed fictional objects that those terms des-

---

<sup>67</sup> Roughly, in assuming (3) to be true, we all commit ourselves to the same assumption that the same *de re* judgment is true that is about the same actual-world term "Anna Karenina". As seen, that assumption is that there is an object *x* such that *that term* is a particular proper name of that object *x* in the precise sense of "particular proper name" explained above in fn. 54 and 58. In addition, we all commit ourselves to the same assumption that there is a unique reality that the text of the novel describes. It follows from these commitments that we are all committed (relative to these assumptions) to the same unique object, which exists in the world of the novel and makes (3) (and also (4)) true. (By appealing to analogous assumptions that apply to the text of several different fictions, a similar strategy can be employed to deal with the case of a single character that occurs in all those fictions.) In the sense in which the negative existential claim "Anna Karenina does not exist" is true, that claim can then itself be understood as expressing a literally true claim, relative to the assumption. Given the assumption, we take the object *ak* to exist in what we have assumed to be the world of the novel; and then we note that that object does not exist in the actual world. So, given the assumption, that claim is as it appears to be on its face—namely, a true singular claim asserting nonexistence. (But because it amounts to such a claim not in reality but only relative to the assumption, that claim is not a freestanding claim that is literally true at the actual world independently of any assumptions. So we do not now confront the problem of nonbeing that often faces accounts that take such negative-existential singular claims to express literal truths.)

<sup>68</sup> As reflection shows, and as already suggested in §4, we ordinarily use "character" not to refer to any abstract entity of the realist sort but, instead, to designate the central distinctive figures (human, animal, alien, mechanical) that stories are, *de re*, about. Those are the figures in which we are absorbed as we engage ourselves with these stories and their worlds. (This fact also explains why, as critics of realist abstract character-objects note, we can say, without uttering a necessary falsehood, that "the main character of the novel is a woman". Here see Howell (2005); Everett (2005, p. 644); R. M. Sainsbury (2005, pp. 209–15.) The examples noted in the present paragraph of the main text provide examples of external claims about fictional objects, in the sense indicated in §5 at fn. 44. As observed there, it is such claims that many philosophers take, I think rightly, to provide the strongest positive arguments for realism. My comments in the main text and in the previous footnote indicate how I would handle various of these claims, using the view of fictional objects that I have developed above. But, as also remarked in §5, I postpone further discussion until Howell (2010).

ignite. But that empty view is not what the present theory holds. The present theory holds, to the contrary, that the assumptions in question are usually made non-consciously as we process the text. It is only as a result of our semantic descent (itself non-conscious) that we come consciously to judge our singular claim that Anna Karenina is a woman to be true. And from that conscious point of view, and without our realizing how we have reached it, that singular claim appears to us to express a genuine, actual-world singular proposition about an object, Anna Karenina.

Evaluated at the actual world as a bare, self-standing claim that is not based on any assumptions, that singular claim of course expresses no proposition at all, as we have noted. But we do not realize that we have made this singular claim only relative to a hidden set of assumptions. So, operating unwittingly on the basis of those assumptions, we take that claim, in a realist spirit, to express a literal truth about the object Anna Karenina. As we do so, we evaluate that claim at the world of the novel. And, evaluated relative to that world, that claim does indeed express the proposition  $\langle ak, \textit{being a woman} \rangle$ ; and that claim is indeed then true. (In addition, and relative to these assumptions, we can then see ourselves as making a *de re* judgment about Anna Karenina in the same sort of way that we made such a judgment about Eleanor Roosevelt.) But of course our claim that “Anna Karenina is a woman” really expresses such a proposition (and the object *ak* and such a proposition really exist) *only* relative to the truth of the assumptions in question.<sup>69</sup> So our literal and

---

<sup>69</sup> Moreover, note that the present idea is *not* that (a) there is (in some Meinongian sense) a genuine object *ak* and a genuine proposition  $\langle ak, \textit{being a woman} \rangle$  that have being in a categorical way that is independent of the assumptions in question, those entities then (b) being somehow introduced into the discussion when we make those assumptions. No such entities exist or have being in any categorical way. Rather, the existence of the object and of the proposition is wholly dependent on the truth of the assumptions. These entities exist only relative to that truth. And because, in actual fact, the assumptions are false, the entities do not actually exist at all. (I assume that, unless the assumptions are true, there is no reason to hold that these entities exist. Thus, and as the discussion above shows, the object *ak* does not exist within any fictional worlds in such a way that it is true, in the actual world, that (3) holds and the actual-world term “Anna Karenina” really does rigidly designate that object as that object exists in such a world. And I see no other reason to think that that object exists in such a world.) The assumptions in question enter the discussion only on the basis of inferences made on the basis of the usually nonconscious, false assumptions that, in reading the text, we spontaneously make (or that we perhaps make as a part of some make-believe, although we do not have to take the assumptions in that way). In addition, note that the actual-world term “*ak*”, as it is used in my present, theoretical comments about how readers of *Anna*

true claim about a genuine Anna Karenina object constitutes a literal, true claim only insofar as it is based on the assumptions that we have, usually unbeknownst to ourselves, made. And that claim has a semantic content only relative to these assumptions.

To the extent that, as the realist view holds, our actual-world talk “about Anna Karenina” expresses any literal truths, that talk is thus in effect carried out as a result of the process of semantic descent noted above. And so it is carried out relative to the assumptions that (3) is true and that the relevant terms in the text rigidly designate objects in the world of the fiction. But this talk is *not* itself (and it need not be understood as) a part of any make-believe. (When we claim, in the actual world, that “Anna Karenina is a woman”, we are not making-believe or pretending that that sentence is true, for example.) Nor, contrary to the other standard make-believe treatment of such talk, does our talk constitute any sort of assertion about what it is correct to claim within any game of make-believe that we are playing (or that we could be playing) with the text of the novel.<sup>70</sup> That talk is no more a part of (or an assertion about) a make-believe than are the assertions that we make when we reason, in a deductive argument, from given premises—premises that we take to be true—to a conclusion.

This last point is worth spelling out a bit further. Suppose that we reason as follows:

*Premises (assumed):* (a) any  $x$  is such if  $x$  is Greek, then  $x$  is human; (b) any  $x$  is such if  $x$  is human, then  $x$  is mortal; (c) Socrates is Greek.

---

*Karenina* proceed, is of course not itself a genuine singular term designating a particular object. Instead, it is in effect a term introduced through the existential instantiation of the second quantifier in the readers’ assumption that there is a world, and there is an object in that world, that is such that “Anna Karenina” denotes that object in that world. (By fn. 58, that object will be a unique object. The actual-world term “Anna Karenina” itself—the term that Tolstoy introduced and that readers use in talking about the character—functions in (4) as a rigid designator of that unique object, given the readers’ assumption. But of course it is not such a designator independently of the truth of that assumption. For the points here, see also Howell, 2010.)

<sup>70</sup> Nor, again, does our talk constitute any assertion to the effect that the underlying metalinguistic points in (3) hold and that the assumptions in question have in fact been made. Of course if the present view is right, our being in a position to make the actual-world claim that Anna Karenina is a woman and to take that claim to express a singular proposition depends on the holding of those points. And it depends also on those assumptions’ being made. But that claim itself (and my making of it) in no way asserts or is used, pragmatically, to imply that such matters are the case.

*Reasoning from (a), (b), and (c):* (d) if Socrates is Greek, then Socrates is human. (e) Socrates is human. (f) If Socrates is human, then Socrates is mortal. (g) Socrates is mortal.

In this familiar derivation of Socrates' mortality, (d) of course follows from (a) by Universal Instantiation; (e) follows from (c) and (d) by Modus Ponens; (f) follows from (b) by Universal Instantiation; and (g) follows from (e) and (f) by Modus Ponens.<sup>71</sup> Thus each of the steps (d) to (g) after the assumptions (a) to (c) ultimately depends on those assumptions for its assertion in the derivation. But none of these steps is a bit of pretense or make-believe. (Nor does the claim made in any of these steps express a commentary on what it would be true to assert within a make-believe. Nor, for that matter, does the claim made in any of these steps express an assertion about what it is correct to assert within this sort of derivation.) Rather, the claims made in all of these steps constitute literal, actual-world assertions, although each claim is asserted only given (or relative to) the assumptions (a) to (c).<sup>72</sup> Nor need these assumptions themselves be made as part of a case of pretending, making-believe, simulating, or imagining, in any of the senses that would make those terms suitable to express a properly formulated make-believe theory. Rather, in making these assumptions we simply *take them to be the case*, in a perfectly literal sense, in order to derive what follows.<sup>73</sup> The same goes for our (nonconscious) assumption of the metalinguistic (3) and of the truth of (3). And it then

---

<sup>71</sup> The comments in the present sentence that explain the steps in this derivation are of course not themselves part of the derivation. The derivation amounts simply to the sequence of steps (d) to (g), each of which follows from previous steps (including the premises) according to one of the rules of deductive reasoning that is here employed.

<sup>72</sup> For these claims in steps (d) to (g) to be asserted on their own, on the basis of this derivation, the assumptions (a) to (c) must of course be true (as they indeed are in this case). But, in the derivation as it stands (and without our noting anything about the truth or falsity of those assumptions), the claims in steps (d) to (g) are still asserted, although they are asserted only on the basis of their derivation from the assumptions.

<sup>73</sup> Note that, on standard make-believe theory, claims made within the overall make-believe are themselves only pretend assertions. (Or else our claims about characters are indirect actual-world assertions about the correctness of making various assertive moves within a game of make-believe.) However, assertions made on the basis of an actual-world assumption are ordinary, literal actual-world assertions and do not themselves become mere make-believes or claims about make-believe. To make a claim on the basis of an assumption is not to perform a new, non-assertive speech act with that claim or to prefix that claim implicitly with the operator "we make-believe that". (To take a further analogy, suppose that I assume that there was a unique person who first climbed Mt. Everest. And then I reason about that person by inferring, for example that, given further premises, "the first climber had

applies to our subsequent derivation of the object-language (4) on the basis of (3) and of our further assumption that, roughly, there is a unique world with respect to which (3) is true. (Of course in the *Anna Karenina* case we do not make these assumptions in order to prove any result. We make the assumptions non-consciously in the course of reading the text and exploring, through that reading, the world that the text presents to us, given the assumption.)

Despite the remarks in the last several paragraphs, it is conceivable, I suppose, that someone might still argue that the present view is not that much different from make-believe or pretense accounts of our claims about fictional objects. On the present view, such claims are not asserted on their own, in the actual world, but only as parts of an overall activity of making the assumptions and then proceeding as if those assumptions were true. Yet, it might be objected, isn't behaving in this way very much like performing a Waltonian activity of making-believe that the text of a fiction is true and then producing sentences such as "Anna Karenina is a woman" merely as parts of that activity? But, however natural this objection may seem to those who are immersed in pretense theory, the remarks in the last few paragraphs show that it is mistaken.

After all, and as we observed in §2, Walton's view is that in producing "Anna Karenina is a woman" in the actual world, we are (at least in many cases) merely copying the behavior that we would employ if we were making a genuine assertion about a real woman named "Anna Karenina". And we copy that behavior because we are simply carrying out a pretense of asserting that there exists a woman by that name.<sup>74</sup> However, on the present view this idea is wrong. When we produce such a claim in the actual world, we are *not* copying any sort of assertive behavior while nevertheless not making a genuine assertion of our own. Nor—if we are not literally copying a genuine case of assertive behavior—are we in some

---

a great deal of tenacity and endurance". Then this last claim is an ordinary literal assertion, made on the basis of my assumption, that will be true given the truth of the assumption and of those further premises. This claim does not have some special, further psychological or pretense status.)

<sup>74</sup> Of course, and as noted above, Walton and his followers agree that we may indeed make a claim such as "Anna Karenina is a woman" as an actual-world assertion. But then we are simply commenting on what it would be correct to assert within a game of make-believe played with the text of the novel. This idea is not at issue here, however. (And, as also argued above, we are not making any such comment in the usual case in which we make the claim that "Anna Karenina is a woman" as an actual-world assertion.)

other way merely pretending or “making-as-if” the sentence is true, to use an expression of Yablo’s.<sup>75</sup> Rather, and as I have emphasized above, we are genuinely and literally asserting, by means of our utterance of that sentence, the singular claim that Anna Karenina is a woman. As was just stressed in the discussion of the (a)–to–(g) syllogism, the fact that we

<sup>75</sup> Yablo (2005, p. 96ff., and 1998, p. 251). The idea derives from Walton (1990); see also Walton (1997). As I have indicated above, on my present view, and at the most basic level, we use these sentences to make genuine assertions, and we do not copy the behavior of someone who asserts such sentences without actually asserting them ourselves. (Nor, if we somehow do not literally copy such assertive behavior, do we engage in some other sort of simulation or making-as-if activity that is still not itself the making of a genuine assertion.) As indicated in fn. 57 and 60, the present view allows that such copying or making-as-if can also occur in some cases. It is just that it is not the core of what goes on when we make claims such as “Anna Karenina is a woman”. Make-believe theorists often talk of imagining, making-believe, pretending, making-as-if, and so on, in fluid ways that make it difficult to decide what does and does not count as a clear case of a genuine make-believe theory. (Thus Walton, 1990, chap. 1, countenances non-occurrent, non-episodic, unconscious imaginings; and he suggests at p. 21 that the term “‘imagining’ can, if nothing else, serve as a placeholder for a notion yet to be fully clarified.”) Such fluidity is not necessarily a bad thing—it may be cognitively useful during theory-building. And in various broad, loose senses, I suppose that my discussion of assuming (3) and then deriving (4) might be described, using Yablo’s general, abstract phrase, as a case in which we “make-as-if” (3) (and then (4)) is true. However, Yablo’s official explanations of making-as-if appeal to Walton’s ideas about pretending and imagining. Thus Yablo (2005) says that “the thread that links as-if games together is that they call upon their participants to pretend or imagine that certain things are the case” (p. 96; the same sentence, with “as-if” replaced by “make-believe” occurs in Yablo, 1998, pp. 245–6). And, in any case, I take the distinctive core of clear, genuine make-believe views—when they are applied to the case of fiction—to be that when we utter sentences such as “Anna Karenina is a woman”, we do *not* make any literal assertions. Rather, we simply make-believe or pretend or do-as-if we are making such assertions. (As before, I ignore the here-irrelevant cases in which, according to Walton and his followers, when I produce such a sentence, I am literally asserting something about what action it is proper to perform within the relevant game of make-believe.) This core usage seems present in various of Yablo’s explanations of his own view. Thus Yablo (2001, p. 84, fn. 11) describes a certain action of saying sentence *S*—an action which I take to be of his make-as-if sort—as the action of “*saying* it [sentence *S*] meaning: *S* (pssst—judge this by its faithfulness to reality as we are supposed to *imagine* it when doing arithmetic)” (my emphases). If that sort of claim (with its attached “pssst” qualifier) is what we really *mean* when we produce *S*—or when we make claims such as (4) (“Anna Karenina is a woman”)—then our production of *S* and of such claims does not seem to be a simple, actual-world act of asserting these claims as true. But, on my view, our usual issuing of (4), on the basis of our assuming (3), is such an actual-world act of assertion, and it does not lose that status when we take (4) (and (3)) to be evaluated for truth not at the actual-world but at the world of the fiction. (See also fn. 77 below.) So I take it that my view is not a case of a Yablo-style make-as-if position, if such positions are cases of genuine make-believe theories of fiction. What counts, for make-believe theorists, as making-believe and making-as-if needs further discussion, I think.

are making that assertion relative to a set of assumptions does not mean that our action is not one of producing a literal, actual-world assertion. It means only that that assertion is not a freestanding, categorical assertion but is made relative to the assumptions in question.

We can put matters here clearly by returning to the comparison between the present *Anna-Karenina* object-language claim (4) and the Eleanor-Roosevelt object-language claim (2). On my present view, both of these claims are made relative to actual-world assumptions—that (1) is true, in the case of the Roosevelt claim, and that (3) is true, in the case of the *Karenina* claim. In the case of the Roosevelt claim, the actual-world steps in, in the sense noted above, to guarantee that the assertion made in (2) does indeed express the true singular proposition, which holds in that world, that Eleanor Roosevelt is a woman. In the case of the *Karenina* claim, the actual world cannot so step in. Instead, we shift the semantic evaluation of that claim to the world that we take to exist and to make the text of the story true. Relative to the assumption that there is such a world, the *Karenina* claim counts as a genuine and true assertion. It is an assertion that holds true of an object in that world and that expresses the singular proposition that that object is a woman. The shift in evaluation here does not alter the actual-world, assertive status of the claim itself or turn it into some sort of bit of make-believe or making-as-if. It simply changes the range of objects about which we take ourselves to be talking.<sup>76</sup> Logically

---

<sup>76</sup> Of course make-believe theorists suppose that when I make-believe that the *Karenina* claim is true, I in effect shift the domain of evaluation of that claim from the actual world to the world at which I pretend that that claim holds true. So my present view shares with such theories the general idea of domain shifting. But, as I have stressed, these theories are built around the claim that when, in the actual world, I issue the *Karenina* claim, I am not asserting that *Karenina* is a woman. Rather, I am only making-believe that—or making as if—I am asserting that point. (Indeed, they seem to accept the idea of domain shifting simply as part of their view that such make-believe occurs.) And here the present view diverges sharply from such theories. (See also fn. 77 below.) In a future essay, I hope to develop these points further. I also hope to apply the present view to the general problems of empty names and to Peter Geach's Hob-Nob puzzle. And I intend to say more about our assumption that there exists a world of the story. One last remark is worth noting here, however: on pretense views, the function of operators such as "according to *Anna Karenina*, it is true that" is as follows. They allow us to indicate in the actual world, in a literally true way, the make-believe, singular-proposition content of claims such as the *Karenina* one. (On pretense views, those claims of course have no actual-world propositional content when they are made as parts of the pretense itself.) On standard realist views, we use such operators to indicate the actual-world, singular-proposition content (with regard to the relevant abstract objects, as characters) that such claims have. On my present view, such operators need not

and semantically, the situation is, in the relevant respects, thus similar to that in which, while standing in my room, I say that “all the books are in Russian”; but I take that sentence (a literally true assertion) to be evaluated not with respect to the objects in my room but with respect to objects in the university library (or with respect to objects in some library that we intend to build). Or I take that sentence to be evaluated with respect to one time and place in history rather than with respect to another.<sup>77</sup>

---

yet enter into the basic case in which, in the actual world, we read the text of the fiction, non-consciously make an assumption such as (3), and then descend, semantically, to the singular claim that Anna Karenina is a woman. Rather, we introduce these operators later, in order that we can mark the shift, in our evaluation and understanding of these claims, from the actual world to the relevant fictional world. In order to avoid confusion, we need a linguistic means of making this shift explicit. In other cases (for example, when two fictions make differing claims about the same character), we also need such a means. We need this means for roughly the same sort of reason that we may need a means of saying consistently, for example, that it is true that all the books are in Russian (said with regard to the university library shelf) but that it is not true that all the books are in Russian (said with regard to the room in which I’m working).

<sup>77</sup> *Objection:* you admit that both our assumption that (3) is true and our subsequent, derived claim (4) (“Anna Karenina is a woman”) are claims that are evaluated with respect to an assumed fictional world. But that fictional world is not actual. Nor is it any part of, or any realm within, our actual world. So for us, in the actual world, to issue (4) and to take (3) and (4) to be evaluated with respect to this assumed fictional world, you have to suppose that we effect domain-shifting (of the world of evaluation of (3) and (4)). But we can bring about such shifting only by taking (3) and then (4) in effect to be prefixed by some intensional operator such as “imagine/make-believe/make-as-if it is true that ...”. So, after all, our issuing of (3) (and then (4)) *is*, in your hands, ultimately part of a make-believe or a making as if. Contrary to what you’ve argued, your theory in the end is just a variant on make-believe theories such as Walton’s and Yablo’s. *Reply:* Not so. It is true that *one* way, in the actual world, of getting claims such as (4) (and (3)) evaluated for truth at a fictional world rather than at the actual world is to prefix those claims, implicitly or explicitly (as we issue them), by such an operator. Doing so does indeed turn our issuing of (4) into a piece of make-believe. Moreover, I agree that we might proceed in such a way (compare fn. 57 and 60). And, as we become engrossed in a fiction and its world, we may conceivably do so often. However, and as I have argued, I do not think that such is at the core of our response to fiction. Rather, when we read the text, we simply and spontaneously, for whatever reason, take the term “Anna Karenina” to function as a rigid designator of an individual. And so, without engaging in any make-believe, we make the assumption (3) (usually in a non-conscious way) and move on to (4) by semantic descent. As we do so, our issuing of (4) functions, in the actual world, as a literal assertion that is (relative to our assumption of (3)) about the object Anna Karenina—the entity, in the fictional world, which we have assumed the term “Anna Karenina” rigidly to designate. When we take (4) to be evaluated for truth at the assumed fictional world and not at the actual world, our actual-world issuing of (4) does not lose this assertive status. Our issuing of (4) does not become merely part of a make-believe that we are asserting that (4) is true. To make explicit (or to change) the domain at which a claim is to be evaluated for truth is not, in and of itself, to change the assertive (or other speech-act)



We may note, finally, that the Everett problems do not arise given the present view. The characters of which we speak do not exist in the actual world. (Nor, for that matter, do they occur as nonexistent entities in some categorically given Meinongian realm.) These characters are taken to exist only relative to the relevant assumptions. So no actual-world ontic indeterminacies or inconsistencies come into play. Thus suppose that we make (in a non-conscious way) the relevant assumptions about the fiction *f* above. Then, given those assumptions, there will be a disjunctive proposition that is true with respect to the relevant fictional world and that says that (either Jekyll = Hyde or Jekyll  $\neq$  Hyde). However, given the assumptions (and given *f*'s silence about whether Jekyll is indeed Hyde), we do not have as true, with respect to that world, either the proposition that Jekyll = Hyde or the proposition that Jekyll  $\neq$  Hyde. Such a situation is impossible in the actual world, which (let us suppose) is fully determinate and consistent. But no problem arises here as long as we do not suppose that the fictional world must itself be fully determinate and consistent. And we suppose no such thing.<sup>78</sup> Moreover, we can then handle the Slynx and Jules-Jim problems by similar appeals to propositions that exist—and are true at the fictional worlds—only relative to the relevant assumptions.<sup>79</sup>

---

status of the claim itself. And it is not to change that status even in those cases, like the present one, in which the new domain is itself something that we simply assume to exist but that does not itself, in actuality, exist. As a reader, I issue (4) by semantic descent from my (usually non-conscious) assumption (3). As I do so, I take (4) (and also (3)) to hold true at the fictional world that I assume exists and renders true the text of the novel. This process is a process of my making an assumption about (4) itself in the course of my issuance of (4). But it is not a process that turns my original act of issuing (4) into some sort of mere make-believe (or make-as-if), performed by me, to the effect that I am asserting that (4) is true. Rather, as I issue (4), I so-to-speak have my eye on the relevant, assumed fictional world, and I assert that (4) holds with respect to that world. The case is no different from that in which, while asserting that my computer is turned on, I assume or otherwise indicate that that assertion holds true in the actual world itself. In neither case does the assertive status of my claim disappear when I indicate that my claim holds true at a certain domain.

<sup>78</sup> Of course anti-realists such as Everett agree. He in effect points out (pp. 642–3) that fictional worlds—for him, pretense worlds—can be indeterminate and inconsistent.

<sup>79</sup> Finally, the identity, distinctness, or indeterminacy of characters remains, on the present account, a matter that is wholly determined by the identity, distinctness, or indeterminacy of characters as they occur in the relevant fictional world. (Compare §3 and §4 above.) However, on the present account this point does not hold about characters considered as actual-world, abstract objects. Instead, it holds simply because the characters *are* the relevant entities that, given the assumptions in question, exist in those worlds.

## 7 Conclusion

Where does this discussion leave us in regard to realist views of fictional objects? As I have noted, actualist realism must itself be abandoned. Evaluated at the actual world, our talk of fictional objects does not express any singular propositions about such objects. In this sense, we must agree with the anti-realist opposition and, in particular, with the implications of recent make-believe accounts of fiction. However, we are not therefore forced to agree with the standard, make-believe view of the nature of claims such as “Anna Karenina is a woman” or “there are more characters in *Anna Karenina* than in *Madame Bovary*”. We can maintain the realist view that these claims express literal truths if we take these claims to express such truths only relative to the assumptions that we make (usually non-consciously) in the course of reading the relevant fictions. The present view thus embodies neither realism nor any standard make-believe approach. It is better thought of as a descendant of both positions. It has various loose ends as it stands, and I have not developed it in full detail. But it provides an alternative that allows us to accept what seems true in these two positions while avoiding the difficulties that exist for them both.<sup>80</sup>

---

<sup>80</sup> This paper has benefited greatly from discussions with Brad Armour-Garb and from a conversation with Amie Thomasson and her later e-mailed remarks. Comments by John Woods and others at the April 2009 APA symposium on fictions in literature, mathematics, and science were very helpful. So were e-mailed responses from Kendall Walton and Peter van Inwagen. In addition, I appreciate the encouragement extended to me by my colleagues at Moscow State University during 2007-2008. And I thank the Fulbright Program, the SUNY Center on Russia and the United States, and the University at Albany for making my stay at MSU possible. I valued the chance to lecture on fiction in spring 2006 at the Kaliningrad Immanuel Kant State University. My work on Kant at the Institute for Advanced Study in summer 2006 also contributed indirectly to the ideas in this paper, and I am grateful for the opportunity to have been in residence there. Reactions by David Braun, Walton, and Armour-Garb to a 2005 realist version of my §§5–6 ideas also helped. (See fn. 63.) None of these fine people and institutions is responsible for whatever imperfections may remain in the discussion above.

# The Fiction of Creationism

FREDERICK KROON

## 1 Introduction

According to the semantic-ontological doctrine of *creationism in fiction* (“creationism” for short)<sup>1</sup>, fictional objects exist—they *really* exist, they don’t just exist in relevant stories. And they exist as a result of being created by the authors of these stories. More precisely, creationism holds that i) if something is a fictional object, then it exists as a result of being created through the imaginative activities of the author(s) of the work in which the object appears,<sup>2</sup> and ii) there exist fictional objects, among them fictional objects designated by fictional names. So creationism is an

---

<sup>1</sup> I shall use the shorthand when it is clear that fiction is the target. (As I emphasize later, creationism is a type of doctrine that might be advanced for other discourses as well.) The term “creationism” has been used informally for some years, and was adopted more or less officially by Yagisawa (2001) and so used by, for example, Caplan (2004) and Voltolini (2009). The creationist literature is a rapidly growing one, and includes the book-length treatments of Thomasson (1999) and Voltolini (2006) as well as such seminal papers as Salmon (1998). Kripke (1973) also advocates a version of creationism.

<sup>2</sup> I here leave out the role that consumers of fiction might play in a creationist account of fictional objects; the importance of such a role is clear from the possibility of fictional objects that are created through a certain non-fictional work having acquired the status of fiction by being so treated by readers. (For an influential creationist account of the role of consumers of fiction, see Thomasson, 1999.) Along a different dimension of generality, note that the category *fictional object* is best thought of as encompassing not only fictional *individuals* (named or merely described), including those too insignificant to be labelled *characters* in the standard sense, but also fictional places, kinds, events, and so on. We should also include fictional *fictional* characters like *Hamlet*’s Gonzago, since they too are thought to be created, this time as fictional objects within a fiction (see Voltolini, 2009, pp. 43–4).

amalgam of the thesis that there are fictional objects (*fictional realism*), including those named by fictional names, and the thesis that all fictional objects have their genesis in certain imaginative activities. We might say that the specifically *creationist* part of this brand of fictional realism results from a literalist reading of the latter claim: just as a sculpture is an artefact created over a period of time by a sculptor, so a fictional object is an artefact created over a period of time by an author—not a concrete object this time, of course, but an abstract object, wrought not with hands but through the imagination.<sup>3</sup> Creationism therefore takes fictional objects to be *abstract created artefacts* (indeed, some call it the *artefactual(ist)* theory of fictional objects).

So described, creationism is clearly very different from a classical (neo-) Meinongian theory of fiction,<sup>4</sup> which holds that there is no time at which a fictional object begins to exist (fictional objects do not exist at all for classical Meinongianism); at best, a fictional object can be singled out for use in a fictional work by being selected to this end; the author contributes the selection, not the object. But the views of creationists and Meinongians are very similar in other ways, and there are contexts at which even hardened creationists will liken their view to a version of Meinongianism. Both approaches, after all, admit to the genuine objecthood of fictional objects, as well as to a sense in which these objects don't exist. No less significantly, both acknowledge two kinds of property-attributions. Meinongians invoke the nuclear/extranuclear distinction or (in the case of Zalta's *abstract object* version of Meinongianism) the encoding/exemplifying distinction.<sup>5</sup> For classical Meinongianism, Hamlet has the *nuclear* property of being, during his life, a Danish prince and the *extranuclear* property of being more famous than any existing prince of Denmark, while on Zalta's version Hamlet *encodes* the property of being, during his life, a prince of Denmark and *exemplifies* the property of being more famous than any existing prince of Denmark. Creationists admit their own version of this duality, and sometimes say that fictional objects

---

<sup>3</sup> Such an analogy is often used by creationists; see, for example, Fine (1982, p. 130). (Despite the analogy, Fine himself does not think that authors bring characters into existence but only into actuality. The difference between these two notions is not easy to understand.)

<sup>4</sup> Parsons (1980) and Routley (1980) offer influential neo-Meinongian treatments of fiction. Priest (2005) offers another version of Routley's approach.

<sup>5</sup> See, for example, Zalta (1988).

*have* properties but also *hold* properties. An object like Hamlet *has* such properties as being more famous than any existing prince of Denmark, and it *holds* such properties as being, during his life, a Danish prince. In addition, creationists think that while Hamlet doesn't *have* this latter property (no abstract object can be a flesh-and-blood Danish prince), Hamlet does have it according to Shakespeare's *Hamlet*.<sup>6</sup> So there is broad agreement about the properties we can, in one sense or another, properly attribute to fictional objects. But Meinongians and creationists disagree sharply about some of the (extranuclear) properties it is proper to attribute to fictional objects. For creationists, Hamlet has or exemplifies being an abstract fictional character that did not exist until it was created by Shakespeare; not so for Meinongians.

Fictionalism rejects this picture of the semantic underpinnings of our talk of fiction and its objects. More precisely, it rejects the idea that fictional realism offers the best semantic account of our talk of fiction and its objects, and a fortiori that a species of realism like creationism offers the best account of our talk of fiction and its objects. Fictionalists think that there is some overarching story or perhaps some overarching game of make-believe according to which such objects exist and have their various properties, and that this is enough to give substance to our talk of truth in this area. Some, but not all, fictionalists also declare that it is *only* in the fiction that such objects exist, that there really are no such objects. (Sometimes fictionalism about Xs is in part defined in terms of the denial that there are Xs.)

Fictional realism, especially of the creationist variety, is an increasingly popular view of the semantics of fiction among analytic philosophers. And of the various kinds of fictional realism, it is probably creationism that has in recent years acquired the largest following.<sup>7</sup> My own preference is for a version of fictionalism, but in this paper I will not so much argue for fictionalism as pose a problem for perhaps the best-known argument for creationism. The problem involves parity-of-reasoning considerations, and the argument it very naturally suggests is that, since the

---

<sup>6</sup> On some formulations, the notion of *holding* a property is rejected in favour of the relativised notion of *having* a property according to the relevant story. See, for example, Thomasson (1999).

<sup>7</sup> See, for example, Braun (2005), Goodman (2003), Howell (1996), Kripke (1973), Predelli (2002), Salmon (1998), Soames (2002), Thomasson (1999), and Voltolini (2006, 2009).

kind of considerations that allegedly support creationism would also, by parity of reasoning, support quite unacceptable versions of realism, we should be sceptical of the support they allegedly provide to creationism. My own view is that the considerations in question are best seen as providing support to a rather different kind of theory—a fictionalist creationism—and in the last section of the paper I briefly describe and motivate this theory.

## 2 A Popular Argument for Creationism in Fiction

Historically, perhaps the most compelling argument for an abstract object version of fictional realism is due to van Inwagen. Van Inwagen reminds us of sentences that are *prima facie* true (true *simpliciter*, not just true in some work of fiction), and that commit us to the existence of objects of fiction if taken at face value—sentences of literary criticism like:

- (1) There are characters in some 19th-century novels who are presented with a greater wealth of physical detail than is any character in any 18th-century novel

and:

- (2) Some characters in novels are closely modelled on actual people, while others are wholly products of the literary imagination, and it is usually impossible to tell which characters fall into which of these categories by textual analysis alone.

According to van Inwagen, these sentences taken at face value entail that there *exist* characters in 19th-century novels and characters who are closely modelled on actual people: the quantifiers in (1) and (2) are just the existentially loaded classical quantifiers. Other literary-critical characterisations entail that other kinds of characters exist. Assuming, as van Inwagen thinks, that such sentences should be taken at face value as assertions that quantify over fictional characters, it follows that there exist characters, and hence fictional objects, of the corresponding kinds. He also thinks that there are true literary-critical sentences featuring fictional names rather

than quantification that again should be taken at face value and show the same commitment to fictional objects. Take, for example, a sentence like:

- (3) Mrs. Gamp [...] is the most fully developed of the masculine anti-woman visible in all Dickens's novels.<sup>8</sup>

(The same commitment can, of course, be derived by using quantificational sentences:

- (4) There are a number of masculine anti-woman characters in Dickens's novels, and the most fully developed of these is Mrs. Gamp,

which directly entails:

- (5) There is/exists such a character as Mrs. Gamp.)

Van Inwagen thinks that the *prima facie* truth of such sentences of literary criticism provides compelling reason for accepting the existence of fictional characters (indeed, of *all* fictional characters), including the existence of particular characters such as Mrs. Gamp and the existence of fictional characters of various distinctive literary-critical kinds. If so, it is difficult to resist the further conclusion to creationism. Maybe it is hard to find compelling examples of literary-critical statements that directly imply that fictional objects are artefacts. What is not hard, however, is finding *prima facie* true literary-critical statements that imply that fictional objects were created by authors. For example, a literary-critical claim like (1) is surely on a par with a claim like:

- (6) When authors create fictional characters, they present them with more or less physical detail, but in the 19th-century there were authors who presented some of the characters they created with a greater wealth of physical detail than had been done in any 18th-century novel.

There is no reason to suspect that (6) is any less worthy of being taken at face value than (1)–(3), this time as a claim that attributes a relationship of creating between authors and fictional characters; such a claim is certainly no less assertible. And once we allow such claims it is only a short step to

---

<sup>8</sup> Van Inwagen (1977, p. 301), quoting Manning (1971, p. 79).

fully fledged creationism; for fictional objects that are *created* by authors are created artefacts—to be an artefact is to be an entity that has been created through the creative artistic activities of agents.

I shall call the inference from the availability of sentences such as (1)–(6) to the truth of creationism in fiction the “*Master Argument for Creationism in Fiction*” (MAC<sub>F</sub>)—the epithet “*Master*” reflects the fact that while it is by no means the only argument found in the creationist literature there is a sense in which it holds pride of place. Suppose, for example, that you are persuaded by van Inwagen’s very influential argument for an abstract-object version of fictional realism. Given the problems that attend the idea that authors have the ability to create abstract objects,<sup>9</sup> your embracing creationism would be surprising if you did not also think you that there were *prima facie* truths, such as (6), that license the thought that such objects are created by authors. In fact, an appeal to MAC<sub>F</sub> seems to feature implicitly whenever creationists try to explain how the creation of fictional objects occurs. Searle, for example, thinks there is a causal connection between pretense and creation:

By pretending to refer to people, and to recount events about them, the author creates fictional characters and events. (Searle, 1974–1975/1979, p. 73)

Similarly, Schiffer claims that:

Fictional entities are created in a straightforward and unproblematic way by the pretending use of names. (Schiffer, 1996, p. 157).

Such views gain much of their intuitive force from the pervasiveness of talk about authors’ creating fictional characters and events (as in (6)), talk that gives rise to the strong presumption that there are fictional entities and that authors somehow create them. The pretense story strikes many philosophers as compelling not just because it aptly describes the activity of fiction-making but also because it explains how this kind of creation might be effected. Searle’s and Schiffer’s remarks should, I think, be seen in that context; that is, should be seen as providing an explanation of the

---

<sup>9</sup> Van Inwagen himself is wary of the inference to creationism because of just such problems (van Inwagen, 2003, p. 153f.).



claim that there are fictional entities and that authors create them, where this claim is established by something like  $MAC_F$ .

Just how pretense can explain the genesis of fictional objects is a controversial matter. Schiffer, for example, thinks there is a constitutive connection between the notion of pretending and the idea of fictional objects. Others offer more detailed hypotheses about how pretending on the part of authors create fictional objects and what this means for the existence and identity conditions of fictional objects (see, for example, Thomasson, 1999, and Voltolini, 2006). Such views face a number of problems, but in the present paper I am not concerned with the hypotheses or the problems they face.<sup>10</sup> I am concerned only with the background argument— $MAC_F$ —that takes the creation of fictional objects to be a fact in the first place. My target in this paper will be argument  $MAC_F$  itself.

### 3 A Problem for the Argument I: The Case of Diaphanous Entities

Despite its apparent popularity, there is reason to think that as it stands  $MAC_F$  is flawed: the argumentative strategy it instantiates over-generates. For there are other sentences possessing the same logical structure as (1)–(6) that seem no less assertible but that quantify over/refer to “entities”—I’ll call them *diaphanous*—that even hardened creationists would be reluctant to admit as bona fide entities. Such sentences subserve what I shall call *MAC-type arguments*, but this time for creationism about entities that we should repudiate. If so,  $MAC_F$  needs to be supplemented in some way to rule out such putative counterexamples.

To begin with, notice how the need for supplementation already shows up on the version of creationism defended by Nathan Salmon, himself one of its foremost advocates, and one of the most liberal as far as the *reach* of creationism is concerned. After encouraging the adoption of “a very inclusive attitude towards such applicants for Existence as fictional characters, mythical planets, fabricated boyfriends, and flights of fancy” (Salmon, 1998, p. 305), Salmon then wonders: “Where does it all end?”

---

<sup>10</sup> For criticisms of the idea that pretending can bring about the existence of fictional objects, see Yagisawa (2001, pp. 154–7). Brock (2010) gives a more general range of criticisms of the idea of created fictional objects.

The answer he gives is that nothing whatsoever corresponds to names that we introduce to stand for things when there is no expectation that there is such a thing, and when we are not story-telling. An example of a name that is, in his terms, “thoroughly non-referring” in this sense is the name “Nappy”, introduced to stand for the actual Emperor of France if there is such a person, and to be non-referring otherwise, where the introducer believes no such person exists:

Barring a fairly radical skepticism, we know that there is no such person as Nappy. Nappy is not a fictional character, not a mythical character, not a fabrication, not a flight of fancy. There is a very good reason why Nappy is none of these things. Not to put too fine a point on it, Nappy does not exist. (Salmon, 1998, p. 306)

It is not hard to show, however, that this way of imposing limits on creationism is inherently unstable. For names are communal currency. If I encounter a sentence containing the name “Nappy”, without knowing the context and while catching a glimpse of the whiteboard in Salmon’s office, I may well decide that “Nappy” stands for someone. If I then want to inquire about that person, I’ll use the very same name to make my inquiry, intending (as Kripke taught us) to use the name to refer to the very person that this earlier use of the name referred to. And it results in the uncomfortable situation that such later uses have, or can have, all the hallmarks of the sort of uses that must stand for abstract entities on Salmon’s way of thinking. As Salmon might himself point out: “Nappy isn’t a real person; he is just a person I made up to help me convince philosophers that some negative existentials involve thoroughly non-referring names”. Or, to put the point in an explicitly quantificational manner,

- (7) There are all kinds of things that are posited simply as a result of people misapprehending what others say; things like Nappy, construed by some as a real person but created by me to convince philosophers that some negative existentials involve thoroughly non-referring names.

Using a MAC-type argument, we can now deduce that “Nappy” on this use stands for an abstract created entity. But if it does, earlier uses of the name on my part and Salmon’s part, including the use of “Nappy” in Salmon’s

target sentence “Nappy doesn’t exist”, stand for the same abstract entity. After all (again, a point of Kripkean orthodoxy), there is only a *single* name in contention here.<sup>11</sup>

Here is an extreme way of making the point. One can be wrong not only about the referential status of a *name* but even about the nominal status of a *word*. Take a word like “max” as in “He went to the max[imum]” (meaning something like: “He gave it his all”). Suppose that Smith overhears the “max” part as I loudly urge my most recent crop of PhD graduates to “Go the max when giving your job talk! Don’t be wishy-washy!” Smith begins to wonder: “Who is Max and why shout at him?” After Smith informs various friends about my behavior, one of them finally approaches me to hear my side, and I answer:

- (8) You and Smith’s other friends needn’t worry about poor Max. Max doesn’t exist: he is just one of Smith’s unwitting inventions (there are others!), the result of his mishearing my instruction to “go to the max” as talk about a real person called “Max”.

Again, (8) yields a MAC-type argument for the conclusion that there exists an abstract object Max: an unwitting creation, whose origins lie in Smith’s confusing a noun for a proper name.

In my view, the conclusion that there are such things as Nappy and Max is unacceptable, and I take it that many creationists would agree—if not about Nappy, then at least about Max.<sup>12</sup> The situations characterized in (7) and especially (8) are simply too thin to deserve description in realist, let alone creationist terms: the “entities” in question are unacceptably diaphanous. Although we may use such nouns as “creation”, “invention”, and the like, as well as the machinery of quantification, to characterize the mistake made by those who in all seriousness employ these words as genuine names, the use of such language seems at best figurative and doesn’t

<sup>11</sup> Kroon (2003) contains an earlier version of this sceptical argument against Salmon. See also Caplan (2004), which provides a sustained argument for the weaker claim that if fictional objects are created entities for the reasons Salmon gives, then there is every reason to think that entities like Nappy are also created entities. Caplan is concerned only to argue that creationists about fiction should accept these further commitments; he doesn’t pass judgement on creationism itself.

<sup>12</sup> Braun (2005) and Caplan (2004) both report that Salmon says he aimed to be neutral in Salmon (1998) about whether Nappy exists as an object of the imagination (despite the fact that he calls “Nappy” a “thoroughly non-referring” name). An “object” like Max, however, cannot even be classified as an object of the imagination, except in some very weak sense.

license the conclusion that there are such objects as Nappy and Max. And this in turn provides us with a reason for being suspicious of  $\text{MAC}_F$  as it stands. It alerts us to the fact that we shouldn't trust our first impressions as to the literal truth of sentences like (1)–(6). Why should they be treated any more seriously than (7) and (8)?

Perhaps, however, van Inwagen intended his argument for realism to rest on a more substantial premise than the availability of prima facie true quantificational locutions. Perhaps he assumed that there was an appropriate filter on the prima facie truths we are to take as literally true. After all, van Inwagen takes the entities provided for by his argument to be *theoretical entities of literary criticism*<sup>13</sup> and that suggests that the sentences that feature in the argument are rather special: they form part of a particular sort of discourse, the sober quasi-scientific discourse of literary criticism. By contrast, the wording used in (7) and (8) might well strike us as relatively strained and ad hoc, designed to undermine the very referential practice that it is in some sense party to, and not part of a serious, robust description of the world. (We might, as I did, label the language *figurative* rather than *literal*, but these are so far little more than labels. The question is how to identify assertively uttered sentences that bear their literal interpretation, and are thus able to reveal genuine ontological commitments.) To mark this point, let us say that sentences like (1)–(6) satisfy the condition of belonging to *a sustained, stable and sober quasi-scientific discourse* (condition **S**, for short). We might say that condition **S** functions as a filter on when we are to take prima facie truths at face value, able to sustain inferences to ontological claims, and (hence) as literal rather than figurative truths. It is this fact, I am now suggesting, that plays a quite pivotal background role in  $\text{MAC}_F$ . If so, it may be the failure of (7) and (8) and their ilk to satisfy condition **S** that explains why they do not provide support for realism about “objects” like Nappy and Max.

---

<sup>13</sup> Van Inwagen (1977, p. 302).

## 4 A Problem for the Argument II: The Case of Imaginary Companions

I have suggested that the sober nature of discourse about fictional entities found in literary-critical discussion might well have a critical role to play in  $MAC_F$ . Because sentences like (7) and (8) don't belong to sober, stable discourses, on this construal of the argument they no longer threaten  $MAC_F$ .

Unfortunately, this is not enough. For even if sentences like (7) and (8) no longer threaten the argument, other sentences, belonging to other discourses, do.

Thus consider the case of imaginary friends or companions: friends or companions who don't really exist but that children *imagine* as existing. It is now well understood that the invention of imaginary friends is extremely widespread among young children (close to 25% of four year olds have such friends).<sup>14</sup> Such friends come in all shapes, sizes and looks, and have a bewildering range of personal characteristics. These characteristics often determine how the child interacts with the friend. A powerful friend, for example, might be called upon to ward off perceived dangers, while a weak friend might be bossed around or treated with particular sympathy. Consider, for example, Humberto Nagera's case<sup>15</sup> of a child's imaginary blind brother whom the child invented after he had punched his own brother and was accused of nearly blinding him:

The brother is younger, like his own brother, but his blindness makes him totally dependent on him; he cannot go anywhere without him and never wants to leave him. Being with his older brother, walking with him, feeling his arm over his shoulders, or sitting close to him, is the happiest experience for the blind brother. (Taylor, 2001, p. 67)

In such descriptions, we can distinguish properties that the child gave the imaginary companion in creating him and properties that are based on the child's choices about how he and such a companion will interact (for

<sup>14</sup> See Taylor (2001) for some of the evidence for such claims.

<sup>15</sup> Nagera has been one of the early investigators of the phenomenon of children and their imaginary companions (see especially Nagera, 1969).

example, the child might decide to go for a walk with his blind brother on a certain day; as a result, the blind brother might well be particularly happy that day, while the child himself might feel pleased and proud as he imagines the accolades from passers-by who notice his treatment of his blind brother).

In addition to claims about the personal characteristics of imaginary companions and the actual behaviour of such companions in the light of these characteristics, there is another category of statement commonly made about imaginary companions. People make statements about the *creation* of such companions: why they are created, at what point they are created, why they are created with this or that set of characteristics, whether the creation of such companions is psychologically harmful, and so on. Parents of children with imaginary companions have long had an interest in the last question, but there is now a large and sophisticated psychological literature devoted to all such questions. Here, for example, is a representative quote from Marjorie Taylor's influential *Imaginary Companions and the Children Who Create Them*:

I believe that fun and companionship are the primary reasons most children create imaginary companions. Some of the imaginary beings who fulfil these purposes are zany characters like the Cat in the Hat.<sup>16</sup> They reflect the child's idiosyncratic interests, have characteristics that are not particularly logical or internally consistent, and evolve to suit the whims of their creators. (Taylor, 2001, p. 64)

We can now state the problem that these various facts pose for  $\text{MAC}_F$ . First of all, a point about discourse about imaginary companions and **S**. Earlier we said that for the argument to work there must be appropriate statements that are *prima facie* true and satisfy condition **S**. Even a slight acquaintance with the literature shows that the first part of this claim is true. Discourse about imaginary companions is replete with the kind of quantificational and categorical claims appealed to in  $\text{MAC}_F$ . Here, for example, is Taylor again, this time with a claim that closely resembles a literary-critical claim like (2):

---

<sup>16</sup> Note that the Cat in the Hat is a fictional character in the Seuss story *The Cat in the Hat*, but also (at least, on the standard reading of this story) an imaginary companion to the story's narrator.

[Some] imaginary companions are more closely modelled after playmates of the child's own age, size and gender. (p. 64)

Indeed, all of the statements that van Inwagen uses in his classic paper have their analogues in statements that psychologists have made about imaginary companions. But not only is discourse about imaginary companions replete with such quantificational claims; it is also clear that they meet condition **S**—discourse about imaginary companions has all the hallmarks of being stable, sober, and sustained, with experimental and theoretical research on the topic of imaginary companions now a burgeoning part of developmental psychology.

If  $\text{MAC}_F$  works for creationism in fiction, then, by parity of reasoning, there is also a good MAC-type argument for creationism about imaginary companions. That is, we should count imaginary companions as genuine abstract artefacts created by the children who invent them. The only way to resist this argument, it seems, is to impose yet another requirement on taking such discourses at face value, another way to filter out the potential for MAC-type arguments. For the moment, however, it is hard to see what this can be. If anything, the discourse that developmental psychologists participate in when talking about imaginary companions should be treated as *more* scientifically respectable than the discourse of literary criticism: this is science after all, encompassing both experimental and theoretical components. From this point of view, discourse about imaginary companions is in better shape than the discourse of literary criticism. Against the latter, Yagisawa has charged that:

[L]iterary criticism is not even a social science. It is not a science of any kind. Its main aim is not to discover truths, but to help enhance aesthetic and other kinds of experience by the readers of literary or fictional works. (Yagisawa, 2001, p. 164)

No such evaluation can attach to the psychological literature devoted to imaginary companions.

Nonetheless, I take it there are overwhelming reasons to resist creationism about imaginary companions. Even if fictional objects are created artefacts, this can't be true about imaginary companions. It is often said that we can't interact with fictional characters—we can only have intentional attitudes towards them, and this asymmetry is consistent with,

and may even suggest, their being abstract. (More on this later.) But there is no such asymmetry in the case of imaginary companions. In reporting on their dealings with imaginary companions, children don't only report on their liking or loathing of the companions. They report on their playing with the companions, or watching their companions get into mischief. This symmetry is retained in the sober reflections on imaginary companions by psychologists. Taylor comments, for example, in a passage whose beginning I quoted earlier, that:

[Some] imaginary companions are more closely modelled after playmates of the child's own age, size and gender—happy active, and loving children with everyday names like Joel or Susan. Who could be a better partner in play than an imaginary friend? ... The child doesn't have to worry about an imaginary companion getting cranky and threatening to take his or her toys and go home. On the other hand, the child can walk out on the imaginary companion at any point without repercussion—the companion will be cheerfully ready to start up again at a moment's notice. (p. 65)

Note how the way in which imaginary companions are said to be modelled on real children in this passage makes it clear that the companions are supposed to be concrete individuals, just like real children. Even the demise of imaginary companions is described in entirely concrete terms by the psychologists writing about such companions. In Taylor's words,

Children do not seem to mourn the passing of imaginary companions. This point underlies the utilitarian nature of their usefulness. They tend to be abandoned when they have outlived their usefulness.<sup>17</sup> (p. 120)

In short, the vocabulary that psychologists as well as their subjects use to describe imaginary companions is highly concrete, with creation seen as a special "bringing into existence" of friends who have the special feature of being imaginary, with the demise of such creatures described in equally concrete terms. The contrast this reveals with the way creationism in fiction conceptualizes its objects can be usefully set out as follows.

---

<sup>17</sup> Children have also been known to arrange the death of their companions (sometimes simply by announcing their funeral; see Taylor, 2001, pp. 84–5).



Recall that for creationism in fiction, fictional objects *have* some of their properties and *hold* others: the former are the properties that they have absolutely, apart from the work, and are properties appropriate to their status as abstract objects; the latter are the properties that the objects have in the fictional work in which they appear, and are properties appropriate to the kinds of objects they are in those works—typically, they are properties of concrete objects such as the property of being a flesh-and-blood Danish prince. In terms of this division, the properties that imaginary companions *hold* are the properties with which children invest them when they create them, while the properties they *have* are the other properties, including properties that reflect aspects of the way their creators relate to them and the way psychologists theorize about such relationships. But as we have seen, these other properties *are no less properties of concrete objects*: there simply is no concrete/abstract division in the case of imaginary companions matching the division that creationists posit in the case of fictional objects (or if there is, it is so theoretically sophisticated that it seems simply perverse to attribute it to the way children and their parents, no less than psychologists, talk about imaginary companions). It is not plausible, therefore, to construe talk of imaginary companions as talk of special abstract objects created through children’s imaginative activities.

We can now draw the two strands of the foregoing discussion together. (i) If  $\text{MAC}_F$  is a compelling argument for creationism in fiction then, by parity of reasoning, there is also a compelling MAC-type argument for creationism about imaginary companions. (ii) There is no compelling MAC-type argument for creationism about imaginary companions: talk about imaginary companions is not talk about abstract created artefacts of some kind. Hence,  $\text{MAC}_F$  is not a compelling argument for creationism in fiction. Whatever force the argument seems to have should be explained in other ways.

Before we ask how, then, to understand the kind of locutions that have been used to subserve a van Inwagen-style argument for creationism in fiction, let me briefly revisit the question of whether there might be conditions other than *S* to rule out a MAC-type argument applied to discourse about imaginary companions. Such a condition would challenge the soundness of the above argument by removing the grounds for premise (i). Earlier I expressed scepticism that we could find such a condition, but it might now appear that the argument given for premise (ii) yields an

obvious candidate. Unlike talk of fictional objects, talk of imaginary companions portrays the creators of imaginary companions as able to interact with their creations. Arguably, only a discourse where there are no *prima facie* truths about the interaction between creators and the objects they create should be admitted as licensing a MAC-type argument for creationism about such entities (we might call this the “No interactionist truths” condition on the discourse, **N**). The suggestion is that this excludes our being able to run a MAC-type argument for creationism about imaginary companions, while leaving  $\text{MAC}_F$  itself unscathed.

I suspect, however, that **N** is plausible only on a very selective reading of what is found in literary-critical discourse. Literary criticism is in fact replete with claims that suggest that authors interact with their creations in somewhat the same way as children interact with imaginary companions. Authors don’t just create their characters. They might also *live* with them in order to get to know them better, in the course of which they might grow to love or despise them. Novelist Alice Walker reports that while writing *The Color Purple* she lived with her characters for a year, accommodating them even to the extent that she moved from New York City to Northern California because they didn’t like the city congestion and the tall buildings. More generally, “fictional characters have been known to haunt, amuse, teach and infuriate their creators”.<sup>18</sup> And some characters acquire so much fame that creation has in a sense become a two-way process:

Dante, like Petrarch, falls in love with his own creation. What else can Beatrice be? And since she is the *Comedy*’s greatest originality, does she not in turn create Dante? (Bloom, 1994, p. 103)

Eventually, authors might kill off their characters (at the stroke of a pen, without remorse, etc.) because the characters have acquired traits that deeply annoy the authors, or because the authors feel that the characters have begun to exert too much power over them—because, as it might

---

<sup>18</sup> The first line of “Who Created Whom? Characters That Talk Back”(The New York Times, May 31, 1987). Taylor, Hodges & Kohányi (2003) catalogue a range of cases (like that of Alice Walker) in which fiction writers experience what they call “the illusion of [characters’] independent agency” (IIA). Fifty writers were interviewed about the development of their characters and their memories of childhood imaginary companions, with forty-six reporting at least some experience of IIA.

seem, the characters now direct the stories in which they appear, not the authors.

These are all familiar claims, and literary criticism turns out to be full of instances of such claims. No doubt creationists about fiction will see them as figurative rather than literal, but there is no good reason to distinguish them on such grounds from the literary-critical statements ((1)–(6), for example) that sustain  $\text{MAC}_F$ .<sup>19</sup> Indeed (here recalling Yagisawa’s criticism of creationism cited earlier), to the extent that such statements talk about the psychology of authors they belong to a part of literary criticism that is explanatory in a psychological sense rather than simply aesthetic, and so are, if anything, more scientific in spirit than (1)–(6). The only reason why they are not highlighted as much in literary criticism is that literary criticism is often far more product-focused—not surprisingly, given the interest in the aesthetic properties of literary products.

In summary: Both the theoretical discourses under consideration—about imaginary companions and about fictional objects—admit interactionist truths. So if  $\text{MAC}_F$  is to have a fighting chance, we should *reject* **N** as a further filter to be placed on discourses that yield to MAC-type arguments for creationism. With premise (i) thus left unscathed, its combination with premise (ii) provides us with a strong reason to reject  $\text{MAC}_F$ .

## 5 Fictional Creationism

I have argued that since there is no plausible MAC-type argument for creationism about imaginary companions, there is no plausible argument of this type for creationism in fiction either. But nothing I have said is thereby intended to denigrate the kind of talk that subserves MAC-type arguments applied to these discourses. (And, of course, nothing in what I have said shows that there are no other arguments for creationism, perhaps working in conjunction with an argument like  $\text{MAC}_F$ .) Indeed, there is this much to be said for both creationism in fiction and creationism about imaginary companions: the creationist locutions that subserve such arguments deal with significant and serious theoretical issues—issues that deserve to be

---

<sup>19</sup> Apart from anything else, standard literary-critical statements themselves often contain elements that strike readers as less than fully literal, suggesting that they are not to be taken fully at face value (not, at least, without subjecting them to quite drastic reformulation). See Yagisawa (2001, pp. 164–5), for discussion.

labelled scientific in at least the broad sense. What I have disagreed with is not the seriousness and significance of such locutions and discourses but a familiar kind of argument for a popular realist account of the “entities” apparently assumed in one of these discourses (that of literary criticism).

To return to the question I asked towards the end of the last section, how, then, are we to understand such locutions? This is, of course, a large and difficult question, and not one I plan to discuss in any detail. Assuming, however, that there are no other, more compelling arguments for creationism in fiction, I suggest that we might just take our cue from some of the features of our discourse about imaginary companions—features that fuelled the doubts I have expressed about  $MAC_F$ . It is plain, in my view, that psychologists’ talk of imaginary companions succeeds because it enters into the spirit of the pretense that children themselves are involved in when they “create imaginary companions”. There strictly are no such individuals, but the fiction that children really do create or invent such individuals is one that is worth adopting if we are trying to understand—from the inside, as it were—how children’s imaginative activities affect the way they live their lives. So we should be fictionalists about such talk rather than creationists: truths about the imaginary companions that children create are true only from the point of view of such a fiction. In short, we should be fictional creationists.

Likewise, we should be fictionalists about discourse involving talk of fictional characters and their creation, and for much the same reason. Authors pretend, in their writing, that there are such flesh-and-blood individuals as Hamlet and Holmes, just as children pretend they have their little invisible friends. And both parties want their audiences to pretend along with them, although for very different reasons. On this perspective, when critics or psychologists come along with such attributions as *being created/made up by so-and-so*, *being modelled on such-and-such*, *being killed off/resurrected for such-and-so a reason*, and so on, they too are continuing the pretense, this time in order to describe, from the inside, the ground, content, style, and history of the creative imaginative activities in question. Such pretense seems continuous with the kind of proto-psychological and proto-critical pretense that children as young as two engage in when commenting on imaginary companions or story-characters. They will readily say things like: “Otto is not real, he is just a *pretend*-friend; I made him up because I wanted someone to play with”. Similarly,

a frightened child is quite likely to ask her parent for reassurance when being told a story containing a frightening character: “The monster isn’t real, is it? He is just pretend?” Children are good at pretend-play; locutions of this sort suggest they are also adept at locating and describing features of the pretense from inside the pretense.

Of course, much more ought to be said in defense of such a fictionalist perspective, and this brief characterization in fact fits a number of different ways of being fictionalist.<sup>20</sup> If I am right, however, the phenomenon of imaginary companions presents us with a case where a perspective of this kind is intuitively rather more appealing than some kind of abstract realism about imaginary companions. Given the deep correspondences we have noted between theoretical talk about imaginary companions and literary-critical talk about fictional characters, there is good reason to find a fictionalist creationism about fictional objects no less appealing.<sup>21,22</sup>

---

<sup>20</sup> There is, for example, the *semantic* pretense approach championed in Walton’s seminal work on make-believe (see, for example, Walton, 1990, and Crimmins, 1998) as well as the *pragmatic* pretense approach (as in, for example, Kroon, 2004). For a fictionalism in the same anti-realist spirit as pretense-oriented fictionalisms but eschewing pretense, see Brock (2002). Brock offers paraphrases of claims like van Inwagen’s (1) and (2) that use special in-the-fiction operators; given a target statement  $\Phi$ , the appropriate paraphrase declares: “In the fiction according to which (concrete) realism about entities of such-and-such a kind holds,  $\Phi$ ”. Given its sophisticated paraphrases, such a fictionalism cannot easily be extended to children’s early uses of statements like “Otto isn’t real; I made him up”. That, I think, counts as a mark against it. See also Sainsbury (2009) for a rather different anti-realist account framed in terms of the idea of presupposition-relative truth.

<sup>21</sup> Good but *defeasible* reason. Defeaters might include other, more compelling arguments for creationism (whether creationism about fiction, imaginary companions, or both). If sustained, such a fictionalist perspective suggests that talk of fictions is not so very different after all from the kind of talk about diaphanous entities like Nappy and Max that initially fuelled our worries about  $\text{MAC}_{\text{F}}$ . For utterances of (7) and (8) also appear to involve pretense—this time a “going along” with the language of those who seriously think there are such things, in order to make a point about the disreputable origins of this way of talking.

<sup>22</sup> I am grateful for helpful comments on an earlier draft of this paper from an audience at the *World Congress of Philosophy* in Seoul in 2008. Special thanks to Stuart Brock, Jeff Goodman, and Alberto Voltolini.



# Virtual Worlds and Interactive Fictions

GRANT TAVINOR

## 1 Virtuality and Fiction

There is an obvious plausibility to the idea that many digital artefacts are works of fiction. The videogame *Grand Theft Auto IV* depicts a series of events in a place called Liberty City that are imaginary rather than real. Niko Bellic, the protagonist in the game, is a fictional character much like any number of fictional characters to be found in film and literature; indeed he has thematic precedents in those traditional fictive media. The game also depicts a fictional narrative that follows Niko's arrival as an immigrant in Liberty City, his immediate frustration when he learns that life in the city is not quite what it was made out to be by his cousin Roman, and his search through the city to resolve one of the dark secrets of his past. None of this really occurs, rather it is fictional that it does so.

But *Grand Theft Auto IV* is in many respects quite unlike other works of fiction. The setting of the game, Liberty City, is represented by a richly dynamic 3D graphical model, filled with citizens, traffic, buildings and parks, and a broadcast media that can all be interacted with in detailed ways. The city has a changing weather and time of day, has a complicated geographical layout, and can be traversed on foot, in a car, by using the subway, or even flying a helicopter. Niko, unlike protagonists in more familiar fictions, is under the control of the player: just what Niko does,

where he goes in the city, and the activities he pursues, is substantially up to the player. Hence the player is represented in the gameworld through a fictional proxy so that he can get drunk with Roman and get a taxi home, browse internet dating sites to arrange dinner dates, and, of course, get up to the criminal hi-jinks for which the game is notorious. Furthermore, the narrative that plays out in the game depends on the actions of the player. At certain junctures of the game, the player is asked to make a choice that has ramifications for the direction in which the narrative develops, and so the player, as the protagonist, has a role in shaping the story.

All of this means that what we would think of as the fictive scenarios in videogames are robustly represented, interactive, and open-ended in a way that they are not in the case of fictions such as novels, TV shows, and films. This representational robustness, interactivity, and openness sets these items apart from traditional kinds of fiction, so much so that when people refer to these digital works they often employ a different terminology to that used to discuss films and novels. Videogames and other kinds of digital works are referred to, in somewhat common parlance, as *virtual worlds*, *simulations*, or instances of *cyberspace*, the latter term owing to the fiction of William Gibson. This vocabulary has been employed in a widely ranging theoretical literature that pertains to these digital developments (Castronova, 2005; Heim, 1991; Wertheim, 1999; Woolley, 1992; Zhai, 1998). Indeed, some writers go so far as to say that this new technical vocabulary is appropriate in a way that traditional terms are not: games scholar Espen Aarseth has claimed that the various features of digital artefacts picked out above—their representational and participative robustness—makes them virtual *rather than* fictional (Aarseth, 1997, 2005).

Thus, there is an ostensible variation in the ways we might characterise videogames and other digital artefacts as presenting fictional worlds, virtual worlds, instances of cyberspace, and so on. Moreover, there is a lack of clarity on what all of these terms actually mean and how they relate. We are left with a puzzle: what is the relationship between the concept of *fictional worlds*, familiar from much analytic philosophical theorising about works of fictions, and the theoretical lexicon of new media theory that refers to *virtual worlds*, *simulations*, and *cyberspace*? To put this puzzle in concrete terms, what sort of thing is *Liberty City*, and what is it to



*interact* with such a world in the ways that *Grand Theft Auto IV* makes possible?

## 2 Divergences

In a first step to addressing these issues, I will narrow my focus in this essay to the term *virtual*, as I believe it is this concept that allows for the greatest purchase on the issues; *cyberspace* is intractably metaphorical, and *simulation* too specific in its reference to a kind of use to which we might put virtual items, for either to locate what is really at stake here. *Virtual* has the virtue of being widespread and relatively generic, and also of capturing in its derivation a key part of the issues to be explored here. Given this focus, it is worthwhile making the case for the incompatibility of the classifications *virtual* and *fictional* so we can see why it is that some new media theorists are opposed identifying the two. By thus surveying the apparent divergences, I think we will come to a pretty clear understanding of how the two concepts do in fact relate. Why then, might we be tempted to say that Liberty City is virtual *rather than* fictional?

The principal claimed difference between fictional and virtual items—and it is the difference with which I introduced and motivated this essay—is that virtual items allow for a kind of interactive and participative involvement not seen in traditional forms of fiction such as novels and films. Aarseth is one theorist to positively argue for the distinction between virtual and fictional items on this basis (1997, 2005). He rejects the idea that videogames are merely a species of “interactive fiction” (1997, p. 50), as some previous writers have been tempted to say (Murray, 1998) and claims instead that virtual items are “ontologically different” to fictions (2005, p. 1). Aarseth’s choice of phrase, “ontologically different”, has a hint of metaphysical realism to it, as though virtuality is some distinctive mode of being. He is not alone in making this implication, with a number of theorists claiming that virtual representations are metaphysically significant in some respect (Heim, 1991; Woolley, 1992; Wertheim, 1999; Zhai, 1998). Such metaphysically robust readings of the nature of virtuality almost always lead to mysterious or bizarre views. Benjamin Woolley, to take just one example, draws connections between virtual realities and, amongst other things, the Copenhagen Interpretation of quantum

physics and Baudrillard's anti-realist provocations, to argue that virtual reality stresses our common-sense notion of what is real (1992). In the interest of explanatory parsimony I believe that we should try to avoid the conclusion that new digital technologies have such a metaphysical significance.

In a deflationary and charitable reading, one could take Aarseth's reference to ontology as redundant, and assume the phrase "ontologically different" to simply mean "are different to". But I am not sure that a metaphysically robust interpretation of virtuality can be entirely avoided here, because Aarseth's arguments do seem to depend on drawing the distinction between real, virtual, and fictional items in a way in which these terms are ontologically arrayed. Furthermore, some of the issues discussed later in this essay, on their surface do seem to court a certain amount of metaphysical mystery: what is it, precisely, to *interact with a virtual world*? Indeed, deflating the often-professed metaphysical significance of virtuality is one of the key motivations for assessing its conceptual relationship to fiction, where, though some writers are tempted to take a robust conception of "fictional objects" (Meinong, 1960) such metaphysical issues can be worked through with considerable success (Brock, 2002; Kroon, 2004).

Why does Aarseth think virtual items are different to fictional ones? Virtual items, Aarseth contends, are unlike fictions in that they sustain a level of participative involvement—though he is unwilling to use the term "interaction"—that is not characteristic of fictions. He argues that virtual items "can typically be acted upon in ways that fictional content is *not* acted upon" (2005, p. 1; emphasis in original). Furthermore, "simulations allow us to test their limits, comprehend causalities, establish strategies, and effect changes, in ways clearly denied by fictions, but quite like reality" (2005, p. 2). One of his key examples is of the difference between dragons as represented in literature and as represented in videogames. In *The Hobbit*, Smaug is represented by the way of descriptions (though my copy of the book also includes Tolkien's own ink drawings of the dragon). In the videogame *EverQuest* a dragon is represented by a dynamic graphical model. Aarseth claims that:

One dragon is clearly fictional, but the other is simulated. One is there to read about, or watch on a TV or movie screen, the

other is there to be played with. One is made solely of signs, the other of signs *and* a dynamic model, that will specify its behavior and respond to our input. (Aarseth, 2005, p. 2; emphasis in original)

Hence, that the dragon in *EverQuest* can be “played with” suggests for Aarseth that it has a different ontological status to fictional dragons, which do not allow for such involvement. We might say that Smaug cannot be “played with”, because, as the philosopher Alex Neill has argued in a different connection, there is an “ontological gap” between appreciators and fictional characters: they do not exist in the same world (1994, p. 181). But according to Aarseth no such gap exists in the case of the *EverQuest* dragon, and this implies that it is of a different ontological status to fictions such as Smaug: rather than fictional, it is a virtual dragon.

Similarly, some of the things that in fictions are clearly not real seem to have some kind of real existence in digital artefacts. Aarseth’s own example is of a *maze* or *labyrinth* (2005, p. 3). Mazes exist in any number of works of fiction, where they often constitute a setting for the fictional events. But Aarseth contends that mazes in videogames—where they are common features—are quite different to those in fictions because they support the functions for which mazes are really constructed. The maze featured in the final scenes of Stanley Kubrick’s movie *The Shining* is an evocative setting for a set of narrative events, but not one that has the interactive qualities that real mazes do, or which could be navigated as one might navigate a real maze. But with a videogame representation of a maze, such as in *Pac-Man*, one might actually navigate their way through the maze—and in this case, doing so is a prerequisite of playing the game. The maze in *Pac-Man* thus seems to be something other than a fictional maze, because it is real rather than merely fictional that one might engage with it *as a maze*. There is a connection here to the example of Liberty City, which presents a maze-like structure in the form of its complex highway system and street layout; one of the great challenges in playing the game is learning the way around the city, in much the same way one would learn their way around Liberty City’s real world analogue New York City. Whereas one might *get lost* in a fiction like a novel in a metaphorical way—that is, become *absorbed* in it—one can get lost in Liberty City in an almost literal way, that is, disorientated, or unsure of

where one is. Furthermore, the player can consult the map that is included with the game to discover *where he is* in the city, or *how to get to* some other location. This potential to be genuinely and not only metaphorically lost in a virtual world again implies that the former worlds have an ontological status that the latter do not, and for this reason Aarseth concludes that they are virtual rather than fictional.

Virtual currency is a particularly vivid example of the apparent ontological difference between fictional and virtual items. Currency exists in many videogames, perhaps the best example coming from the massively multiplayer online role-playing game *World of Warcraft*. The fictional currency in *World of Warcraft* is comprised of copper, silver, and gold. Currency is collected by gathering resources and killing monsters and can be spent to obtain new weapons, armour, spells, and other items. Within the gameworld, players can bid their currency in auctions for rare items, or trade directly with other players. But significantly, the currency is also subject to *real* financial transactions, though usually illicit ones, where players exchange gameworld currency for real world currency, usually with third parties who collect the gameworld currency by “gold farming”. Buying farmed gold provides players with a means to avoid the significant investment of time (or *grinding*) needed to progress in the game by legitimate means, by outsourcing that *virtual labour* to companies organized for the sole purpose of the activity, and which are sometimes labelled “sweatshops” (Barboza, 2005). It is even possible in some cases to calculate an exchange rate between gameworld and real world currencies (Castronova, 2005). Fictional financial transactions are preceded outside of videogames, of course—the fictional money in *Monopoly*, or the pretend money used by children in games of make-believe—and people have long traded fictions such as novels and films. But in the *World of Warcraft* the depiction of currency is robust in a way such that it sustains trade in much the same way that real currency might. Surely then, if one can trade virtual money for real money, they must be *ontologically commensurate* in a way that real and fictional items are not, and we should say that the *World of Warcraft* currency is something other than fictional: it is virtual.

A last and related difference that we might think establishes a difference between fictional and virtual items is that virtual items may not involve the imagination in the way fictions typically do. A fiction such as

*The Hobbit* warrants that the reader employs his imagination in filling in the detail of the imaginative scenario presented, as where the descriptions of Smaug found in *The Hobbit* leave much of the dragon's appearance indeterminate:

There he lay, a vast red-golden dragon, fast asleep; a thrumming came from his jaws and nostrils, and wisps of smoke, but his fires were low in slumber. Beneath him under all his limbs and his huge coiled tail, and about him on all sides stretching away across the unseen floors, lay countless piles of precious things, gold wrought and unwrought, gems and jewels, and silver red-stained in the ruddy light.

From this, the reader might imagine Smaug in various ways. Videogames, especially of the graphically rich recent variety, are likely to represent their dragons in a far more detailed and determinate way. In such a videogame a player does not need to imagine what the dragon he is battling looks like, because it is presented in a graphical way in which it can almost literally be inspected from all angles. Indeed, videogames are sometimes criticised for the fact that they preclude an imaginative involvement of the kind that children can have with books. But the point of the argument here is that whereas in literary fictions the involvement in the fictional world is mediated by the imagination, in videogames and other digital artefacts we seem to have a *direct* perceptual engagement with the virtual world, and hence, a transaction with that world that is dissimilar to previous cases of fiction.

In *virtual reality*—a fully immersive and perceptually replete instance of virtual depiction that has been a constant fancy of both science fiction and digital media theory—the lack of an imaginative involvement of appreciators is most striking. Michael Heim notes that in such cases we “[...] do not have to imagine [ourselves] bodily entering a virtual world” (Heim, 1991, p. 135). This putatively *direct* perceptual engagement with virtual reality is sometimes referred to as “telepresence” (Minsky, 1980). Some writers have even gone as far to assert that virtual worlds and the real world might be indistinguishable, and in such cases the apparent *directness* of perceptual access to virtual worlds takes on an ontological tenor. Echoing classic empirical idealism, Philip Zhai suggests that an ontological

distinction between perceptually indistinguishable real and virtual worlds would be redundant (Zhai, 1998).

It is clear enough that there are differences between virtual items and traditional fictions on the basis of which we can distinguish the two classes. But the nature of the distinction is not yet clear. It remains a possibility that the differences are, in respect of the status of virtual items as fictions, merely contingent. That is, rather than being differences that establish that virtual items are never fictions, they may amount to differences that show that when virtual items are fictions, they are of a novel kind. Moreover, some of the above attempts to cash out the distinction are difficult on their own terms, in that in attempting to draw an ontological dissimilarity between fictional and virtual items the idea of virtuality is explained in a potentially mysterious way. It is worthwhile, then, to explore the ways in which virtual and fictional items converge to see if this clarifies the conceptual status of both.

### 3 Convergences

The arguments of the previous section attempt to capture the difference between fiction and virtuality in terms of media and the modes of interaction those media make possible. Putatively, a virtual videogame dragon differs to a literary fictional dragon, because the former, in being represented by a dynamic graphical model, is responsive to the player; whereas the latter, being represented by a set of descriptions and perhaps drawings, is unresponsive. The player can involve themselves with the former dragon, perhaps by battling it, in a way that a reader of a book or viewer of a film cannot. But there are strong continuities between old and new media that warrant the extension of the term *fiction* to at least some instances of the new media, and indeed, the basis of these continuities suggests a proper characterisation of fictionality. Virtual worlds may differ to traditional fictions like novels and films in being representationally replete and interactive, but they share with those previous forms a pragmatic context of depicting situations and objects with an imagined existence only.

Fictions depend for their nature on considerations of representational pragmatics. The usual way to show this is to reflect on the fact that a pair of formally identical representation tokens—portrait paintings say—can

differ in their status *vis-à-vis* fiction. Imagine two portraits, one painted to represent a mythical figure such as Odysseus, but based on the sitting of a model, and the second painted to represent the model himself. Perhaps the sitter for the painting of the mythical figure liked the original painting so much that he requested that the painter repeat the effort to depict him dressed in the mythical garb. In this case, there might exist two indiscernible paintings, one which depicts a fictional person, and the other depicting a real person in fancy dress. It is not the media or representational form that makes one fiction and the other non-fiction, because they share the representational form of portraiture; rather, it is a fact about their intended function that distinguishes them, a fact which surely has to do with what the painter had in mind when producing the artefact: the fictional work is painted to depict a person with an imagined existence only, the non-fictional portrait to depict a person who actually exists. Thus, a pair of formally or perceptually matched items may be fictional and non-fictional depending on their intended functions.

A number of philosophers, though differing on the fine-details, follow this basic understanding of the fiction/non-fiction distinction to characterise fiction as engaging a specific kind of mental attitude. Peter Lamarque calls this attitude a “fictional stance” (Lamarque, 1996). Noël Carroll takes the cognitive attitude to be comprised of “unasserted thoughts” (Carroll, 1990, 1998). For Greg Currie and Kendall Walton the appreciative attitude is seen as “make-believe” (Currie, 1990; Walton, 1978, 1990). All of these positions are similar in seeing as crucial to fiction an ability to hold representational content before the mind, not so that it is believed, but so that it is subject to quite a different type of attitude, with behavioural consequences quite different to those normally associated with beliefs.

Within this body of theory there are also the resources to explain how such imaginative attitudes might utilise or be augmented by linguistic or pictorial representations. Walton argues that in fictions, the imagination is enriched by the intentional design by artists of physical “props” (1990, p.21). Such props, be they the linguistic tokens of a story or the pictorial tokens employed in paintings with fictive subjects, are used to ground and advance games of make-believe. Thus the imagination gives rise to sophisticated fictions like novels, films, and paintings when artists create props to depict fictional situations in an enduring, robust, and detailed way. A painting, for example, builds on a basic imagined scenario by

adding painterly detail so that viewers can get a vivid sense of a person or set of events having no real existence; a film employs moving images to do much the same, but because of its temporal nature, it does so in a manner quite different to static paintings. The differences in these media lead to differences in the way appreciators engage with their imaginary content, and gives each form unique artistic strengths.

On this basis, we can chart how a single kind of fictional item might be represented by different fictive media, seeing the connecting pragmatic thread that runs through different fictive forms. Take the example of a fictional city, any number of which can be drawn from the history of fiction, some of them almost as familiar in name as real cities: Gotham City, Bedrock, Emerald City, Metropolis, South Park, Bikini Bottom, Liberty City, and so on. When it comes to the ways in which these places are represented, there are clear variations in terms of the depictive richness of the representations, and of the modes of interpretation and interaction they support and encourage. In fact, some fictional cities are *transmedial*, and so we can chart a single example to see the scope of these differences.

Minas Tirith is a city that plays a key role in the narrative of *The Lord of the Rings*, principally by being the seat of power of the realm of Gondor and the site of the pivotal Battle of Pelennor Fields. In the original novel, the city is represented through detailed descriptions. This literary form engages the reader's imagination, asking him to fill in the scene, perhaps employing what Walton calls "principles of generation" (1990, pp. 144–69). But because the representational content is fixed by the author at the time of writing, the reader is limited to reconstructing and interpreting the fiction, and the fiction has a "work world" that is common to all readings (Walton, 1990, p. 58). In versions of the book where the city is also illustrated, the imaginative activities of the reader are more strongly guided, potentially excluding some of the imaginative options open to the reader of the non-illustrated novel. Though readers are certainly not cognitively passive when they read the book, they are passive in respect to what is represented by the work itself, and furthermore, they *interact* with the book only in the trivial sense that they perceptually and cognitively engage with it and turn its pages.

*The Lord of the Rings* finds its most famous transmedial instantiation in Peter Jackson's films, which though differing in many details, also depict the city of Minas Tirith. In the filmed *The Return of the King*, the city



is given a graphical depiction that captures in a determinate way the look of the city. In one particularly memorable sequence a camera swoops to follow Gandalf as he rides his horse up the rocky tiers cut into the mountainside on which the city perches, accompanied by the vaulting brasses of the soundtrack. Jackson's treatment of the source fiction was widely praised for how it visualised the fictional world of the books. The film has its own distinctive modes of appreciation, particularly deriving from its presentation of quickly paced spectacle and action sequences, but like the book, the film depicts a fictional work world which is insensitive to audience's involvement. This is because, like the novel, the fictional world is *fixed* in various respects during the production of the film, in this case by the aggregated acts of all those involved in the production of the film, be they writers, actors, art designers, computer graphics artists, directors or editors. Because of this fixity, the principal mode of involvement in both the written and the filmed *The Lord of the Rings* is the *following* of a narrative; the metaphor capturing the *subservient* imaginative role of the audience.

A number of licensed videogames were produced in conjunction with Jackson's *The Lord of the Rings*. One of these, the videogame version of *The Return of the King*, depicts Minas Tirith as a gameplay setting. The Minas Tirith of the videogame is recognisably the city of the film, both being based on the same preparatory artwork. But in the videogame case, the representational prop is comprised of a computer program that contains a complicated mix of functional and representational facets, from 3D polygonal models, virtual cameras, animations, interactive affordances, gameplay mechanics, to shaders and other aspects of the graphics pipeline that determine what is finally rendered on the screen when the game is played. Because of these aspects, the work's representational prop is not rendered or fixed until the player has had his input into the game's software, and so what is ultimately depicted of the fiction really depends on how the game is played.

Though *The Return of the King* is a relatively *linear* game, in defining quite strictly what one must do in a given level—mostly a matter of fighting orcs and other monsters—two playings of the game are overwhelmingly likely to be unique in fictional detail. Other forms of videogame, in particular *sandbox* games such as *Grand Theft Auto* or *Fallout 3*, allow for more extensively divergent playings. The representational form of

videogames thus alters the mode of engagement supported by their fictive works; rather than a passive follower of a narrative, the player is protagonist, directing to some extent the events of the fictional world. Furthermore, the principal mode of interaction the player has with the fictional world is not the interpretation of a narrative—though *The Return of the King* does present much of the narrative of the film, mostly through non-interactive cut-scenes—but the playing of a game. How videogames *situate* their games in fictions is a topic that I have discussed elsewhere (2009, pp. 86–109).

These transmedial instantiations of the fiction of Minas Tirith—in book, movie, and videogame form—though differing in media and modes of interaction, are united by one obvious fact: they depict a place that has no real existence. Thus a fictional city might be represented in any number of ways, and though there are differences in the ways we interpret and interact with the scenario—just how much imagination is involved, what level of interactive involvement the audience has with the scenario, what the ultimate goal of the interaction is—there is a common factor in all the cases: the city has a merely imagined rather than real existence.

It is here that we have the proper response to one of the claimed differences between virtual and fictional items: that the former do not involve the imagination, and that we might have a *direct* engagement with their worlds. This claim depends on smudging a variation in the meaning of *imagination* and its cognates: that is, it relies on a conflation of the *psychological* concept of *imagery*, and the *existential* term *imaginary*. When we say that reading *The Hobbit* is an imaginative experience in the way that playing *Grand Theft Auto* is not, we surely mean that the former calls on our imaginative powers to fill out the fictional scenario with imagistic detail in the way that the latter does not. Similarly, to think that one has a *direct* engagement with Liberty City is a judgment that makes sense only if it refers to the imagistically rich nature of our access to the *Grand Theft Auto IV* fiction. This difference in the extent of the imaginative involvement in the respective fictions is undeniable, but to establish that the fictional worlds of Middle Earth and Liberty City are *ontologically dissimilar*, as the new media theorist might want to say, relies on the existential sense of *imaginary*, on which it is clear that Middle Earth and Liberty City do not differ: both are *merely* imaginary.

That Liberty City counts as imaginary in the existential sense of the term is the basis of the initial plausibility, noted in the beginning of this essay, that videogames are works of fiction. As such, the virtual nature of a work does not have a direct bearing on the ontological status of its depicted world, which is a fact of representational pragmatics that is untouched by considerations of media. Rather, the virtuality of a work or medium derives from the kind of psychological or interactive involvement that it makes possible.

## 4 Virtual Media

If *virtual* is a term that properly applies to media, how is the term best cashed out? To answer this question it might be best to perform a simple conceptual analysis of the term. Indeed, looking at the usage of *virtual* shows it to have a number of senses, the conflation of which probably abets some of the confusion about the nature of videogames and other digital media (Tavinor, 2009, pp. 48–51). One suspects that the term *virtual* is often used merely as a generic reference to new computer technologies. Perhaps we would be better to use the word *digital* for this generic sense, given its established use in referring to digital watches, cameras, and other technological artefacts.

But the concept of *virtuality* also has a more specific use that conflates many the issues discussed in this essay. Arguably, in some new media theories, *virtual* is extended in a way that it comes to refer to issues properly understood under the rubric of *fiction*: *virtual world* is sometimes most credibly understood as a near synonym of *fictional world*. James Newman's discussion of virtuality, for example, juxtaposes the reality of real spaces and the unreal nature of virtual spaces without ever referring to the fictionality of the latter (2004, pp. 107–11). In this sense, a virtual space is one that is imaginary and does not exist. But this leads to a confusion: what is a non-existent space that it can be *traversed*, as virtual spaces in Liberty City apparently can be?

*Virtual*, in a primary sense, means *being such practically or in effect*, and it is this sense which gives us the guidance needed in this discussion. A virtual disaster is an event which is for all practical purposes a disaster. In computer technology, the term entered the lexicon through

reference to *virtual computers* and *virtual memory*. A *virtual computer* involves running a computer program on pen and paper rather than on actual hardware. The virtual computer can be treated as an actual computer for functional purposes, because the pen and paper process mimics the same rules as a hardware running of the program (indeed, if the term *computer* is functionally defined, then the pen and paper computer *is* a real computer). *Virtual memory* gives the impression of being continuous memory, but instead is merely apt to be treated as such for practical purposes because it replicates the functional features of continuous memory. Similarly, a virtual internet shop has all the functions of a traditional shop, without the traditional brick and mortar outlet. *Virtual*, on all of these occasions, is framed to capture the fact that an item might stand as a *functional proxy* for a target—the target being whatever is represented or replicated—allowing for the same kind of functional engagement as the target, but in a non-literal merely *as-if* sense.

Somewhere along the way, virtuality began to refer to digital media representations, and this is now one of its key uses. Following on from the sense of virtuality defined above, a virtual representation is one capable of reproducing structural aspects of its target, so that it is apt to be treated as the target. More strongly, we might say that genuinely virtual representations are *structural isomorphs* of their targets. Such virtual items can be interacted with to produce changes in the representation that correspond to interactions and effects that are possible with the target. And so a virtual representation such as a 3D topographical model of a city can be treated, for some purposes at least, like an actual city because it is structurally isomorphic in various respects with the actual city (the geography, street addresses, and so on share a structure with the actual city in terms of its spatial layout). In epistemological terms, the virtual city is as good (or even better) than an actual city for learning the city's layout because one can explore it in the way one would explore the actual city. Google Maps' *Street View* has this very function in that it presents in a virtual way actual locations such as New York City so that these locations can be virtually explored (by clicking a pointer with a mouse) as one might explore the city itself. That is, it can be treated as an *interactive proxy* of what it represents.

The potential of digital representations such as 3D graphics to depict such structural features explains the naturalness of employing the term

*virtual* in a digital context. But it should be noted that under this definition of *virtual*, such representations can be found in non-digital media. I think we should say that many things, like pen and paper role-playing games, maps, and even toys, involve virtual representations, because they do reproduce the structural properties of their targets in a way that allows for virtual interaction: a sandpit dump truck, for example, might have an articulated tray that can be used to tip out sand.

To make sure this definition of virtuality is meaningful it is worthwhile considering non-virtual representation. The descriptive sentences found in novels are good examples of non-virtual representations. The comparison with the map example just introduced nicely picks out the difference. The sentence, “Lincoln is twenty kilometres southwest of Christchurch”, depends for its meaning on reference and syntax, but the structure of the syntactic combination in this case is not a candidate for being “practically treated as” the fact that is represented. Hence, sentences of this kind do not constitute virtual representations. A map showing the relationship of Lincoln to Christchurch, however, bears a structural relationship that does *map onto* the facts it represents. And, in this case, one might use the map to stand in as a functional proxy the facts it represents, for example by placing it on the ground aligned with north, so that, assuming one is standing in Christchurch, one could discern the rough direction of Lincoln; or by tracing their finger along a street to find a best path to a desired destination. One could not have such an engagement with the sentence, “Lincoln is twenty kilometres southwest of Christchurch”, because here the geographical information is *channelled* through a combinatorial syntax and that bears only a highly conventional relationship to the fact represented.

*Virtual cameras* are a crucial aspect of much digital media, and again allow us to see what is meant by the notion of virtuality in respect to media. Really, videogames do not involve cameras, rather *virtual camera* is an idiom employed by games designers to describe certain aspects of three-dimensional computer representations (Kerlow, 2000, pp. 88–91). In particular, the virtual camera is crucial in opening up the possibility of three-dimensional spaces, and allowing virtual movement through those spaces. Alongside polygonal 3D objects, the virtual camera is one of the key developments in virtual representation, and illustrates the definition of virtuality given here in that the structures it employs are the algorithm-

mic transformations of various vector functions of a 3D graphical model. This software function is apt to be treated as a camera, because these algorithmic geometrical manipulations, and their subsequent display on a 2D screen, can be made to match quite closely the changes that would occur if an actual camera was used to film a real scene. In *Grand Theft Auto IV*, one of the functions of the virtual camera is to represent the point of view of the player's fictional proxy, but it is also given virtual cinematic uses in the cut-scenes of the game; and in such cases even artefacts of the filming process are sometimes simulated to achieve cinematographic effects such as depth of focus variations or lens glare.

Virtual representations can depict both real and fictional things, however. The already mentioned *Street View* is perhaps the best example of a virtual representation used to depict real things, in this case being an interactive topographical representation of real locations. Another good example are the unmanned combat and surveillance aircraft that are controlled as one would control a flight simulator; in such cases the controls often consist of a *virtual cockpit* and graphical view of the environment in which the plane is *really* flying. Clearly such a representation depicts things with a real existence, though in a virtual way, because of the very real and deadly consequences of manipulating the virtual controls so as to attack a target. But consider the close resemblance to combat flight videogames where a player can have a near identical interactive engagement, the difference being that the "targets" in the videogame have no existence in reality. Indeed, the videogame *Call of Duty: Modern Warfare* has an interlude in which the player acts as a gunner in an AC-130H Spectre gunship, picking out targets using night vision in a sequence unnervingly reminiscent of real combat footage.

Properly delineated, *virtual* and *fictional* refer to different things: the former to a feature of some media, that is, their structural and interactive isomorphism with their targets, and the latter to a feature of pragmatic uses of media, namely, the depiction of things that are only imagined to exist. Specified in this way, virtuality and fiction are contingently and partially overlapping rather than opposed categories, sometimes coinciding—in the case of many videogames—but often not. Videogames are thus often *virtual fictions*, and this reformulation of the concepts of *fiction* and *virtuality* as having a contingently overlapping rather than oppositional relationship accounts for the media differences between videogames and other fictions,

but does not lead us to reject to very strong intuitions that videogames—with their opportunities for killing enemies, and battling dragons—are fictional.

## 5 Interactive Fiction

Returning to the initial example of Liberty City, it is clear enough that when virtuality and fiction do coincide, they have the potential to produce distinctive artefacts in the form of *virtual fictional worlds*. Their representational nature makes virtual items ideal for depicting fictional worlds in which one can play a game or toy with the details, that is, *interactive fictions*.

There are a number of sources of doubt, however, about the utility of *interactivity* as an accurate or even theoretically meaningful term; these worries place constraints on any theory of fictive interaction. Dominic Lopes notes that because of its sheer ubiquity as a technological buzzword the term *interactive* is prone to abuse, and is of limited theoretical use without specifying a substantive meaning (2001, p. 67). Furthermore, in maintaining that virtual fictions are interactive, one might seem to claim, problematically, that traditional media are “passive” in some respect. But as I have already noted here, all art is interactive insofar as it involves appreciators in a physical and cognitive engagement with a work (Carroll, 1998, pp. 30–49; Walton, 1990, pp. 144–69). The interactivity of interactive fictions must be more substantive than mere physical, perceptual, and interpretative involvement, therefore. Finally, and strikingly, if one claims that the interactivity of a fiction is comprised of the apparent ability of a player to perform fictional activities, like “driving cars”, or “battling with dragons”, then they will be faced with the immediate difficulty that players do not really do these things. But though this most obvious way in which videogames might appear to be interactive fictions is illusory, the explication of *interactivity* ultimately settled on here must explain how videogames do give *the appearance* of the player having such a role in a fictional world.

I think that the needed sense can be worked towards by introducing Lopes’ theory of computer art (2001, 2009). Lopes argues that a number of recent artworks, exploiting the representational potential of computers,

allow appreciators modes of interactive engagement that “no other artistic media can enjoy” (2004, p. 112). Lopes’ theory is developed to address computer artworks, but it has clear application in the case of videogames because his paradigm examples of “strong interactivity” are traditional games such as chess. Using the example of games to distinguish “strongly interactive” digital works and “weakly interactive” ones, he argues that,

Games are “strongly interactive” because their users’ inputs help determine the subsequent state of play. Whereas in weakly interactive media the user’s input determines which structure is accessed or the sequence in which it is accessed, in strongly interactive media we may say that the structure itself is shaped in part by the interactor’s choices. Thus strongly interactive works are those whose structural properties are partly determined by the interactor’s actions. (Lopes, 2001, p. 68)

Lopes concludes that much of what is referred to in the digital realm as *interactive* is only weakly interactive because the interactivity amounts to the audience having some say in the order a predefined structure is accessed. The interactivity of some digital artworks is more substantive however, because it involves the appreciator having a discernable impact on the structure of the work itself. In the case of digital art, the structures in question are those that are behind “whatever intrinsic or representational properties it has the apprehension of which are necessary for aesthetic engagement with it” (2001, p. 68).

Videogames do not merely involve choosing the order in which the representational structures of the game are experienced, but involve the player partly determining the content of a game’s instance, and so it is clear enough that videogames do count as “strongly interactive” in Lopes’ sense. In *Fallout 3*, unless I decide to attack the slaver community of Paradise Falls, no such events are depicted in the work. And just how my attack progresses is determined by the plan I put into effect, and the contingencies of how it plays out on the occasion; the resulting fictive events are likely to be utterly unique to my playing. *Fallout 3* as a work plays a crucial role in determining the *type* of events that occur in a playing (and some world events may be common to all playings) but playing the game is more than simply navigating a predefined fictive structure; the computational prop behind the game is has a high degree of representational



contingency, and this is one of the key reasons why the game is held in the high esteem that it is. Thus, videogames count as interactive fictions in this theoretically robust sense.

This sense of fictive interactivity might be augmented by overlaying it with Walton's prop-based theory of fiction (Tavinor, 2009). Fictive interactivity could thus be characterised as the situation in which the content of the fiction is determined in part through the player's interaction in a participative "game world" (Walton, 1990, p. 58). Videogaming props exist in a state of representational limbo as computer programs with the potential to depict a range of playings, and from these a gamer may generate a world through their playing. This is a matter the game's fictive prop—the software and hardware—depicting a fictional world to incorporate the fictive actions of the player which are fed into the prop through the controls of the game. But unlike most traditional fictions, there seems to be no single "work world" shared between instances of the game because the player's activities *project* into the fictional world that is depicted by the fictive prop. As Stephen Davies notes, in interactive fiction, "actions authorized for and performed by the player can affect events in the work-world, not only in the gameworld. In other words, the player's actions can affect the content of the fiction" (Davies, 2009).

To say that one *interacts with a fiction* also implies that videogamers have some sort of physical engagement with the fictional objects in gameworlds, so that they might "drive a car to Broker" or "battle a dragon". Players genuinely interact with gamepads and pieces of computer hardware and software, but in what respect can they be said to interact with virtual items like guns, cars, and dragons? Indeed, we *are* naturally tempted to describe our activities in gameworlds in such paradoxical terms, even though we are quite aware that no such objects and events really exist. These observations make it clear that there are two related issues of interactivity at stake here: first, the nature of interactive media, where we might question what it is that makes an artefact like a videogame an interactive work; and second, issues that arise with interactive fictions specifically, where we can question what it is to interact *with* fictional items like dragons and mazes. Furthermore, it is in this context that we can best address the metaphysically provocative reading of virtual worlds that prompted this paper, because it is this apparent interaction with fictional things in the case of videogames and other digital artefacts—the ability to

“play with” a virtual dragon—that prompts Aarseth and other new media theorists to conclude that videogames are something other than fiction.

The solution to this specific problem of fictive interaction can also be derived from a combination of Walton’s theory of fiction and the notion of virtual props. In claiming that appreciating fictions involves the *nesting* of a “work world” within a participative “game world” fiction, Walton is led to the conclusion that many of our statements about our apparent interactions with fictions are to be properly regarded as fictional (1990, pp. 59, 215). If a movie viewer is tempted to avow his fear of a horror movie blob of slime, this apparent *intentional relationship* is rightly characterised as “fictional fear” (1978). This same fictional proviso is also appropriate for the apparent *causal relationships* seen in videogames. Players do not really interact with guns, and cars, and dragons, rather they interact with a virtual fictive prop, an act which *makes fictional* of a game of make-believe, interactions with guns, cars, and dragons. Though the following is a bit of a simplification, the virtual fictive prop that represents a dragon in a videogame fiction, and which is the real object of interaction, is an animated figure on a screen, and the interaction is achieved because the figure is transformable through graphical algorithms that are mapped onto the functions of the controller that the player manipulates (Tavinor, 2009, pp. 61–85). The player can “strike the dragon with his sword”, because the virtual prop—a 3D graphical model affixed with “affordances” (Tavinor, 2009, p. 80)—coordinates the visual display to give this impression. A key part of this explanation of apparent fictional interaction is the role that the “player-character” plays in providing gamers with a fictional gameworld *proxy*, and one about whom they can make fictional attributions. It might be noted that the term “player-character” itself bridges what is in reality an *idiomatic gap* between fictional and non-fictional language (Tavinor, 2009, p. 70).

Though drawing on Walton’s theory, a difference of my explanation of virtual interaction is that where Walton’s main concern is intentional states such as emotions, videogames and other digital artefacts make possible physical or causal interactions because of the virtual nature of their props. Indeed, this understanding may further bolster Walton’s claim that many of our apparent interactions with fictions are properly regarded as fictional, which as most will be aware is a claim that has encountered a great deal of resistance (Carroll, 1990; Lamarque, 1996). Because of the

tangibility of causal interactions it is clearly the case that battling a dragon in a videogame is a fictional activity; but then it should be clear that to say that “I was terrified of the dragon as I battled it”, is equally fictional. The metaphysical mystery of fictional interaction—how one can interact with a non-existent world—is thus solved, not by bridging a supposed “ontological gap” between appreciators and fictional worlds, but stripping fictional worlds of their ontological significance, and understanding that the interaction of players with gameworld objects like dragons and cars is *fictional interaction*.

To be a fully satisfactory account of the apparent ontological problems that prompted this essay, this explanation of fictional interaction owes us an account of what seemed to be the most mysterious cases of virtual interaction: the exploring of mazes in videogames like *Pac-Man*, and the trading of fictional for real money in the case of *World of Warcraft*. In the limited space here I will address the latter issue, though I hope the reader will see how, with suitable amendments, the theory of *virtual fictions* is apt to explain the case of mazes and labyrinths in videogames.

Surely, the new media theorist will say, the fact that purportedly fictional gold—gathered by killing monsters and collecting treasure—can be traded for real currency, implies that the two currencies are ontologically commensurate, and so the gameworld currency must be something other than fictional. The explanation of virtual money draws on most of the conceptual resources I have developed in this essay, and it has two key parts: first, carefully delineating what is fictional and what is real of *World of Warcraft* and noting the crucial role of real props in generating what is fictional of the game’s world; and second, emphasising the richly articulated nature of the virtual props found in this videogame.

The *gold* in *World of Warcraft* is fictional. It is clearly fictional, because one collects gold by killing monsters and gathering resources like herbs, and one spends it on swords and armour; these things only exist in the fictional world of the game. But these activities and objects can be fictional of the game because the virtual props of *World of Warcraft*—which are real—are articulated in a way to support this fictional resource gathering and trade. This is because the virtual prop grounding the fictional trade of the *World of Warcraft*, in line with my earlier definition of virtuality, replicates the structure of its target, in this case, the conditions that are crucial to real currency: *discreteness*, *divisibility*, and *transferability*.

These properties are the basis of the structural isomorphism between the virtual fictive props in *World of Warcraft* and genuine currency. Ultimately bearing these properties are pieces of computer code and their graphical representations that the game's designers have very carefully constructed for the very purposes of sustaining fictional trade, which is a key part of the gameplay in *World of Warcraft*. But because the fictive prop at the basis of *World of Warcraft* has these features of discreteness, divisibility, and transferability, it can also support real trade. Thus a prop—in this case a functionally discrete piece of computer code—*makes fictional* the existence of a piece of gold in the world of the game, one that can be fictionally spent to acquire items like swords and armour, but which in reality can be traded for real-world currency. Whereas with previous forms of fiction, one might buy or sell the fiction as a whole—by purchasing a novel, or a DVD—and in a way that had no impact on the participative fictive game one played with the work, in modern interactive fiction one can buy and sell discrete parts of the fiction in a way that itself makes things fictional of that game. Indeed with some videogames and virtual worlds, such as Sony Playstation's *Home*, the business model depends on this real trade of fictive props.

From this we can construct a general theory of how to resolve the apparent metaphysical realism seen in some discussions of virtual worlds, and hence, deflate the apparent divergence in virtual and fictional worlds with which I motivated this essay. First, *fictional* and *virtual* need to be sufficiently demarcated: fictions, being defined in terms of the pragmatics of representation, are those things that represent *existentially imaginary* events and worlds; *virtuality* is defined in terms of the structural qualities of certain media, namely, the interactive isomorphism of the representation and its target. Defined in this way, fiction and virtuality are contingently and partially overlapping categories. Second, it should be noted that the distinctive nature of virtual media supports fictive representations that allow players to have an impact on what is represented by the fiction: they are often *interactive fictions*. Finally, the apparent paradox of our subsequent descriptions of interactive fictions—where we might say that we “drove a car to Broker”, “battled with a dragon”, and so on—is avoided by the understanding that these things are fictional rather than true. Players do not really interact with the worlds of videogames, it is only fictional they do so.

# Fiction, Indispensability and Truths

MANUEL REBUSCHI & MARION RENAULD

## Introduction

The indispensability of mathematics thesis, associated with the names of Quine and Putnam and implicitly involving Quine's ontological commitment criterion, leads us to accommodate mathematical entities within our ontology. The combination of this criterion and that thesis seems to have been a prevalent paradigm in analytic philosophy, at least as far as ontological issues are concerned. In this context, the specific case of fictional narratives may seem easy to deal with: because they are dispensable, those narratives cannot properly be said to be true (an idea already found in Frege), hence do not commit us to anything. The purpose of the paper, however, is to argue that the combination of the thesis and the criterion does not apply to fiction in such a straightforward way. The task of encompassing fictional narratives requires two moments. On the one hand, we must admit a kind of indispensability, therefore a kind of truth for fiction, while on the other hand, we must deny all kind of ontological commitment towards fictional entities.

The argument we present in this paper is not *ad hoc*, that is, not tailor-made for fiction. It is grounded in a critical discussion of Quine's combination, to the effect that its context of application is excessively narrow. Following Frege, Quine indeed restricts his logical analysis to a very spe-

cial use of language: that of scientific or theorized knowledge of the world. Moreover, both authors agree that natural language (hereafter NL) is not appropriate as a medium for expressing our theories, and that we should either go for a *Begriffsschrift* freed from all the possible ambiguities of vernacular languages (Frege), or else translate (or “regiment”) our theoretical sentences into a *canonical language* like that of first-order logic (Quine).

Quite obviously, the case of theoretical or scientific sentences does not exhaust all the possible uses of language, and more generally, these uses are neither confined to expressing our knowledge, nor to describing the world. We are involved in the *practical* game of finding multiple *interpretations* of the world and of our fellow creatures, and these interpretations do not seem reducible to more or less naturalized descriptions. Here we assume a traditional distinction between explaining and understanding, that is, between the positive methods of natural sciences and the hermeneutic methods of moral sciences (see Descombes, 1995, p. 51). Quine himself endorses a distinction along these lines with his idea of a “double standard” (Quine, 1960, §45), of a sharp separation between what can be known (theorized, naturalized, etc.) and intentional issues (issues that are practically manageable, but from which no knowledge can be gained).

Indispensability—a good path toward truth—need not be restricted to the only descriptive, epistemic and theoretical uses of language. Our goal is to extend it to any kind of discourse that may be involved in *understanding* the world, and in relating to the world in a practical way (if we extend the point to non-linguistic know-how). We propose to characterize those discourses that are indispensable to our apprehension of the world as being true, and this, we think, includes discourses of a non-theoretical kind. To avoid the threat of relativism—defining truth in terms of utility or usefulness not only in our descriptions, but also in our interpretations—we draw inspiration from Dennett (1987), whose purpose is to build bridges between interpretation and description over the double standard gap: although it remains on the “dark side” of knowledge, intentionality takes part in our theoretical explanations, and does this on the basis of a usefulness criterion.

On the other hand, in everyday life all of us take part in linguistic *practices* where some of the assumptions of formal, canonical or ideal languages are purely and simply absent. The strong connection drawn be-

tween ontological commitment and the notion of a bound variable seems to depend on the somewhat ethereal atmosphere of first-order logical languages. Strawson's (1950) well-known criticism of Russell's theory of descriptions (1905) remarkably demonstrates the importance of pragmatic features in linguistic understanding, features which, at that time, were absent from formal languages. NL semanticists have, since then, included numerous phenomena related to the use of language in context. The resulting picture of language as we speak it is much richer than the picture Quine was relying on as he was shaping his criterion for ontological commitment. We propose to embrace the richness and to abandon the criterion in question.

Our strategy will be as follows. Two aspects of Quine's indispensability argument, which turn out to be mutually independent, need to be disentangled. On the one hand, we have an entailment *from indispensability to truth*: if a certain (theoretical) discourse—like mathematics—is indispensable for our holistic explanation of the world, the discourse is to be acknowledged as true. On the other hand, we have an implication *between truth and ontology*: this is Quine's criterion of ontological commitment. We claim that the connection between indispensability and truth must be generalized, and that the link between truth and ontology must be removed. The resulting position on fiction will be that we may well speak of truth when it comes to fiction, yet this yields no particular ontological commitment toward fictional entities.

The argument of the paper comes in two steps. In the first part, we argue negatively against the universality of Quine's criterion, that is, we argue for an ontologically neutral interpretation of NL quantification (and more generally, of NL semantics). This part of the paper does not bear specifically on fictional discourse since it is intended to account for the way language works in general. The idea is that at the level of interpretation, one can handle quantifiers, proper names, and even truth, without any commitment regarding the things in the outside world. In the second part of the paper, we argue positively for the indispensability, hence truth in a strong sense, of fictional discourse. The purpose, then, will be to shed light on the kind of truth that non-descriptive uses of language may possess.

## PART I - TRUTH WITHOUT ONTOLOGICAL COMMITMENT

Quine's ontological commitment criterion corresponds to a partial view of language and as such cannot be applied to any case in general, especially to the case of fictional discourse—or so shall we argue in this first part. The criterion does seem appropriate when dealing with very idealized situations where context does not play any role and where utterances can be interpreted using only semantic (versus pragmatic) resources. In other words, as long as context stays out, Quine's criterion might be in—and vice versa.

As part of the background, we stress a contrast between two major approaches to semantics: (i) a *representational* approach, according to which semantic theories are meant to account for the relations between language and reality, and (ii) a *procedural* approach, according to which semantics should provide explanations of the speakers' understanding of the meaning of sentences and discourses. As we shall see, Quine's approach is of the first kind, whereas natural language semanticists are often tempted to turn to approaches of the second kind. This “procedural turn” is expected to yield important consequences concerning the purported connection between semantics and ontology.

### 1 The Universalistic Tradition

The representational view of semantics clearly dominates the origins of logical analysis. Frege's approach to semantics and Russell's view are similar in that on both conceptions, the semantic correlates of (at least singular) expressions automatically come with an ontological load. And according to many of their followers and commentators, after accurate regimentation has been made, not only singular terms, but *every* linguistic expression triggers an ontological commitment. On this received view, Frege and Russell are thought to have admitted semantic correlates for general terms (concepts and senses of concepts, or Universals) and for sentences (truth-values and thoughts, or Propositions), as being so ontologically relevant as singular terms. Consequently, ontology simply is semantically theorized language (and reciprocally). When the semantics is dualistic, like in Frege, the world splits into (at least) two parts, one



intensional, the other extensional);<sup>1</sup> when the semantics is monistic, like in Russell, a unique world suffices.

This “ontologizing” understanding of semantics, which seems to have been shared by many after Frege and Russell, rests on strong meta-theoretical choices regarding language analysis, for instance the particularly intricate treatment of empty singular terms found in Frege’s work, who wanted to avoid all commitment to fictional entities. Though it is more concerned with epistemic features, Russell’s drastic reduction of the category of genuine singular terms is in the same vein, and so are contemporary “neo-Russellian” criticisms of the legitimacy of individual concepts.

At odds with Frege’ and Russell’s views, Quine explicitly restricts ontological commitment to *one* syntactic category, namely *variables*. This is his well-known criterion: “To be is to be the value of a variable” (Quine, 1948/1961). Quine’s strategy makes it possible to use predicates and to make statements without thereby being committed toward such platonic objects as Universals or Propositions. In this respect Quine’s account may seem to be somewhat similar to nominalism, although it still is full-fledged realism about particulars. Thanks to Quine’s accordance with universalism,<sup>2</sup> his syntactically delimited ontology is still conceived of as a genuine ontology: because there is *one* language—first-order logic being the canonical regimentation language of science—and because there is *one* world—cf. Quine’s lack of interest in model-theory or his rejection of modal logic—, language always relates to reality. For this reason, restricting ontological commitment to variables does not imply departure from the classical view of semantics as representational.

After the so-called “linguistic turn” in philosophy, Quine’s conception of ontology and (extensional) semantics became the most prominent framework for different families of metaphysical realists, who consider some or all semantic items as *bona fide*, that is, ontologically relevant ob-

---

<sup>1</sup> Each part may, in turn, split even further according to syntactic categories (singular terms, general terms, sentences). Also, Frege’s ontology adds a further world corresponding to the connotative or “subjective” aspect of language (though one that is not accounted for by the semantics).

<sup>2</sup> See van Heijenoort (1967) and Hintikka (1988) for the distinction between two traditions in logic: the *universalistic tradition* and the *model-theoretic tradition*, corresponding to two rival views of language: *language-as-medium* vs. *language-as-calculus*, respectively.

jects.<sup>3</sup> What is more, technical issues within logical semantics, such as the status of individual concepts, are generally challenged while still assuming Quine's criterion. It should nonetheless be noted that Quine never really claimed to account for linguistic meaning, a notion he thought a bit suspect, and that his proposed regimentation explicitly departs from sense preservation or synonymy (Quine, 1960, §33).

Quine's support to the universalistic ideas actually involves a strong connection between logic and ontology, whereas the subsequent meta-theoretical constraints seem to have forced him to depart from the project of a complete explanation of natural language semantics (and from a considerable part of the technical developments in logic). The universalistic conception of semantics thus appears to be *strongly normative*, since *what there (presumably) is* plays a coercitive role throughout language analysis, and leads to revisionist strategies. To put it in a nutshell, Quine's conception implies that ontological considerations come first and that semantic theories (in particular, procedural semantics) should either adapt or die.

## 2 The Model-Theoretic Turn

Tarski's work opened new perspectives in and *on* semantics. His systematic meta-theoretical approach to semantics allows, by itself, for the possibility of language-shifts; going even further, model theory allows for world-shifts, that is, for language reinterpretations. Extending Quine's criterion to the resulting model-theoretic frame amounts to identifying ontology with the domain of an interpretation structure for a formalized language. This identification may be labeled *Model Realism* (hereafter, MR). Though a heir of universalism, MR seems to have been the dominant position since the semantic turn initiated by Tarski. According to this view, semantics should consist of (or depend on) representational theories.

Let us call that language- and semantics-relative "ontology"—i.e. the constituents of the domain(s) of a structure—a *Technical Semantic Ontology* (hereafter, TSO). Does it make sense to think of TSO as genuine

---

<sup>3</sup> Many contemporary logical accounts of the semantics of fiction do not cut off from this general conception. See Gochet (2010) for a brief overview.

*ontology*, that is to say, as *what there is*?<sup>4</sup> Intuitively, if the thought is to make sense, we need to choose *a single* language and *a single* semantic interpretation of that language as being the “best” ones.<sup>5</sup> In other words, TSO seems to be genuine ontology ... if we choose to go back to universalism, though.

At first glance, Tarski-type semantics could very well explain the genuine relation that holds between (some) language(s) and (some) world(s). It seems as if it could explicate truth in terms of a correspondence between language and world (regimented in the form of an interpretation structure). Thus, it could—and actually did—reinforce the “ontologizing” conception of semantics associated with Quine’s criterion.

On the other hand, Tarski-style semantics allows for natural and direct solutions to certain problems that arose from within Fregean and Russellian semantics, like the problem of empty singular terms. For instance, providing a truth-functional semantics for fictional statements is as simple and easy as quantifying over a domain of fictional objects; whether we prefer a direct-referential or a descriptivist theory, the very same domain can provide (fictional) “referents” for (fictional) proper names conceived of as rigid designators. Does this kind of tolerance imply an ontological commitment to fictional entities? If Quine’s criterion is to carry over to model theory, then TSO (here “fictional objects”) will make for a genuine ontology. As a result, semanticists should either assume a very liberal and luxuriant ontology—taking fictional entities at face value—, or else try to avoid this commitment by means of a theoretical revision. Likewise for possible-worlds semantics, depending on whether or not we are willing to be committed to “possible objects”.

So, ontology-oriented or representational semantics too appears to impose very strong constraints, embodied in the meta-theoretical choice of splitting usual singular terms into two categories according to the appropriate ontology.<sup>6</sup> Despite (possible) appearances, a descriptivist account

---

<sup>4</sup> In Carnap’s words (1950), the issue is to know whether *internal questions* of existence should be identified with *external questions*.

<sup>5</sup> In a way similar to Quine’s avoidance of semantic variability: he roughly rules out possible (non-actual) worlds in favour of the holistic language of science.

<sup>6</sup> Where the boundary between these two categories is to be drawn is a matter that may differ from philosopher to philosopher. An extreme case is that of those advocates of genuine “non-existent objects” who reduce the class of genuinely vacuous singular terms to the empty class.

of usual proper names *à la* Russell is no exception to the rule: if a definite description is vacuous and occupies a subject position, the whole sentence is systematically false, which entails lots of undesirable consequences.

This way of complicating semantic theories already is a questionable matter. We also think it is illegitimate. A strong argument for a “non-committal” semantics—hence against MR—lies in the *theoretical uniformity* it allows.<sup>7</sup> Taken in isolation, this reason is not sufficient, but can further be warranted on methodological grounds, as follows.

In his criticisms of Donnellan’s seminal paper on descriptive and referential uses of descriptions, Kripke (1977) argues for “unitary theories” against “theories that postulate [semantic] ambiguity”, if the alleged ambiguity is not expected “to be disambiguated by separate and unrelated words in some other language”; so, against Donnellan’s (1966) conception of semantic values that shift with uses of descriptions, Kripke speaks for an account “on pragmatic grounds, encapsulated in the distinction between speaker’s reference and semantic reference” (*op. cit.*, §3c).

By dividing usual singular terms into two classes according to some (presumed) ontology or other, Model Realists postulate a kind of (superfluous) linguistic ambiguity: the semantic values of linguistic expressions are supposed to vary as we go from one *presumed* ontology to another. For example, the same usual proper name “Santa Claus” will receive different *semantic* interpretations when used by a credulous child and when used by his skeptical parents.<sup>8</sup>

To the extent that semantic theories are concerned with *linguistic meaning*, it is quite amazing to think that values may shift according to shifts in (ontological) *beliefs*.<sup>9</sup> In fact, the distinction between the two uses cannot be adequately captured by Tarski-type semantics. What happens, for instance, when the parents and their child talk together about “Santa Claus”?

---

<sup>7</sup> As a proponent of a maximally liberal ontology, Zalta rightly insists on this point: “Once we are able to see that all significant proper names are names of objects, we may simplify the Tarski-style definition of truth for languages in which names of nonexistents appear along with names of existents. The truth conditions may be specified more systematically, since no special precautions need to be taken to distinguish the two kinds of names” (Zalta, 2003, p. 2). (In turn, his support to MR forces him to account for a general “metaphysics of objects”, including nonexistents.)

<sup>8</sup> Moreover, a couterintuitive corollary of MR is that an historical discovery—say, confirming the conjecture that Homer never existed—may yield a *semantic* change.

<sup>9</sup> This contradicts our pre-theoretical notion of linguistic meaning; cf. Evans (1982, pp. 23–4).

How can non-believers *refer* to a non-existent? Of course, they cannot: they can only *pretend* to refer. But if they want to teach the myth as a true story they can and even need to share the same object of belief as their child, namely the semantic value he or she assigns to the mythical name. In other words, whenever one speaks veraciously of “Santa Claus”, be it in a fictional or in a serious way, one has to posit a genuine *semantic* correlate for that name. Fictional or serious uses of singular terms and their related speech acts, pretended or genuine reference, need not and cannot enter the semantics. As Recanati (1996) puts it, “the difference does not matter from a strictly linguistic point of view. For we use the same linguistic material, with the same linguistic meaning, whether we genuinely refer or only pretend to refer. This follows from the very notion of ‘pretense’” (p. 467).

Incidentally, our criticism of the standard processing of empty proper names extends to influential accounts of the semantics of fiction:<sup>10</sup> free logics (Lambert, 2003) as well as substitutional accounts of quantification (Kripke, 1976) rely on the *a priori* and *ad hoc* definition of a class of specific proper names—constants to which no existential generalization applies in the case of free logics, substitutional terms in the case of substitutional quantification—, thus losing uniformity in their proposed semantic analysis of proper names.

### 3 Beyond (First-Order) Formalized Languages

Where does Model Realism come from? It seems to follow directly from a strict application of Quine’s criterion to model theory: if a sentence consisting in an existentially quantified formula is true in a given structure then the value of the existentially quantified variable, which contributes to making the sentence true, must be reckoned as part of the ontology. However, such an application lacks justification. Should we take it for granted that quantifiers exhibit the same behavior whenever model theory applies? It appears not to be the case. Like Frege, Quine sticks to a very restrictive class of language use—viz. genuine descriptive sentences—, and his criterion depends on the regimentation of such sentences within

---

<sup>10</sup> See Fontaine & Rahman (2010) for an overview and a recent account of the semantics of fictional statements.

an adequate logical language.<sup>11</sup> But of course, many language-games do not square with the one favoured by Quine.

### **3.1 Taking Context into Account**

Strawson (1950) challenged Russell's (1905) standard analysis of descriptions. Definite descriptions, in their normal use, are usually incomplete, like "the table", and despite this, we do succeed in referring to individual objects. However, there is more than one table in the universe. So, on the Russellian construal, the description "the table" should miss its target. It is incomplete in that in order to successfully denote an individual, it requires completion by means of an indexical, as in "the table in this room", or by other means, as in "the table in Russell's kitchen".

Different strategies have been followed to avoid the difficulty (see Reimer, 1998). One of them, favoured by lots of semanticists, is to say that the interpretation of an incomplete description is made possible by considering the description in its context of use: the context plays a role in restricting the domain of relevant objects. For instance, interpreting "the table", as used in a room, is made possible because the context of use implicitly restricts the course of values to the objects in that room. In other contexts, for example while talking about Russell's kitchen, the domain can be restricted by the discourse preceding the use of the description, in which case the description in this use is naturally interpreted as referring to the table in Russell's kitchen.

As a matter of fact, an analysis along these lines works not only for definite descriptions, but also for other quantified expressions. When a speaker says to her audience "Everybody's tired now", her statement is not (in general) about everybody existing in the universe, but about the relevant people who constitute her audience. Here too, the context plays the crucial role of restricting the domain of interpretation of the quantifier. Nevertheless, this is not the exclusive role of context as far as natural language quantifiers are concerned: context is also required in order to disambiguate between collective and distributive interpretations of generalized quantifiers (see Hofweber, 2000), e.g. between:

---

<sup>11</sup> Since translating NL sentences into first-order logic is not expected to yield synonymy, in a sense Quine's criterion is not to be expected to account for the meaning of NL quantifiers.

(1) Four philosophers carried three pianos

and:

(2) Four philosophers carried three books

where the usual reading of “three” (either three for each philosopher, or three for all) depends on what things were being carried.

Thus, the key role of context in natural language is not connected to the massive use of indexicals or of underspecified expressions only. Frege’s project of a *Begriffsschrift* devoid of ambiguous symbols, as well as Quine’s intention to make use of *eternal sentences* only, might result in a completely context-independent semantics. Nonetheless, as the above cases indicate, this would not get us rid of the main issue: context plays a pervasive role in the interpretation of natural language quantifiers.

The linguistic situations upon which Quine’s criterion is built are very restrictive and artificial. Like Frege and Russell, Quine restricted his attention to cases in which context plays no role at all, hence to cases in which language can be considered as a device merely designed to describe the world. Those are the typical cases where the representational view of semantics can apply, in conformity with a certain universalistic conception of logic and language. By contrast, the way people use natural language in general, and quantifiers in particular, is heavily context-dependent. The procedural view of semantics is thus quite natural insofar as we are interested in how people really speak and understand language. As we shall now show, from the latter perspective, quantifiers do not bear any ontological commitment. Indeed, members of TSO are conceived of as mere semantic artefacts, and no longer as *bona fide* objects.

### 3.2 Anaphora, Skolemization, and Quantification

The so-called *dynamic turn* in Natural Language (NL) semantics occurred during the 80s, as a response to the difficulties raised by such phenomena as anaphora. Indeed, anaphoric pronouns—i.e. pronouns that somehow “co-refer” with an antecedent expression (proper name, definite or indefinite description)—cannot be fully analyzed within first-order logic. This is for several reasons, including the following ones:

- (i) Anaphora can occur between sentences, whereas variables can only be bound within a single sentence; moreover, in many other cases, logical dependence appears to be at odds with syntactic scopes, as skolemization shows;
- (ii) Anaphora involves information flow, beyond and above purely semantic information.

Inter-sentential binding can be seen through obvious examples where a pronoun refers to an individual introduced in a preceding sentence, like in:

- (3) *A man eats pretzels. He choked on them.*

Some of the simplest cases like this one could certainly be treated within usual first-order logic by means of wide-scope quantifiers: the pronouns (formalized as variables) in the second sentence would then be bound by the quantifiers in the first sentence. The method, however, could hardly be extended to a one-page discourse, and does not work in the general case.

Several theories resort to choice functions and Skolem functions to account for indefinites and anaphoric pronouns. Such functions map non-empty sets onto one of their elements, possibly taking into account parameters that depend on one or several objects in the current domain. Whatever the technical implementation, what is important here is that the functions are required to process wide scope (Reinhart, 1997) as well as narrow scope indefinites (Schlenker, 2006; Winter, 2004). In the above example, a (constant) function  $f$  would be introduced in processing “*a man*”, and would then be used to process the “*he*” in the second sentence. To take a more interesting example, in order to account for:

- (4) Every man loves a (certain) woman—his mother

the following formalization is employed:

- (5)  $\exists f[SK(f) \wedge \forall x[man(x) \rightarrow love(x, f(x, woman))]]$

The primary benefit is that formalizing the sentence in this way makes it possible to accommodate further anaphoric reference to the selected individual, whereas this would not be possible by means of a standard FO formula like:



$$(6) \quad \forall x \exists y [man(x) \rightarrow woman(y) \wedge love(x, y)]$$

where the scope of the existential quantifier ( $\exists y$ ) cannot go beyond the matrix of the formula. Yet, since such formalizations require functional quantification, should we consider that NL quantification amounts to hidden second-order quantification? We do not think so. The entities being quantified over in the previous example are usual individuals—so what?

That puzzling situation is due to the fact that NL quantifiers do not simply encode semantic information, but also involve a properly dynamic component, frequently called *information flow*. Actually, this component can already be found in first-order quantifiers. Hintikka (e.g. 1997) constantly criticizes Frege for holding a short-sighted view of quantifiers as second-order predicates. According to him, Frege missed the point: quantifiers do not only pick out individuals from a domain, they do it in conformity with certain patterns of mutual dependence. For instance, when an existential quantifier ( $\exists y$ ) is within the scope of a universal quantifier ( $\forall x$ ), its value depends on the choice of a value for  $x$ .

Natural language seems to make use of many patterns of dependence and independence between quantifiers, and some of these patterns cannot be formalized in standard first-order logic. This fact basically explains the success of Skolem functions in NL semantics, as well as that of other theories that conceive of quantifier scopes more “freely” than on their classical understanding.<sup>12</sup>

### 3.3 Presupposition

One of the most popular theories in NL semantics since the 1980s is Hans Kamp’s *Discourse Representation Theory* (DRT) (see Kamp, 1981; Kamp & Reyle, 1993). It provides an interesting insight into semantic issues and their connections with ontology. DRT is a two-step semantic theory. The first step, specific to the theory, consists in constructing Discourse Representation Structures (DRSs). This is where anaphora resolution is handled. The second, more standard step is that of model-theoretic evaluation: it connects DRSs to standard model theory. One can choose either

---

<sup>12</sup> Let us mention branching quantifiers (Barwise, 1979), Game-Theoretical Semantics and IF logic (Hintikka & Sandu, 1997), Dynamic Predicate Logic (Groenendijk & Stokhof, 1991).

an extensional interpretation structure or an intensional, possible world interpretation structure.

During the first step, DRT thus adds to the model-theoretic floor a *representational level* consisting of the DRSs. This supplementary level regiments contextual information, e.g. the order of sentences, that does not feature in the model-theoretic evaluation of discourses. A DRS  $K$  is an ordered pair  $\langle U, C \rangle$  composed of a *universe*  $U$ , i.e. a set of *representatives* (*discourse referents*, or *reference markers*), and a set  $C$  of *conditions*, namely the properties and relations ascribed to the reference markers. During discourse interpretation, discourse referents can be added to the universe independently of there being a corresponding entity in the domain of the interpretation model. Therefore, unlike other theories, DRT does not process the anaphoric relation as if it were a genuine co-reference phenomenon: the ontological neutrality of DRSs enables us to account for anaphoric relations even in cases where there is a lack of reference.

More: as *each* singular term gets processed, in the simplest cases DRT systematically introduces a discourse referent into the universe of the corresponding DRS. In other words, the denotation of singular terms is automatically *presupposed* to exist. For instance, because most of the usual definite descriptions are incomplete, they are treated like anaphoric expressions: an antecedent is thus expected in the context of the utterance, i.e. a reference marker available from the DRS, so that the marker of the description can be identified with it. What happens when no marker has been previously introduced? Since the description cannot be resolved, a mechanism of *accommodation* adequately augments the discourse context with a marker that can play the role of the antecedent (Beaver, 1997, pp. 976, 989f.). This mechanism therefore induces context-shifts in order to update the DRSs so that the discourse process can go on. If a marker is introduced by accommodation, it can nevertheless be later cancelled by some revision, as in the following dialogue:

A. Santa Claus is coming.

B. No, he doesn't exist!

Those pragmatic mechanisms of accommodation and cancellation seem to be unavoidable if we want to account for the semantics of discourse. As Peregrin puts it: "For the semantics to become really dynamic, we must

turn denotations not only into ‘context-consumers’, but also into ‘context-producers’—so that an utterance might consume a context produced by a preceding one” (Peregrin, 2000). According to Groenendijk & Stokhof (2000), this is even one of the main characteristics of the dynamic turn: “What *is* new, is the focus on context *change*: interpretation not only *depends* on the context, but also *creates* context” (p. 6).

Resort to DRT can be misleading and suggest that because the theory stipulates an intermediate representational level between language and models, dynamic semantics requires a conceptualist or mentalist construal of semantic interpretation. Semantics and logic would then collapse into psychology. Fortunately, there is nothing of the sort. DRT discourse referents are not to be thought of as “mental entities”: they are neither more nor less than the functional role they play in the theory, namely the role of antecedents in anaphoric processes; since anaphoric processes are fundamentally inferential processes, discourse referents are determined by their *inferential role*.

Can such considerations carry over to model-theoretic individuals? It might be objected that if we go back to the evaluation of DRSs (i.e. to the second step of DRT), standard semantics takes over, so that for instance the discourse referent connected with “Cinderella” has no semantic value. Nevertheless, the example considered in this section indicates that two important features of DRT—the decoupling between the procedural and representational components of meaning, and the determination of the “referents” by their functional role in the evaluation theory—might not be specific to this special framework, but be a characteristic feature of dynamic (procedural) semantics in general.<sup>13</sup>

### 3.4 Understanding, Phenomenology, and Internal Truth

Other analyses, like Hofweber’s (2000) account of NL quantifiers, suggest a similar sharp distinction between two levels. According to Hofweber, a separation needs to be made between the *internal reading* or *inferential*

---

<sup>13</sup> This seems to be Peregrin’s opinion, who argues for a general reduction of “reference” to “inference” *via* “co-reference”: “If we recognize individuals [sets of individuals contextually introduced by dynamic semantics] and their inhabitants as mere tools to account for inferences, then the talk about reference becomes essentially parasitic upon the talk of inference—a *referent* is nothing more than an illustrious clamp holding certain inferentially related expressions together” (Peregrin, 2000).

*role reading* of quantifiers, and their *external reading* or *domain conditions reading*. The first case naturally obtains with sentences conveying partial information, as in “There is someone Fred admires very much”, where Fred’s attitude can be directed towards Sherlock Holmes, while the second case is obvious in “There is something to drink in the fridge”. Just like with DRT, it seems that linguistic understanding in concrete contexts only requires a neutral, procedural view of quantifiers, whereas ontologically committing interpretation is reserved to very specific situations. How are we to interpret those inferential levels?

Several (perhaps conflicting) interpretations of DRSs can be given. Asher (1993, p. 64) mentions three of them: logical form, partial model, and mental representation. As we just said, the latter interpretation can be escaped. Asher himself conceives of DRSs as homomorphic with the structure of mental states, i.e. as partial representations of the cognitive structure conveyed by the interpretation of a discourse. A minimal interpretation of the DRT formalism conceives of DRSs as mere syntactic objects. According to such a view, DRT does not provide an extension of first-order logic in the usual sense: no real increase in the expressive power of the language would uncover new features of the model-theoretic structure.

In any case, the principal goal of DRT remains to account for anaphora resolution and some inferences can be handled at the proper level of DRSs, sparing us a detour through first-order logic (see Cooper et al., 1994, p. 54ff.). Of course, this does not suffice to qualify DRT as a non-syntactic theory, but it nevertheless gives its representational level a semantic flavor. DRSs must be thought of as encoding the process of discourse interpretation that makes further semantic interpretation possible.

Discourse interpretation can be carried out without confronting reality, i.e. without semantic evaluation. Except for underspecification cases, the process of constructing a DRS is almost completely independent from that of evaluation. A semantic restraint of this sort is certainly assumed in the case of anaphora resolution, since anaphoric linking must be solved *before* semantic evaluation can take place. Besides this, the restraint in question has other advantages, like providing a non-trivial interpretation of our talk about non-existent entities, or like resolving contradictions at the representational level without total collapse.

An immediate corollary is the possibility of accounting for the notion of truth at the intermediate level, as being local and internal, taken as another name for the consistency of the DRSs.

Even though we need not endorse a mentalist reading of DRSs (or of other theoretical entities meant to account for the procedural aspect of meaning), it should however be noted that the way DRT processes language and presupposition is somehow akin to phenomenological “bracketing”. In a certain sense indeed, dynamic semantics provides a picture of how people reach an internal perspective on the world through discourse interpretation.

## 4 Towards a Neutral Semantics

As we saw in the previous section, several NL phenomena cannot be analyzed within first-order logic, that is, within the *locus naturalis* of Quine’s criterion. Semantic analysis requires a procedural approach as a first step reflecting the way people interpret or understand NL utterances, before a possible, more classical and ontologically committing account can be given, that connects an interpreted language to the world. The ontologically neutral conception of semantics put forward in the previous section allows many choices that remain unavailable to Model Realists, so that our (procedural) theories suitably account for our linguistic intuitions. According to a neutral view of this kind, procedural semantics only aims at modelling linguistic meaning, whatever the world is like. Put in the Quinian idiom: the neutral view leads to the *naturalization of semantics* and to an anti-revisionist strategy.

Focusing on the cognitive value of *any* usual proper name (empty or not), a naturalized conception of semantics will naturally favor a uniform analysis and account for the aforementioned case of “Santa Claus” in a child’s way. With this broad (ontologically) neutral view, what can be expected from a theory is a meaningful account of the contribution of contextual domains of discourse to the semantics of natural language sentences: ontological considerations about the status of the domains are beyond their subject-matter. If a theory aims at a true description of linguistic meaning, regardless of the ontology, it needs no artificial (meta-

physical) meta-theoretical constraints on the constitution of TSO: neutral semantics is therefore characterized by a kind of meta-theoretical freedom.

TSO variability (“ontological relativity”), as it is induced by model theory, does not need to be restricted. Defined as relative to a structure, the Tarskian concept of truth should not be thought of as the universalistic, absolute one. Nor should we consider such and such TSO as made of *bona fide* objects. Model Realism thus appears to be a coarse transposition of the universalistic outlook to a frame where it becomes meaningless.

We can summarize the findings of the first part of this paper in a few points:

- (i) Semantic values, including truth, can be conceived of either in a procedural way, as being internal to the interpretation level of discourse, or in a representational or referential way, as resulting from some relations of language to the world;
- (ii) Internal semantic values, including internal truth, are sufficient to account for linguistic understanding; they are disconnected from all ontological commitment;
- (iii) Quine’s criterion may be relevant as far as we are interested in representational semantics, but not at all if we are concerned with procedural semantics, that is, with the interpretation level of language.

The intermediate conclusion reached so far is the lack of general entailment between truth and ontological commitment. Applied to fiction, however, this only means that we may avoid commitment to fictional objects by acknowledging a kind of internal (procedural) truth. In the second part of the paper we claim that fictions can be true in a strong and external sense too. To this end, of course, representational semantics must be abandoned—for otherwise, we would have to restore Quine’s criterion and ultimately admit of fictional objects. We defend that the connection between fiction and the external world is not representational, but of another nature, and that because fiction is indispensable in a certain sense, there is also a genuine sense in which a fiction can be considered as true.

## PART II - FROM INDISPENSABILITY TO TRUTH

In this second part, we rely on the Indispensability Argument to argue for the utility of fictions, and from there to the conclusion that there is a sense in which fictions can be true. The issue, then, is no longer so much that of truth *within* a fiction, than that of the truth *of* a fiction. The underlying idea is that because fictional narratives make an interpretative rather than a descriptive use of language, they may not be subject to theoretical reproof.

We exploit a number of considerations drawn from literary criticism and the philosophy of art to back up our argument. We also choose to focus on novels and novelists' thoughts as they will neatly fit our purpose. After specifying what we mean by the "indispensability" of fictions, we then make it more precise what kind of truth(s) fictions appear to us to express.

### **5 The "Indispensability Argument" and Its Revision**

Historically, the so-called argument was developed and meant by Quine (1948/1961, 1960/1976, 1981) and Putnam (1979a, b) as a contribution to the philosophy of mathematics. Because mathematical theories play such a central role in all branches of empirical science, philosophers have argued that they are indispensable in precisely that sense, and that their indispensability gives us good reason to believe in the existence of mathematical entities. The strategy can thus be used to reinforce the thesis of mathematical realism, as Colyvan (2008) notes: "reference to (and quantification over) mathematical entities such as sets, numbers, functions and such, is indispensable to our best scientific theories, and so we ought to be committed to the existence of these mathematical entities" (p. 1). Of course, many other forms of the Indispensability Argument have emerged since then, and adjustments have been proposed, especially with respect to the relevant sense of "indispensability". It is not our purpose to engage into the debates over the classical version of the argument, but to examine how it might fit the case of aesthetic fictions.

First of all, we are not willing to support fictional entities, since we argued in Part I that no ontological commitment necessarily ensues from the truth of a fictional discourse. What remains to be done is to cut off the inference from the indispensability of fictions to the existence of fictional entities, so that we are left with the only idea of indispensability itself. However, it is important once again to note that fictions are neither essentially determined by any kind of applicability to empirical sciences, nor by any kind of purely descriptive or explanatory function. Therefore, the intended weakening of the argument is aimed at dismissing the original link between being indispensable and being indispensable *to our best scientific theories*, which amounts to rejecting the idea of theoretical indispensability.<sup>14</sup> The reason for this is the distinction we assume between two possible uses of natural language: a descriptive use, involved in the classical version of the argument, and an intentional use, involved in the kind of discourse that proceeds from how we practically grasp the meaning of events and behaviors.<sup>15</sup> This contrast allows us to think of fictions as being deeply connected with the latter use, as they do not *seriously* aim at describing the world as it is. So, fictions are not *stricto sensu* indispensable to our best scientific theories, yet they are indispensable to a better understanding of the meaning of life itself. We are thus after something close to the notion of a practical, hermeneutical assessment of fictions.

Besides its ontological reading, the Indispensability Argument can receive another one. If we admit that there is a sense in which mathematics, or fictions, are indispensable, then there is a sense in which mathematical, respectively fictional, discourse can be considered as true too. Just like it seems hard to deny the truth of mathematical theories in the light of their empirical success, likewise it would seem hard to deny the truth of fictions in the light of their practical success—or so shall we argue. This is likely to sound rather odd, though. Fictions are often said to be false, wrong, sometimes malicious, and literally defined as non-serious, and this for the very reason that *they are fictional*. Consequently, our goal is first

---

<sup>14</sup> As Morizot suggested to us, this could be viewed as a generalization and an extension of Quine's holism beyond its original boundaries.

<sup>15</sup> It is worth noting that just like in Part I, our purpose is to ground truth into *interpretation*, although here we are concerned with the interpretation of people, actions, etc., while in Part I the issue was that of linguistic interpretation. Common to this and the previous part is the goal of accounting for truth in a non description-involving way, hence without ontological commitment.



to give an account of the special indispensability of fictions, and secondly to make sense of the idea that fictions can provide us with truths about the world and other people—we shall make it clear what kind of truths they are.

One last thing before we proceed: contrary to mathematics, fictions are not in themselves *a priori* indispensable, in the sense that they are not created in order to explain reality, describe what there is or improve our knowledge of the world.<sup>16</sup> They are not produced to be scientifically useful, neither empirically, nor theoretically. We could thus say that fictions are not indispensable, yet it is indispensable to have fictions: once it has been written or filmed, even if it is not *per se* indispensable, a fiction can help us interpret and understand certain features of the world and of human attitudes.<sup>17</sup> Fictions are *a posteriori* indispensable as a tool for improving our general and intentional understanding of life.

## 6 Indispensability as Irreducible Utility

*J'ai lu beaucoup et je comprends de quoi il s'agit.*

Romain Gary, *Les racines du ciel*

Fictions are usually thought to be useless: telling stories is nearly the same as telling foolishnesses. This opinion cannot only be found in everyday-life, but in philosophy as well, transpiring from a number of logical, moral, and even aesthetical reflections. It is as if it were commonly agreed that, just because they are fictional, fictions were silly, illusionary discourses, and were thereby neither indispensable nor true. Recall Plato's condemnation in the *Republic* (X, 605d–606c), Furetière's (1666/2001) in the 17th century, and Frege and Russell for whom sentences in fiction are never true<sup>18</sup>. As a matter of fact, this kind of allegiance seems to have been the common starting-point for a great number of philosophical approaches

---

<sup>16</sup> This claim can be nuanced a bit more by distinguishing between fiction and myth: the latter might have been initially intended to explain the origins of the world, but it would still be a symbolic rather than a theoretical description.

<sup>17</sup> This idea is close to the epistemic—as opposed to the aesthetic—reading of aesthetic cognitivism. See Darsel & Pouivet (2008).

<sup>18</sup> More accurately, fictional propositions are considered by Russell to be always false, and by Frege to be truth-valueless.

to the status of fictional narratives. More recently, Lamarque & Olsen (1994) have promoted an “anti-truth” line of thought regarding literature, although they defend at the same time a humanistic conception of fictions that acknowledges their indispensability<sup>19</sup>.

On the opposite side, as Bouveresse notes (2008, p. 13), we find “mystical” positions that laud the superiority of fiction in the magical quest for Truth. Only art and literature (or poetry) can reach the profound and essential truth of Life. Fiction might even be Truth itself. We will have nothing to say about this line of thought—too excessive, too marginal.

While objecting to the utility of fictions, people and philosophers adopt a scientific outlook, thereby assuming that fictions rely on the descriptive function of language. If they were really written in a descriptive context, nothing would distinguish them from science or philosophy, and at worst, they could be judged dispensable. In particular, the explanatory mission of 19th century realist and naturalist novels, Balzac’s or Zola’s for instance, and their purported objectivity and impartiality, could be thought to be directly challenging the goals and methods of empirical sciences. The relevant notion of utility would thus be relative, for one may well wonder why those novelists favored a fictional treatment of questions that are best analyzed by human sciences like philosophy, or nowadays psychology.

The problem lies in a confusion between two uses of language. When we say that novels *describe* certain features of the world and of human hearts, the term “description” is not to be taken in its scientific and referential sense. Fictions do not describe anything in this sense: fictions are always a medium for *interpretations*. In other words, those who claim that fictions are useless assume a purely descriptive use of language, and their disappointment in this respect leads them to discredit all kind of indispensability which fictions might be credited with. Yet, if we bring into light that fictions make an intentional use of language and if we thereby manage to tell serious from non-serious discourse (in terms of dissimilarities in methods, ends and objects), then we can show in what sense non-serious, non-descriptive narratives are indispensable.

---

<sup>19</sup> Their own view is that “the concept of truth has no central or ineliminable role in critical practices”, that is to say, “there is no significant place for truth as a critical term applied to works of literature” (1994, p. 1). Lamarque & Olsen are not reluctant to the idea of a specific cognitive function for fiction and literature, but disentangle this question from that of truth. On this point, we might disagree.

One might nonetheless object that we are making an instrumental use of literary creations. We are not, though. Highlighting their possible utility is not the same as reducing them to their practical helpfulness. Fictions do have many other aspects and purposes, including that of absolutely not being created to be useful or used at all, as mere aesthetical products. We now turn to the question of their indispensability in connection with the debate over their truth (or falsity).

## 6.1 The Utility of Fictions

Let us explicate the kind of utility that we promote: practical utility. Fictional creations can certainly be useful in many other ways: according to different authors, they can allay our sorrow or grief, (cor-)respond to a primitive need, incite social revolts, purify our souls, and so on. The point we want to make, however, is rather that fictions, novels in particular, are helpful to our apprehension of life.<sup>20</sup> The reasons why novels are practically useful—to gain a more accurate understanding of the real world and real people—are thus multifaceted. We now propose an inventory of those reasons, based mainly on considerations from literary criticism, ordering them from the least to the most specific function.

**(i) Describing.** Novels have a heuristic value in that it points to, and reveals, aspects of reality that were ignored until then. Novels could thus sometimes be considered rivals to the work done by historians, as many authors have claimed. For example, Victor Hugo defines himself as an “historien des mœurs et des idées”, “des cœurs et des âmes” (1862/1967, Vol. III, p. 10). Novelists, like Jane Austen, Marguerite Yourcenar, Henry James, and almost all writers, would have the power of uncovering interesting traits of human psychology. Fiction seems to provide us with knowledge of human beings and reality, working as an unprecedented source. But as we said before, this way of talking forms a kind of “naïve” view: it is not properly speaking a view on the utility of narratives, and it is not at all there that their specificity lies.

**(ii) Spreading.** Novels convey scientific, philosophical, and moral knowledge, like the so-called didactic poems by Lucretius or Boileau, or

---

<sup>20</sup> This idea echoes Walter Benjamin’s study of the function of the “narrator”. See Benjamin (1936/1991).

like La Fontaine's fables. Their value can thus be didactic, though once again, that value is not specific to them.

**(iii) Enhancing.** Novels can expand and improve our mental faculties (intellectual and sensible), and in particular, as Pouivet (2008) stresses it, our modal competencies. This is one of the ideas developed by the advocates of aesthetic cognitivism. Here, the value of novels is cognitive, but once again, not characteristic.

**(iv) Taking Distances.** Novels question and destabilize dominating viewpoints, offering new fruitful perspectives on what happens or has happened, on what a uniform social discourse is meant to hide, on what appears to be well-established opinions. Their value here is critical: we can think of polemical or provocative stories, or of various notions turned upside down—barbarism in the works of Aeschylus, Flaubert and Coetzee; stupidity in those of Grass and Singer. Highbrow essays could do the job very well, but we still want to argue that there is a special flavor to fictional criticism.

**(v) (Re-)Arranging.** Novelists sometimes try to arrange and give meaningful structure to all kinds of fact or knowledge, willing to reduce the feeling of chaos and absurdity that we may experience about the world and the way we talk about it. Because they want to “give sense”, some authors condense intellectual, social and empirical matters, and include them into a story with seemingly encyclopedic purpose. Think of Musil, Broch, Borges, Flaubert, and today, Senges. Some of them associate that tendency with an obsession for taxonomy. The value of fiction here lies in synthesizing information, a function that is already more specific to fictional works in that they are meant to interpret the world in its heterogeneity.

**(vi) Surveying.** Novels thoroughly clarify features of reality that are already known, by extending our intentional vocabulary. Descombes observes that reading stories provides an “enlightenment of the vocabulary we use to describe human areas” (1987, p. 16). We increase our awareness of various aspects of life that were not so obvious at first. For instance, moralists like La Bruyère or dramatists like Corneille and Molière reveal dark parts in the functioning of our passions. More generally, digressions and meditations within the plot are meant to play this role, as Kundera emphasizes (1986/1995). And the efforts made by many authors to grasp, for instance, the spirit of a city can be analyzed in that perspective (e.g.

Vienna seen by Bernhard, Musil, Jelinek, etc.). It is distinctive of novels that they are analytical, and as such, pedagogical. In fact, they educate our perception by offering us the possibility to pay more attention to ordinary things. For example, Italo Calvino defines his book *Palomar*, as a sample of a “pedagogy of looking and reflecting” (1985/2006).

**(vii) Interpreting.** Last but not least, a novel creates new meanings. It can suggest different interpretations of human behaviors and social phenomena, developing different frames for decoding the world. Polyphonic novels illustrate the point very well, like Kristof’s trilogy, Durrell’s *Quatuor*, or Barnes’s *Talking It Over*: intertwining viewpoints enables the reader to increase the meanings of the story, and as a consequence, the meanings of life. What is more, novels are sometimes bound to a particular metaphysical hypothesis, in a trivial sense of “metaphysical”; they can then help us gain one out of many possible understandings of reality. Kafka, Dostoïevski, Beckett, and others invite us to read the world in light of its absurdity, its tragic-comical side. The value of fictions, then, is hermeneutical, properly speaking this time.

In brief, (vi) and (vii) are what we claim to be the central contributions that fictions make. They jointly characterize what makes their usefulness so specific. Moreover, (iii), (iv) and (v) may be considered either as modalities of, or as means towards, or even as results of that particular usefulness. Point (i) is controversial depending on what we mean by “description”, and (ii) is quite contingent, being exemplified by many novels, but not exclusively by novels. The foregoing inventory as a whole leads us to claim that, *pace* Plato and Quine, fictions are undoubtedly useful.

## 6.2 The Irreducibility of Fictions

The indispensability of fictions is defined by two features: they are (i) useful to understand the world and human existence, and (ii) not reducible for that purpose to any other means of expression.

Novels can be said to be irreducible if and only if their content cannot be paraphrased by any other formulation. Dispensability definitely asks for the possibility of a translation, say, in the form of some canonical notation. This would imply that we can eliminate the fictional in favor of a theoretical, or at least descriptive, use of language. The point is that, although it seems possible to reduce certain statements in a work of fiction

at the local level, it still remains difficult to reduce the fiction at a global level. Whatever the lesson we extract from a story, it does not express at all the whole hermeneutical process in progress. Why is it so? Six features can roughly be sketched out.

**(i) Complexity.** As Kundera claims, “the spirit of the novel is the spirit of complexity” (1986/1995, p. 30). According to him, the “eternal truth of the novel” is to tell the readers that “things are much more complicated than [they] think”. Moreover, Kundera insists that the specific hermeneutical utility of fictions resists the tendency to standardization: “the novel is [just like the world] threatened by the termites of reduction, which do not only reduce the meaning of the world, but also the meaning of the works”. What is crucial, then, is to take into account, and preserve, this multiple interpretability.

This goes hand-by-hand with the ironic, humoristic and playful dimension of the novel, which is a constitutive dimension of the offered understanding. It pushes away and questions our theoretical grasping of real events. Furthermore, because fictions enjoy the privilege of an entire freedom in the treatment of serious themes, authors are allowed to draw aside rational norms, scientific requirements, or reasonable viewpoints in order to create stimulating echoes of meaning. According to Rozakis, novels are attached to the *praxis*, and the function of imagination is to show that there is no straight *a priori* rule to know whether some principles are controversial, or to know when and how they are so (2009, p. 18). So, contrary to the descriptive or theoretical use of language, language in its fictional use partly creates room for paradoxes and exploits their evocative power.

**(ii) Non-actuality.** Novelists look for meanings beyond what really exists or happens. For instance, the way in which Musil presents certain facts in *The Man without qualities* indicates that the bundle of possible meanings for an event, as it could or should have happened, is much more important than the event itself when it comes to grasping the deep meaning of things. The idea is to dismiss the privilege of a literal one-dimensional viewpoint on actual reality. Rozakis also claims that inquiring about human affairs requires fiction, where fiction does not consist in describing existing customs, but in the quest of some future good (2009, p. 15). The entire axiological dimension of fictions is connected to this deviation from current data.

**(iii) Symbolization.** Fictions are generally based on the use of symbols and metaphors. The specificity of such non-literal expressions lies in their having various possible interpretations, at least several variable ones. The novel is the art of detail: its irreducibility comes from that evocative brightness of concrete things, considered at first blush as insignificant dust. And sometimes, from this perspective, we are urged to read the story of one particular man as a relevant mirror of human life in general. The process is therefore much more analogical than logical (in the trivial sense) or explanatory, because the relevant notion of meaning lies in emblems and pictures rather than in formal argumentative systems.

**(iv) Content-Form Dependency: Local Level.** There is a necessary inseparability between what is written and how it is written, between what the story tells us and how it tells it to us. This is the reason for choosing narrative means rather than theoretical or descriptive means. Descombes notes that an idea becomes genuinely novelistic when the writer has found a way to “analyze” it, that is to say, to turn it into a schematic scenario (1987, p. 90). The point is that the specificity of fictions results from a narrative analysis of the concepts, ideas, or questions involved, for example, breakdown of values in Broch’s famous book, or adultery in *Anna Karenina*. Therefore, it is through the story that the reader follows the gradations, the developments, and the contents of one or several (theoretical) thoughts. In this sense, it is difficult to imagine how one could strictly paraphrase what a fiction says, and therefore impossible to reduce its structural aspect—that which helps us understand a character’s (and by extension a human being’s) reactions or choices. This proceeds from a kind of heterogeneous dynamic impulse. Rivière (1913/1999) distinguishes between poetical and novelistic emotions, precisely because the latter kind ensues this chronological movement: we understand neither instantaneously nor directly, but progressively; and the content is not fixed, but changing.

As a consequence, fictions allow us to test the full meaning of some concepts, by incorporating them into a plot, or into a character. Many features of those concepts—for example, the practical effects they could have in “real” life—could not be disclosed by *a priori* theoretical means. In other words, the “veil” of imagination plays the role of an intentional adjuvant impeding (1) the possibility of being paraphrased by a documentary or by a systematic theory, and (2) the possibility for the fictional dis-

course as a whole to be of a descriptive kind, although it may include various descriptions. Thanks to this specificity, what seems to be descriptive, hence eliminable, happens to be mingled into an inseparable intentional and aesthetic frame which turns the content into a contribution to our understanding of the world. A slogan might be:

*Different means of expressing the same idea yield different understandings of that idea.*

In general, a movie, a comic book, a cartoon, and a novel do not shed the same practical lights on us.

**(v) Content-Form Dependency: Global Level.** The work of representation itself seems to endorse meanings.<sup>21</sup> We might think the indirect imitative process to possess hermeneutic value by itself, for it structurally organizes our words and our world. The spirit of nomenclature that OuLiPo writers display rests on the idea that formal properties, even arbitrarily chosen, allow us to read reality from different angles, thus offering different intentional hypotheses. Likewise, the stylistic and poetic aspects of fictional discourse, relying on rhythms and sonorities, also involve a kind of intentional understanding that can hardly be paraphrased. How could the musicality of a tale be rightly eliminated?

This point relates to Goodman's idea that understanding and creation go together (1978). According to Morizot (2008), the idea is that (1) contemplating a work of art or following an artist's intentions do not exhaust what has to be understood about the artwork, and that (2) creation is the most fruitful form of understanding since it first has to create its own conditions of relevance, and then to give them communicable and sharable forms. The same applies to works of fiction. Neither the many possible interpretations of a myth nor the multiple readings of a story could exhaust their respective meanings.

**(vi) Incompleteness.** Finally, novels are like open frames. They produce a sort of fragmented meaning, unfinished, intermittent, that is beyond systematization. This partly explains why their reception by the reader and the work she does are so important. In other words, fictional contents cannot be reduced to positive fixed knowledge, because fictional statements

---

<sup>21</sup> This is one of the main reasons why translating literary works may prove to be problematic. We do not address this issue here.



are always incomplete, in need of interaction with us. It is incomplete because (1) there are questions about a story that are neither checkable nor answerable, and (2) there are hypothetical interpretations that depend on the inner and the outer context, i.e. on the context of the story itself, and on that of the reader. In accordance with (1), we may say that, contrary to descriptive discourse which is underdetermined, therefore asking for completion, fictional discourse is undetermined, therefore allowing multiple interpretations. This multiple interpretability shows in the second aforementioned aspect of incompleteness. Indeed, in accordance with (2), the meaning of current fictional data would, in itself, be interpretable in various ways.

That does not mean that we could carelessly propose any possible interpretation, without paying attention to the letter of the text. We simply mean that fictions are like centripetal media: their meanings, just like the meaning of reality itself, cannot be reduced, and their open dimension gives right to Calvino when he writes that novels are more clever than novelists. The fact that we can interpret one story again and again suggests that we can do the same with real world events and real life choices. A slogan can summarize the idea:

*Different interpretations of the same fiction yield different abilities to interpret human life.*

This plural hermeneutics of fictional narratives comes from their being relative to the readers, their age, thematic focus, and so on. In short, we claim that there is no unique interpretation of a novel, and that no fiction offers a unique interpretation of the world.

Again, although we went from the least to the greatest specificity of the practical utility of novels, we must keep in mind that each of the above features contributes to support the idea that fictional discourse cannot be eliminated or translated into a purely descriptive language. Fictions thus break out as being really indispensable for our grasping of the intentional side of the world.

### 6.3 The Relation of Fiction to Reality

The main problem with fictions having practical utility is, perhaps, that we have to admit of a necessary connection between fiction and reality. To be more precise, the question is, first, “Are fictions linked to the real world?”, and second, “How?”, whereas fiction and reality are commonly seen as opposite terms. Our thesis does imply a certain porosity between the two, maintaining at the same time a distinction between them. We disagree with opposite theses like formalism, like the self-reference of fictions, or like Baudelaire’s claim about the autotelic value of works of art or literature (1857/1976). In particular, Genette’s idea of water-tight fictional worlds (1991, pp. 58–60) appears to us to be erroneous, as it implies that novels cannot be helpful to anything but themselves and literary history (see also Macdonald, 1954).

Aesthetic cognitivism too assumes certain connections between reality and fiction. The possibility of learning from a novel about the real world presumably requires a link of this kind. And even when novelists set out to talk about how things could ideally happen instead of how they really do, we may still think that this moral and modal stance remains closely linked to real life. According to Fraisse & Mouralis (2001) the abstract, imagined reality is no less real, because “its advent is considered possible” (p. 152). Measures of likelihood and credibility regarding a novelist’s moral considerations then entail a dialogue between what falls *within* and what falls *outside* the fictional frame.

What fictions help us understand is part of reality. But it differs from descriptive and theoretical explanations in that it is much more a kind of interpretation than genuine knowledge of reality. Intentional discourse in fiction can neither be corroborated nor falsified by experience, empirical observation, or experiment. Yet, we may judge a fiction to lack credibility or authenticity, typically when the interpretation it gives of reality seems partial or incoherent. Validity, then, results from the application of some criteria, such as the unity of the narrative plot, the respect of good manners, the laws of the literary genus, and of course, the way it *echoes* real data—about human feelings, behavior, social functioning, etc.

Consequently, fiction and reality are connected in several respects; they enrich each other. Fictional universes have lots of similarities with the real one—objects, rules, individuals, spaces—and, on the other hand,

reality blends fictional parts—names of certain characters, possible interpretations of certain events. Even if the latter connection seems to be more of an analogical kind, it is a genuine connection, and this idea is all we need to establish our point! Fictions are indispensable because they are useful and non-reducible, and they are indispensable to our understanding of a unique shared reality.

If so, we can follow the line of the Indispensability Argument and directly claim that fictions, precisely because they are specifically indispensable for a better understanding of the world, can be true. We thus made our way to fictional discourse within the limits of the serious practical contributions they can have. Incidentally, this shows that our concept of truth is not antirealist: even when applied to non-descriptive discourses it remains linked to reality via practice, as stated by the Indispensability Argument.<sup>22</sup> But of course, all this may sound odd: fictions tell us something true while being at the same time false or illusionary creations of our minds. So, something more precise must be said about the nature of fictional truth.

Does it follow that the truth-as-correspondence scheme, based on Aristotle's definition of *adequatio intellectus et rei*—to say of what it is that it is—, is true? Obviously, this can hardly be the case here, because fictional discourse, by definition, does not stand in a direct descriptive relation to reality. Is fictional truth a merely built-up truth, one that holds only with the fictional domain? In this case, the problem would be for us to deal with those statements from the story that are literally true. In short, how is the concept of truth modified when applied to fiction, and in what sense can something coming from a fiction be true? From now on, we will take the indispensability of fiction for granted. But this indispensability only delineates the extension of the concept of truth we have in mind, and one will expect more than that. The account we sketch in the next section, more intensional in character, crucially takes into account what literary criticism, philosophy of art and effective practices of writers and readers have to say on the matter.

---

<sup>22</sup> Our proposal mainly consists in *widening* the scope of the concept of *truth*. The truth of descriptive discourse can be considered compatible with the Indispensability Argument, therefore compatible with our account.

## 7 Truths *of* and *from* Fiction

According to Lamarque & Olsen (1994), there is no reason to consider the notion of truth as central when we analyze fictions as critics and try to characterize their aesthetic value. Again, we do not claim to assess the power of fictions in terms of their degree of truth, but simply want to argue that because they are practically useful, they often present us with true interpretations of the world and true intentional hypotheses about human matters. The kind of truth they involve is objective rather than personal, even if there can be different subjective ways in which different characters see the world. Truths from fictions are public and debatable, being spread and rendered objective by the writing itself.

There are many sorts of truths in a story: fictional truths have much in common with non-fictional serious referential truths. Again, we must sort out those that are specific to a narrative treatment of certain subject-matters. Fictional truths are, just like their utility, characterized by the essential interdependency between content and form. If the propositional content of a work of fiction could be made autonomous, fictional truth could be suspected of being reducible, in the sense of capable of occurring in a theoretical discourse.

Some distinctions are therefore necessary to determine the very nature of truth in fiction: on the one hand, the difference between propositional and non-propositional truths, and on the other hand, that between literal and non-literal truths. We will later see how this double distinction works. Moreover, it is important to keep in mind the slight difference between what has to do with truth and what has to do with verisimilitude—as Aristotle defines it in his *Poetics*—, for the novelist’s moral and modal quest crucially depends on this.<sup>23</sup>

### 7.1 The Eliminability of Factual Propositional Truths

When we read a novel, we sometimes run into statements that are literally true, that is, which are referential and to which the truth-correspondence scheme applies. These propositions are what Schaeffer (2008) calls “the tracks of the non-fictional universe within the fictional universe” (p. 79).

---

<sup>23</sup> See Genette (1969, p. 74) for the difference between the notion of verisimilitude and that of motivation.

They can actually be checked against empirical facts. At this first level being true amounts to sticking to reality.

The entire documentary dimension of novels falls under this kind of truth. The plot will have a social and historical setting that has really existed or currently exists. This is the case of naturalist novels, which tell us about the conditions of French society at the time. Fiction, then, is expected to provide a true account to historians: novelists do the job of registering various facts. Zola's use of exact names for the tools used by each corporate body, or Hugo's choice of including slang in his writings, or DeLillo's approach to financial operations which really happen in our world, may serve to illustrate the point. Moreover, novelists sometimes use real persons as characters in their story, including biographical data in the narrative. For example, Stridsberg recently wrote a novel whose main character is Valérie Solanas, the woman who actually tried to murder Andy Warhol.

Another kind of literal truth comes from the theoretical dimension of various fictions. The didactic aim of many of them, the encyclopedic tendency of those novels that include fixed knowledge, the essayist aspect realized by the presence of digressions and abstract meditations in the story are examples of how serious discourse may enter fictional worlds.

However, fictions are not indispensable for that precise job: the kind of literal truth that comes from the theoretical dimension of a fiction is not specific to fiction, since we can reduce it to serious true statements about facts that are already known. Nevertheless, we can make two comments:

(i) Factual propositional truths in works of fiction allow for an account of effective practices, those of readers and writers (see Fraisse & Mouralis, 2001, p. 196). A reader can find and learn true information during her fictional pastime. A writer may want to use her novel to spread true but insufficiently known facts. For example, part of the literature from former colonies can be thought to proceed from the willingness to show the reality of a culture or of a nation.

(ii) This kind of referential truth is what we set out to use to account for the literal truth of true statements in works of fiction. In other words, it is clear that fictional truths are based on many ordinary empirical truths on which we obviously agree. When a novelist writes that her character is eating this or that dish, or walking on such and such a street, this is not fancier than anything we could find in our everyday life. However, this

is exactly what science-fiction or fantastic novels modify, for example, specifying that in such and such a fictional world, human beings can fly or eat pure red energy-bubbles.

## **7.2 Interpreting Facts: A Two-Level Truth**

Besides non-fictional truths, another kind of truth seems to emerge from a fictional frame of investigation. This is the case of some statements which can be considered as literally true (i.e. non metaphorical), although they are non-referential and not testable empirically. Such statements rather propose a certain interpretation, a particular understanding of the meaning of real-life events and behaviors, relative to a general point of view or to the particular point of view of a character.

Indeed, when Balzac (1842/1965) confesses that he tried not only to reproduce a sort of mirror of society, but also to explain something like the “hidden meaning” of the architecture of social and psychological facts, he was then intending to make true, yet non-descriptive statements. What is expressed by the “present tense of general truth” makes for a second-level truth that no longer corresponds to a kind of adequacy to the facts, but to the relevance of many explanatory hypotheses offered by the fictional frame. As he wanted to describe the spirit of his time, Balzac was actually interpreting facts, and we may consider his interpretations to constitute a form of novelistic truth. So, efforts towards a general understanding of a particular situation lead us to propositional though irreducible truths.

Besides constructing true expressions about social contexts, novelists are also interested in uncovering truths about our feelings and other vague impressions, about human passions and states of mind. Fictions offer psychological depictions, for which truth is no longer factual but has to do with the inner life of a person. This feature seems to be genuinely specific to fictions because only a fictional treatment can develop that sort of thought. Sade claims that the novel tells us the truth about people once they have taken off their “mask”, and he considers the resulting depiction to be not only more interesting but even “truer” (1799/1970, p. 54). This intentional truth is therefore only possible through the narrative dynamic dimension of fictions; and so, the process of characterizing—by which we here mean the process of specifying as well as creating a character— involves the production of what Barthes (1978) calls the “truth of affects”.

We might have some reservations about such truths being literal; they are mostly literal. But from the perspective we adopt at this point, we could also include the novelists' frequent practice of giving linguistic ability to those who lack it in the real world: angels, the dead, trees, animals, and so on. In such cases, true statements appear to be quite metaphorical. They are a form of intuitive understanding of what is beyond us, an attempt to make sense of complex situations.

All in all, the reason why a fictional treatment is desirable has to do with the nature of the objects in question: the meanings of events as they are experienced, perceived or projected in many ways, human hearts, inanimate entities, dreams and subconscious thoughts, as well as life itself, the meanings of life and reasons to accept the existence we have, cannot be the *rei* in the conception of truth as *adequatio*. However, their occurrence within a fictional discourse can have as a result that we end up close to something true, for they are merged into the general tendency to understanding. "Being true" here means that we judge it to be as likely as what we really feel or think, how we behave in our lives, or what things mean to us. But this kind of truth is not reducible to clear and exact formulas; they are vaguely referential truths, mixed with their narrative expressions.

Because fictions direct our attention to moral and axiological questions, as Pavel (2003), Bouveresse (2008), Nussbaum (1995), and many others have claimed, we need to modify and extend the concept of truth so as to account for that fact. As we formerly said, there is no single valid interpretation of human attitudes, no scientific truth about the relationships between human beings, or about one's correct position in the world. We do not know what the true fixed principles are that would lead us to a good life! Nevertheless, because fictions are indispensable to clear up our mind about such deep existential questions, it is tempting to think that fictions tell us something true on those questions.

In this perspective, Bouveresse considers that a novel fulfils this function by "setting a problem" in the most likely, plausible and relevant way (2008, p. 49). The likelihood in question is closely related to the idea of the "truth of a person", and passes through the representation of his or her entire story, as English moral realists have tried to do. In this respect, we side with Rozakis (2009) in thinking that fictions, through endlessly iterated reconstructions and redefinitions of the "possible content of an entire life" (p. 78), propose different interpretations of human behaviors and rea-

sons for why we act in such and such a way. Truths from fiction are of an hermeneutic kind, therefore neither eliminable nor definitive.

### 7.3 Irreducible and Specific Truths: Four Features

The kind of truth that is really specific to fiction is quite general and intuitive, and rather of a conceptual order—here, we follow Bouveresse (2008). It is difficult to isolate what is true in a fictional story: what is true appears to us as a fluid whole, and is not *stricto sensu* expressed in propositional form. So, we could say that truths from fictions have to do with “showing” rather than “saying”, even if it may seem paradoxical in the case of novels. We simply have in mind what Wittgenstein suggests when he claims (yet in another context) that what cannot be said can however be shown (1921/1961, §6.522): in a way, what is true in a work of fiction cannot be captured with a few words, but only read between the lines. We now present four features that delineate the specificity of the concept of fictional truth.

**(i) Links to Falsity and Game.** It may sound contradictory to claim that what is true depends in this case on what is literally false. But this simply means that the way truths come to light depends on the illusionary dimension of stories. As Schaeffer (2008) outlines, the particular modes of reception and mental processes that give us access to fictional universes are immersion and simulation; these two modes determine what characteristic aspects fictional truths have.

So, illusion is indispensable to fiction, which in turn is indispensable to the elaboration of interpretative truths. The pretense dimension partly defines certain properties of fictional truths. We must do as if fictional statements were literally true, and only then can we understand how they could be true of the real world too, in an analogical as well as intentional way. We have to invent stories and characters in order to create an appropriate frame for true statements that cannot be expressed otherwise.

What is more, this feature is closely related to the ironic and playful dimension of fictional truths. Many authors enjoy blurring the supposedly clear-cut boundaries between what is true and what is false, between what is real and what is fictional—as explicitly stated in prefaces by Montesquieu, Laclos, or Rousseau (Montalbetti, 2001). Fictional truths seem to exhibit an ironic side.



**(ii) Links to Questioning Form.** This feature has its roots mostly in the incompleteness and complexity of fictions. Fictional truths, in this respect, can be characterized as unfinished broken intuitions with an evasive and changing dimension. They appear to be problematic and non-decidable, undeterminable and sometimes paradoxical. And, even worse, they do not always follow logical rules. In this perspective, fictional truths can be considered unsystematic and quite provisional.

But then, can we properly speak of “truths”? Yes, for the intentional aspect of language, as we already argued, do not respond to the same requirements as descriptive or theoretical discourse. In the case of fictions, there is no unique meaning, and novels allow for different possible interpretations of the world. For these reasons, it is important to insist on the plural form of the term, “truths”, for it would be erroneous to claim the existence of one Truth for a fiction. Furthermore, novelists create a sort of unknown truths and many of them often confess not to know themselves what kind of interpretation they will end up with: truths from fictions are not conclusions from well-formed premises.

**(iii) Links to Words.** Fictional truths are inseparable from linguistic work.<sup>24</sup> When we read a novel, we can uncover narrative truths, i.e. truths developed within a narrative scheme, but we can also discover truths that depend on the words themselves, such as analogical or metaphorical truths. This verbal dimension is, perhaps, closely associated with the natural tendency we have of telling stories in order to make sense—this anthropological aspect of fictions has been studied by Molino & Rafhail-Molino (2003): the intentional use of language allows us to formulate various meaningful hypotheses about ordinary situations. In that case, words are useful to make the “problem setting” clearer and subject to cognitive control.

Novels sometimes allow for a sort of external “typification” of certain names and narrative sequences; the reader isolates and possibly applies them to real-life events or individuals. For example, the Peter Pan type applies analogically to people who refuse to grow older; the Emma Bovary type or the model of Don Quixote characterize people who confuse dreams or fictions with reality; and we may even think of the mental process illustrated by Proust’s *madeleine*. Likewise, Balzacian “types” shed

---

<sup>24</sup> Obviously, this point concerns more specifically the case of novels; movies relate to another kind of medium specificity.

light on human behaviors.<sup>25</sup> Fictions are used to put names on real data, and in this respect too we can speak of a special kind of truth.

Truths from fictions are intrinsically linked to words, for the reason that they express real features for which we might lack the words before we read the works of fiction, that is, such features as emotions (see Roza-kis, 2009, p. 117). In the line of Descombes' idea of a possible extension of our vocabulary by means of fictional language, we could say that when no adequate term correctly and directly describes the nature of human attitudes, a well-developed plot can help us understand what is at stake. Consider, for example, Alexis' long letter in Yourcenar's short story, "Alexis": the character dedicates several long pages to barely succeed, in the end, to express to his wife why their marriage has failed. So, novels do not only aim at naming moral and psychological motivations that could actually be ours; they especially try to design languages that are subtler than the ordinary, and the truths they convey are then in straight connection to them.

(iv) **Links to the Possible.** The fourth and final feature certainly is the most important, for it actually overhangs the others. Let us say right away that the very specificity of the truths that fictions express is linked to the modal and deontic notion of possibility. Truths have to be thought of as dealing with (1) ontological possibilities and (2) interpretative possibilities.

(1) On the one hand, novels seem to present us with various forms of what is possible. This is the case of various utopias and thought experiments and Musil's famous novel is paradigmatic in this respect. The appropriate truths, then, concern what has not happened yet, what could happen, what will perhaps happen, that is, they bear on what does not actually exist. This is why we bring into consideration the concept of verisimilitude which, in a sense, is that of what is possibly true in the real world and really true in a possible fictional universe. This kind of truth cannot be eliminated because it expresses possibilities that have never obtained. They come from the prospective dimension of fictions, inventing and imagining human possibilities of existence. According to Schaeffer

---

<sup>25</sup> We agree with Pouillon (1946, p. 94) that the type is not so much *what we understand* as *what makes us understand*, in the sense that, for instance, we can start to see the real world through Balzac's eyes and find Balzacian characters everywhere.

(2008), a fiction thereby is the “virtual exemplification of a possible being-in-the-world” (p. 80).

Moral possibilities relate to a similar but deontic notion, as both of them seem to be attempts to improve our moral thoughts and choices; the question is “how to live a good life?” According to Bouveresse (2008, p. 63), the purpose of novelistic inquiry precisely is that sort of practical truth, and unlike scientific understanding, the kind of understanding that works of fiction may yield is not propositional and directly concerns the question how we *may* or *must* live.

(2) On the other hand, novels are dealing with the multiplicity of the possible meanings of what there is. Possible interpretations and possible conceptions of the world express this modal hermeneutical kind of truth. This closely relates to Putnam’s idea that novels teach us “to see the world as it looks to someone who is sure that hypothesis is correct” (Putnam, 1978, p. 90). Novels, then, do propose something like a hypothesis, which they take for granted, developing it through the story, and modifying it according to the changes that affect their characters.

Putnam claims that novels direct our attention to new interpretations of old facts and construct a genuine kind of knowledge, “knowledge of a possibility”, “conceptual knowledge” yielded by the writer’s intuition—to be more precise, it is the conjunction of the writer’s intuition and the reader’s reception. A problem for this view is that intuition is not knowledge, at least if knowledge relies on scientifically checked verification. Novels, then, could not express any truth at all.

This problem leads us to prefer another definition of what Putnam calls “knowledge”. Putnam’s claim is correct, but only if by “knowing”, we mean “understanding”. The reason for this is that the intentional use of fictional language can yield a certain understanding of the meaning of life, but one that cannot, by any means, be confirmed by any fact.<sup>26</sup>

So, the truths of fictions are of a general level. They concern real human possibilities as well as possibilities of various interpretations. Both kinds of possibilities are plural, and this makes the truths in question irreducible. We thus preserve the applicability of fictional truths without thereby assuming a directly referential connection with reality. Being true,

---

<sup>26</sup> Fictions can also be thought to improve our modal faculties: they are a way for us to test our modal judgments about what is possible and what seems impossible, in this or that situation.

here, does not mean that life finds accurate representation in the fictional frame, but rather that life itself enters the novel, as Bakhtine (1978) would put it. The truth criterion could then be formulated as follows: Does the work of fiction remain outside life and beside reality, a reality that it is yet supposed to take into consideration?<sup>27</sup>

### *Additional Remarks*

What counts for now is the idea that fictions, because they are practically indispensable, can tell us something true about human beings and the world where they live, think, and feel; they say something true about the meaning of the world and the multiple interpretations we may propose. And all this depends on an intentional use of language that is undoubtedly different from its descriptive use.

An obvious advantage of the resulting view is that the question of truth in fiction and the question of the existence of fictional entities are kept apart. We no longer have to assume that the truths fictions express have anything to do with the correspondence scheme. Those truths proceed from the indispensable role that fictions play, not from a direct reference—necessarily failing anyway—to real empirical facts. The kind of understanding one may get from reading a novel does not only concern what exists, but rather what could possibly exist if this or that happened in this or that situation according to this or that modality, and that understanding bears as well on the possible meanings we could give to what exists. Incidentally, our view entails that fictions that are useless in that they do not improve the reader's interpretative ability in any way are definitely not true—although finding fictions that are absolutely useless in this sense might reveal a difficult task, as even pulp literature might help us understand better certain aspects of things.

A requirement as well as a consequence of our proposal is that we cannot strictly isolate fictional worlds from the actual world any longer. As soon as we claim that fictions are useful to understand real phenomena, we must admit that there is direct contact between these two poles. Fictional worlds do not, however, superimpose possible worlds, for we can think of examples of impossible fictional universes—in fantastic literature, for

---

<sup>27</sup> See also Woolf (1927/1963).

instance. Possible worlds simply allow us to put fictional worlds and the real world together into the same conceptual bag, without involving any kind of equivalence for all that.

## Conclusion

Roughly said, Frege, Russell and Quine rejected fictions as being non-serious, hence unconcerned with truth. However, fictions play a central and particular role in human lives: this is the main reason why we cannot be satisfied with their approaches. We argued that fictional narratives should not be considered descriptive: what one can learn from them irreducibly is nothing but ways of interpreting others and the world. So fictional narratives should not be treated like descriptive or theoretical discourse. However, theoretical discourse *is* the proper target of Quine's criterion of ontological commitment, as well as that of the original indispensability argument.

We proposed to clearly detach two aspects of this argument: (i) the entailment from indispensability to truth, and (ii) the entailment from truth to ontology. The second aspect corresponds to Quine's ontological criterion. If the argument is to transpose from theoretical discourse to fictional discourse, then one may retain (i) but must abandon (ii).

The entailment from truth to ontology, (ii), is specific to theoretical discourse, if not to first-order formalized theories. The development of NL semantics since the 1970s reveals how insufficient, hence inadequate, first-order logic is when it comes to accounting for a variety of linguistic phenomena. Fine-grained analyses, especially those provided since the dynamic turn in NL semantics, emphasized the need to explain how people interpret discourse in context, rather than static descriptions of the meanings of expressions in isolated eternal sentences. In fulfilling this task, semanticists promoted two-level theories that add a kind of "internal level" to the usual model-theoretic one. The way in which linguistic interpretation is handled is thus rendered relatively independent from the representational level, that is, from the external connection between language and world. Obviously, (ii) concerns the representational dimension of semantics, and not at all the procedural one. At the internal level, one

can account for some kind of truth of a discourse, with no relation to the world, hence no ontological commitment.

Nevertheless, we are only halfway through. The relevant notion of truth is indeed a weak one: it is internal and cannot be considered to be serious. Assuming (i), we argued that fiction, being indispensable, is true in a strong sense, totally disconnected from the idea that fictional narratives have a descriptive dimension. Truth in the strong sense comes from some appropriate connection to the world, one that is strictly different from representation. Because a fictional narrative is not a description, the correspondence theory of truth is clearly irrelevant. We used analyses drawn from literary criticism and the philosophy of art to account for the specificity of fictional truths. Truth in the relevant sense comes at a global level and is connected to modality: because fictions offer many interpretations, their truth is that of the interpretative possibilities they let us explore. In a way, this can be conceived of as being in relative accordance with Lewis' and many philosophers' idea that the semantics of fictions can be accounted for in a possible-world setting.<sup>28</sup>

### *Literary Works Mentioned*

Balzac, H. de, *The Human Comedy*. Barnes, J., *Talking It Over*. Broch, H., *The Sleepwalkers*. Calvino, I., *Mr Palomar*. Coetzee, J. M., *Waiting for the Barbarians*. Durrell, L., *The Alexandria Quartet*. Eschyle, *The Persians*. Flaubert, G., *Salamambo*. Gary, R., *Les racines du ciel*. Grass, G., *The Tin Drum*. Kristof, A., *The Notebook; The Proof; The Third Lie: Three Novels*. Musil, R., *The Man without Qualities*. Singer, I. B., *Gimpel the Fool and Other Stories*. Stridsberg, S., *La Faculté des rêves*. Yourcenar, M., *Alexis*. Zola, E., *The Rougon-Macquart*.

---

<sup>28</sup> We would like to thank Jacques Morizot and Roger Pouivet for their comments on an earlier version of this paper, and the editor of this volume for his invitation to contribute and useful suggestions.

# References

- Aarseth, E. (1997). *Cybertext: Perspectives on ergodic literature*. Baltimore: Johns Hopkins University Press.
- (2005). Doors and perception: Fiction vs. simulation in games. Paper presented at the *Conference on Digital Arts and Culture*; Copenhagen, Denmark.
- Adams, F., & Dietrich, L. (2003). Empty names, natural kind terms, and radical externalism. *Proceedings of the 2003 Hawaii International Conference on Arts and Humanities*; Honolulu, Hawaii, January 12–15, 2003.
- (2004). What's in a(n empty) name? *Pacific Philosophical Quarterly*, 85, 125–148.
- Adams, F., & Fuller, G. (1992). Names, contents, and causes. *Mind & Language*, 7, 205–221.
- (2007). Empty names and pragmatic implicatures. *Canadian Journal of Philosophy*, 37, 449–461.
- Adams, F., Fuller, G., & Stecker, R. (1992). The Semantics of thoughts. *Pacific Philosophical Quarterly*, 73, 375–389.
- (1993). Thoughts without objects. *Mind & Language*, 8, 90–104.
- (1997). The Semantics of fictional names. *Pacific Philosophical Quarterly*, 78, 128–148.
- (1999). Object-dependent thoughts, perspectival thoughts, and psychological generalizations. *Dialectica*, 53, 47–59.
- Adams, F., & Stecker, R. (1994). Vacuous singular terms. *Mind & Language*, 9, 387–401.
- Anderson, A. R., Belnap, N. D., & Dunn, J. M. (1975). *Entailment: The logic of relevance and necessity*, Vol. I. Princeton: Princeton University Press.

- Anderson, A. R., Belnap, N. D., & Dunn, J. M. (1992). *Entailment: The logic of relevance and necessity*, Vol. II. Princeton: Princeton University Press.
- Asher, N., (1993). *Reference to abstract objects in discourse*. Dordrecht: Kluwer Academic.
- Bach, K. (2002). Giorgione was so-called because of his name. *Philosophical Perspectives*, 16, 73–103.
- Bakhtine, M. (1978). *Esthétique et théorie du roman* (D. Olivier, French Trans.). Paris: Gallimard.
- Balzac, H. de (1965). Avant-propos à *La comédie humaine* (P. Citron, Ed.). Paris: Seuil. (Original work 1842)
- Barbero, C. (2005). Madame Bovary: Something like a melody. Milan: Edizioni Albo Versorio.
- Barboza, D. (2005, December 9). Ogre to slay? Outsource it to Chinese. *New York Times*, Technology section.
- Barthes, R. (1978). Longtemps je me suis couché de bonne heure. In *Le bruissement de la langue: Essais critiques IV*. Paris: Seuil.
- Barwise, J. (1979). On branching quantifiers in English. *Journal of Philosophical Logic*, 8, 47–80.
- Baudelaire, C. (1976). Notes sur Edgar Poe. In *Oeuvres complètes*, Vol. II (C. Pichois, Ed.). Paris: Gallimard. (Original work published 1857)
- Beaver, D. I. (1997). Presupposition. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 939–1008). Amsterdam: Elsevier.
- Bencivenga, E. (1994). Free logics. In D. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic*, Vol. 3 (pp. 373–426). Dordrecht: Kluwer Academic.
- Benjamin, W. (1991). Le narrateur: Réflexions à propos de l'oeuvre de Nicolas Leskov. In *Écrits français* (pp. 264–298). Paris: Gallimard. (Original work 1936).
- Bloom, H. (1994). *The Western canon: The Books and School of the ages*. New York: Harcourt Brace.
- Bonomi, A. (2008). Fictional contexts. In P. Bouquet, L. Serafini & R. H. Thomason (Eds.), *Perspectives on Contexts* (pp. 215–250). Stanford: CSLI Publications.
- Bouveresse, J. (2008). *La connaissance de l'écrivain: Sur la littérature, la vérité et la vie*. Marseille: Agone.



- Bratman, M. E. (1992). Practical reasoning and acceptance in a context. *Mind*, 101, 1–16.
- Braun, D. (1993). Empty names. *Noûs*, 27, 449–469.
- (2005). Empty names, fictional names, mythical names. *Noûs*, 39, 596–631.
- Brock, S. (2002). Fictionalism about fictional characters. *Noûs*, 36, 1–21.
- (2010). The creationist fiction: The case against creationism about fictional characters. *Philosophical Review*, 119, 337–364.
- Calvino, I. (2006). Interview by P. Fournel. *Magazine Littéraire*, 220, 78–79. (Original interview 1985)
- Caplan, B. (2004). Creatures of fiction, myth, and imagination. *American Philosophical Quarterly*, 41, 331–337.
- Carnap, R., (1950). Empiricism, semantics and ontology. *Revue Internationale de Philosophie*, 4, 20–40.
- Carroll, N. (1990). *The philosophy of horror or paradoxes of the heart*. New York: Routledge.
- (1998). *A philosophy of mass art*. Oxford: Clarendon Press.
- Castronova, E. (2005). *Synthetic worlds: The business and culture of online games*. Chicago, IL: Chicago University Press.
- Cohen, L. J. (1992). *An essay on belief and acceptance*. Oxford: Clarendon Press.
- Colyvan, M. (2008). Indispensability arguments in the philosophy of mathematics. *Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/entries/mathphil-indis/>
- Cooper, R., Crouch, R., van Eijck, J., Fox, C., van Genabith, J., Jaspers, J., Kamp, H., Pinkal, M., Poesio, M., Pulman, S., & Vestre, E. (1994). Describing the Approaches. *FraCaS Deliverable*, D8. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.3456>
- Crimmins, M. (1998). Hesperus and Phosporus: Sense, pretense, and reference. *Philosophical Review*, 107, 1–47.
- Currie, G. (1990). *The nature of fiction*. New York: Cambridge University Press.
- Darsel, S., & Pouivet, R. (Eds.). (2008). *Ce que l'art nous apprend*. Rennes: Presses Universitaires de Rennes.
- Davies, M., & Humberstone, L. (1980). Two notions of necessity. *Philosophical Studies*, 58, 1–30.

- Davies, S. (2009). Responding emotionally to fictions. *Journal of Aesthetics and Art Criticism*, 67, 269–284.
- Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. New York: Viking.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Descombes, V. (1987). *Proust: Philosophie du roman*. Paris: Les Éditions de Minuit.
- (1995). *La denrée mentale*. Paris: Les Éditions de Minuit.
- Deutsch, H. (1991). The Creation Problem. *Topoi*, 10, 209–225.
- (2000). Making up stories. In A. Everett & T. Hofweber (Eds.), *Empty names, fiction and the puzzles of non-existence* (pp. 149–181). Stanford: CSLI Publications.
- Donnellan, K. (1966). Reference and definite descriptions. *The Philosophical Review*, 75, 281–304.
- (1974). Speaking of nothing, *The Philosophical Review*, 83, 3–31.
- Evans, G. (1978). Can there be vague objects? *Analysis*, 38, 208.
- (1982). *The varieties of reference*. Oxford: Clarendon Press.
- Everett, A. (2000). Referentialism and empty names. In Everett & Hofweber (2000), 37–60.
- (2005). Against fictional realism. *Journal of Philosophy*, 102, 624–649.
- Everett, A. & Hofweber, T. (Eds.). (2000). *Empty names, fiction and the puzzles of non-existence*. Stanford: CSLI Publications.
- Fine, K. (1982). The problem of non-existents, I: Internalism. *Topoi*, 1, 97–140.
- Fontaine, N., & Rahman, S. (2010). Fiction, creation and fictionality. *Methodos*, 10. Retrieved from <http://methodos.revues.org/2343>.
- Fraisse, E., & Mouralis, B. (2001). *Questions générales de littérature*. Paris: Seuil.
- Frege, G. (1983). Dialog mit Pünjer Über Existenz. In H. Hermes, F. Kambartel, & F. Kaulbach (Eds.), *Frege, Gottlob: Nachgelassene Schriften und Wissenschaftlicher Briefwechsel*, Vol. 1 (pp. 60–75). Hamburg: Meiner Verlag.
- Friend, S. (2007). Fictional characters. *Philosophy Compass*, 2, 141–156.
- Furetière, A. (2001). *Le roman bourgeois*. Paris: Flammarion. (Original work 1666)

- Gendler, T. S., & Hawthorne, J. (Eds.). (2002). *Conceivability and possibility*. Oxford: Clarendon Press.
- Genette, G. (1969). *Figures II*. Paris: Seuil.
- (1991). *Fiction et diction*. Paris: Seuil.
- Gochet, P. (2010). La théorie de l'objet de Meinong à la lumière de la logique actuelle. In P. Bour et al. (Eds.), *Construction*. London: College Publications.
- Goodman, J. (2003). Where is Sherlock Holmes? *Southern Journal of Philosophy*, 41, 183–198.
- Goodman, N. (1978). *Ways of worldmaking*. Indianapolis: Hackett Publishing Company.
- Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy*, 14, 39–100.
- (2000). Meaning in motion. In K. von Heusinger & U. Egli (Eds.), *Reference and anaphoric relations* (pp. 47–76). Dordrecht: Kluwer Academic.
- Hale, B. (2007). Into the abyss: A critical study towards non-being. *Philosophia Mathematica*, 15, 94–110.
- Heim, M. (1991). The metaphysics of virtual reality. In S. K. Hessel, & J. P. Roth (Eds.), *Virtual reality: Theory, practice, and promise* (pp. 27–34). Westport: Meckler.
- Hintikka, J. (1988). On the development of the model-theoretic viewpoint in logical theory. *Synthese*, 77, 1–36.
- (1997). No scope for scope? *Linguistics and Philosophy*, 20, 515–544.
- Hintikka, J., & Sandu, G.. (1997). Game-theoretical semantics. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 361–410). Amsterdam: Elsevier.
- Hofweber, T. (2000). Quantification and non-existent objects. In Everett & Hofweber (2000), 249–273.
- Howell, R. (1979). Fictional objects: How they are and how they aren't. *Poetics*, 8, 129–177.
- (1983). Review of T. Parsons' *Nonexistent objects*. *Journal of Philosophy*, 80, 163–173.
- (1996). Essay-review of K. Walton's *Mimesis as Make-Believe*. *Synthese*, 109, 413–434.

- (1998). Fiction, semantics of. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*, Vol. 3. New York: Routledge.
- (2005). La finzione e i suoi oggetti. *Discipline Filosofiche*, 15, 161–179.
- (2010). Literary fictions, real and unreal. In Woods (2010).
- Hugo, V. (1967). *Les misérables*, Vol. 3 (R. Journet, Ed.). Paris: Flammarion. (Original work published 1862).
- Jackson, F. (1998). *From metaphysics to ethics: A defense of conceptual analysis*. Oxford: Clarendon Press.
- Kalderon, M. E. (Ed.). (2005). *Fictionalism in metaphysics*. Oxford: Clarendon Press.
- Kamp, H. (1981). A theory of truth and semantic representation. In M. Stokhof, J. Groenendijk, & T. Janssen (Eds.), *Formal methods in the study of language* (pp. 277–322). Amsterdam: Mathematisch Centrum.
- Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Dordrecht: Kluwer Academic.
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–564). New York: Oxford University Press.
- Kerlow, I. V. (2000). *The art of 3-D computer animation and imaging* (2nd ed.). New York: John Wiley & Sons.
- Kripke, S. (1973), Reference and Existence. *John Locke Lectures*. (Transcript available in the Philosophy Library, Oxford University).
- (1976). Is there a problem about substitutional quantification? In G. Evans & J. McDowell (Eds.), *Truth and meaning: Essays in semantics* (pp. 325–419). Oxford: Clarendon Press.
- (1977). Speaker's reference and semantic reference. In P. A. French, T. E. Uehling, Jr., & H. K. Wettstein (Eds.), *Studies in the philosophy of language: Midwest Studies in Philosophy*, Vol. 2 (pp. 255–276). Minneapolis: University of Minnesota Press.
- (1980). *Naming and necessity*. Oxford: Basil Blackwell.
- Kroon, F. (2000). "Disavowal through commitment" theories of negative existentials. In Everett & Hofweber (2000), 249–273.
- (2003). Quantified negative existentials. *Dialectica*, 57, 149–164.
- (2004). Descriptivism, pretense, and the Frege-Russell problems. *The Philosophical Review*, 113, 1–30.

- (2005). Belief about nothing in particular. In Kalderon (2005), 178–203.
- Kundera, M. (1995). *L'art du roman*. Paris: Gallimard. (Original work published 1986)
- Lamarque, P. (1996). *Fictional points of view*. New York: Cornell University Press.
- Lamarque, P., & Olsen, S. H. (1994). *Truth, fiction and literature*. Oxford: Clarendon Press.
- Lambert, K. (2003). *Free logic: Selected essays*. New York: Cambridge University Press.
- Lewis, D. K. (1973). *Counterfactuals*. Cambridge: Harvard University Press.
- (1978). Truth in fiction. *American Philosophical Quarterly*, 15, 37–46. Reprinted in *Philosophical papers* (Vol. 1, 1983, pp. 261–280). New York: Oxford University Press. In-text page references are to the reprint.
- (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339–359. Reprinted in R. Bäuerle, U. Egli & A. von Stechow (Eds.), *Semantics from different points of view* (pp. 172–187). Berlin: Springer-Verlag.
- (1986). *On the plurality of worlds*. Oxford: Blackwell.
- Linsky, B., & Zalta, E. N. (1991). Is Lewis a Meinongian? *Australasian Journal of Philosophy*, 69, 438–453.
- (1994). In defense of the simplest quantified modal logic. In J. Tomberlin (Ed.), *Philosophical Perspectives*, Vol. 8: *Logic and language* (pp. 431–458). Atascadero: Ridgeview Press.
- Lopes, D. M. (2001). The ontology of interactive art. *Journal of Aesthetic Education*, 35, 65–81.
- (2004). Digital art. In L. Floridi (Ed.), *The Blackwell guide to the philosophy of computing and information* (pp. 106–116). Oxford: Blackwell.
- (2009). *A philosophy of computer art*. London: Routledge.
- Loux, M. J., & Zimmerman, D. W. (Eds.). (2003). *The Oxford handbook of metaphysics*. New York: Oxford University Press.
- Macdonald, M. (1954). The language of fiction. In *Proceedings of the Aristotelian Society, Supplementary Volumes*, 28, 165–196.

- Manning, S. B. (1971). *Dickens as satirist*. New Haven: Yale University Press.
- Martinich, A. P., & Stroll, A. (2007). *Much ado about nonexistence: Fiction and reference*. New York: Rowman & Littlefield.
- Meinong, A. (1960). On the theory of objects. In R. M. Chisholm (Ed.), *Realism and the Background of Phenomenology* (pp. 76–117). Glencoe: Free Press.
- Minsky, M. (1980, June). Telepresence. *OMNI Magazine*, 2, 45–51.
- Molino, J., & Rafhail-Molino, R. (2003). *Homo fabulator: Théorie et analyse du récit*. Montréal: Léméac/Actes Sud.
- Montalbetti, C. (2001). *La fiction*. Paris: Flammarion.
- Morizot, J. (2008). Enjeu cognitif et/ou théorie cognitive. In S. Darsel & R. Pouivet (Eds.), *Ce que l'art nous apprend* (pp. 49–66). Rennes: Presses Universitaires de Rennes.
- Murray, J. (1998). *Hamlet on the Holodeck*. Cambridge, Mass.: MIT Press.
- Muskens, R. (1995). *Meaning and partiality*. Stanford: CSLI Publications.
- Nagera, H. (1969). The imaginary companion: Its significance for Ego development and conflict resolution. *The Psychoanalytic Study of the Child*, 24, 165–196.
- Newman, J. (2004). *Videogames*. New York: Routledge.
- Nussbaum, M. (1995). *Poetic justice: The literary imagination and public life*. Boston: Beacon Press.
- Parsons, T. (1980). *Nonexistent objects*. New Haven: Yale University Press.
- Pashler, H. E. (1998). *The psychology of attention*. Cambridge, Mass.: MIT Press.
- Pavel, T. (2003). *La pensée du roman*. Paris: Gallimard.
- Peregrin, J. (2000). Reference and inference: The case of anaphora. In K. von Heusinger & U. Egli (Eds.), *Reference and Anaphoric Relations* (pp. 269–286). Dordrecht: Kluwer Academic.
- Plantinga, A. (1979). De essentia. *Grazer Philosophische Studien*, 7/8, 101–121.
- Pouillon, J. (1946). *Temps et roman*. Paris: Gallimard.
- Pouivet, R. (2008). Esthétique modale. In Darsel & Pouivet (2008), 35–48.

- Predelli, S. (1997). Talk about fiction. *Erkenntnis*, 46, 69–77.
- (2002). “Holmes” and Holmes: A Millian analysis of names from fiction. *Dialectica*, 56, 261–279.
- (2005). *Contexts*. Oxford: Clarendon Press.
- (2008). Modal monsters and talk about fiction. *Journal of Philosophical Logic*, 37, 277–297.
- Priest, G. (1997). Sylvan’s box: A short story and ten morals. *Notre Dame Journal of Formal Logic*, 38, 573–582. Reprinted as section 6.6 of Priest (2005).
- (2001). *An introduction to non-classical logic*. Cambridge: Cambridge University Press.
- (2005). *Towards non-being: The logic and metaphysics of intentionality*. Oxford: Clarendon Press.
- (201+). Against against non-being. *Review of Symbolic Logic*, forthcoming.
- Putnam, H. (1978). Literature, science and reflection. In *Meaning and the moral sciences* (Part 2). London; Boston: Routledge & K. Paul.
- (1979a). What is mathematical truth. In Putnam (1979c), 60–78.
- (1979b). Philosophy of logic. In Putnam (1979c), 323–357.
- (1979c). *Philosophical papers*, Vol. 1: *Mathematics, matter and method* (2nd ed.). Cambridge: Cambridge University Press.
- Quine, W. V. O. (1960). *Word and object*. Cambridge: MIT Press.
- (1961). On what there is. In *From A Logical Point of View* (pp. 1–19), New York: Harper & Row. (Original work published 1948)
- (1971). Quantifiers and propositional attitudes. In L. Linsky (Ed.), *Reference and modality* (pp. 101–111). London: Oxford University Press.
- (1976). Carnap and logical truth. In *The ways of paradox and other essays* (revised ed., pp. 107–132). Cambridge: Harvard University Press. (Original work published 1960)
- (1981). *Theories and things*. Cambridge: Harvard University Press.
- Rast, E. (2007). *Reference and indexicality*. Berlin: Logos Verlag.
- Recanati, F. (1996). Domains of discourse. *Linguistics and Philosophy*, 19, 445–475.
- (2000). *Oratio obliqua, oratio recta: An essay on metarepresentation*. Cambridge: MIT Press.

- Reimer, M. (1998). Quantification and context. *Linguistics and Philosophy*, 21, 96–115.
- Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy*, 20, 335–397.
- Rivière, J. (1999). Le roman d'aventure. In *Études: L'oeuvre critique de Jacques Rivière à la Nouvelle Revue Française, 1909–1924* (pp. 344–347). Paris: Gallimard. (Original work published 1913)
- Routley, R. (1980). *Exploring Meinong's jungle and beyond: An investigation of noneism and the theory of items*. Canberra: Research School of Social Sciences, ANU.
- Rozakis, D. (2009). *Qu'est-ce qu'un roman?* Paris: Vrin.
- Russell, B. (1905). On denoting. *Mind*, 14, 479–493.
- (1959). *Problems of philosophy*. New York: Oxford University Press. (Original work published 1912)
- Ryle, G. (1931–1932). Systematically misleading expressions. *Proceedings of the Aristotelian Society*, 32, 139–170.
- Sade, D. A. F., marquis de (1970). *Idée sur les romans* (J. Glastier, Ed.). Bordeaux: Ducros. (Original work 1799)
- Sainsbury, R. M. (2005). *Reference without referents*. Oxford: Clarendon Press.
- (2009). *Fiction and Fictionalism*. London: Routledge.
- Salmon, N. (1981). *Reference and essence*. Princeton: Princeton University Press.
- (1998). Nonexistence. *Noûs*, 32, 277–319. In Salmon (2005), 50–90.
- (2002). Identity Facts. *Philosophical Topics*, 30, 237–267. In Salmon (2005), 165–191.
- (2005). *Metaphysics, mathematics, and meaning*. Oxford: Clarendon Press.
- Santambrogio, M. (1992). *Forma e oggetto*. Milan: Il Saggiatore.
- Schaeffer, J. M. (2008). Quelles valeurs cognitives pour quels arts? In Darsel & Pouivet (2008), 67–81.
- Schiffer, S. (1996) Language-created language-independent entities. *Philosophical Topics*, 24, 149–167.
- (2003). *The things we mean*. Oxford: Clarendon Press.



- Schlenker, P. (2006). Scopal independence: A note on branching and island-escaping readings of indefinites and disjunctions. *Journal of Semantics*, 23, 281–314.
- Schnieder, B., & von Solodkoff, T. (2009). In defence of fictional realism. *Philosophical Quarterly*, 59, 138–149.
- Searle, J. (1979). The logical status of fictional discourse. In *Expression and meaning* (1979, pp. 58–75). Cambridge: Cambridge University Press. (Original work published 1974–1975)
- Sinowjew, A. A. (1970). *Komplexe Logik*. Berlin: Deutscher Verlag der Wissenschaften.
- Sinowjew, A. A., & Wessel, H. (1975). *Logische Sprachregeln*, Deutscher Verlag der Wissenschaften. Berlin: Deutscher Verlag der Wissenschaften.
- Smith, N. J. J. (2008). Why sense cannot be made of vague identity. *Noûs*, 42, 1–16.
- Soames, S. (2002). *Beyond rigidity: The unfinished semantic agenda of Naming and necessity*. New York: Oxford University Press.
- Stalnaker, R. (1970). Pragmatics. *Synthese*, 22: 31–46.
- (1978). Assertion. In Stalnaker (1999), 78–96.
- (1984). *Inquiry*. Cambridge: MIT Press.
- (1999). *Context and content*. New York: Oxford University Press.
- (2001). On considering a possible world as actual. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 75, 141–156.
- (2002). Common ground. *Linguistics and Philosophy*, 25, 701–721.
- (2003). *Ways a world might be: Metaphysical and anti-metaphysical essays*. New York: Oxford University Press.
- Stanley, J. (2001). Hermeneutic fictionalism. *Midwest Studies in Philosophy*, 25, 36–71.
- Strawson, P. F. (1950). On referring. *Mind*, 59, 320–344.
- Tavinor, G. (2005). Videogames and interactive fiction. *Philosophy and Literature*, 29, 24–40.
- (2009). *The art of videogames*. Malden: Wiley-Blackwell.
- Taylor, M. (2001). *Imaginary companions and the children who create them*. New York: Oxford University Press.
- Taylor, M., Hodges, S. D., & Kohányi, A. (2003). The illusion of independent agency: Do adult fiction writers experience their characters as

- having minds of their own? *Imagination, Cognition and Personality*, 22, 361–380.
- Thomasson, A. L. (1999). *Fiction and metaphysics*. Cambridge: Cambridge University Press.
- (2003a). Fictional characters and literary practices. *British Journal of Aesthetics*, 43, 138–157.
- (2003b). Speaking of fictional characters. *Dialectica*, 57, 205–223.
- (2007). *Ordinary objects*. New York: Oxford University Press.
- (2008). Existence questions. *Philosophical Studies*, 141, 63–78.
- (2010). Fiction, existence and indeterminacy. In Woods (2010).
- van Heijenoort, J. (1967). Logic as language and logic as calculus. *Synthese*, 17, 324–330.
- van Inwagen, P. (1977). Creatures of fiction. *American Philosophical Quarterly*, 14, 299–308.
- (1983). Fiction and metaphysics. *Philosophy and Literature*, 7, 67–77.
- (1990). *Material beings*. Ithaca: Cornell University Press.
- (2003). Existence, ontological commitment, and fictional entities. In Loux & Zimmerman (2003), 131–157.
- Vico, G. B. (1948), *Scienza nuova seconda* [*The New Science of Giambattista Vico*] (3rd ed., T. G. Bergin & M. H. Fisch, Trans.). Ithaca: Cornell University Press. (Original work 1730–1744)
- Voltolini, A. (2006). *How ficta follow fiction: A syncretistic account of fictional entities*. Dordrecht: Springer.
- (2009). The seven consequences of creationism. *Metaphysica*, 10, 27–48.
- Walton, K. L. (1978). Fearing fictions. *Journal of Philosophy*, 75, 5–27.
- (1990). *Mimesis as make-believe*. Cambridge: Harvard University Press.
- (1993). Metaphor and prop-oriented make-believe. *European Journal of Philosophy*, 1, 39–57.
- (1997). Spelunking, simulation, and slime. In M. Hjort & S. Laver (Eds.), *Emotion and the arts* (pp. 37–49). New York: Oxford University Press.
- (2000). Existence as metaphor? In Everett & Hofweber (2000, pp. 69–94).

- (2003). Restricted quantification, negative existentials, and fiction. *Dialectica*, 57, 239–242.
- Wertheim, M. (1999). *The pearly gates of cyberspace*. London: Virago Press.
- Williamson, T. (2002). Vagueness, identity, and Leibniz's law. In A. Bottani, M. Carrara, & P. Giaretta (Eds.), *Individuals, Essence and Identity* (pp. 273–304). Dordrecht: Kluwer Academic.
- (2003). Vagueness in reality. In Loux & Zimmerman (2003, pp. 690–715).
- Winter, Y. (2004). Functional quantification. *Research on Language and Computation*, 2, 331–363.
- Wittgenstein, L. (1961). *Tractatus logico-philosophicus*. (P. Klossowski, French Trans.). Paris: Gallimard. (Original work published 1921)
- (1978). *Remarks on the Foundations of Mathematics*. Oxford: Blackwell.
- Wolterstorff, N. (1980). *Works and worlds of art*. New York: Oxford University Press.
- Woods, J. (1974). *The logic of fiction: A philosophical sounding of deviant logic*. The Hague: Mouton.
- Woods, J. (Ed.). (2010). *Fictions and models: New essays*. Munich: Philosophia Verlag.
- Woods, J., & Alward, P. (2002). The Logic of fiction. In D. Gabbay & F. Guenther (Eds.), *Handbook of Philosophical Logic*, Vol. 11 (2nd ed., pp. 241–316). Dordrecht, Kluwer Academic.
- Wolf, V. (1963). *L'art du roman* (R. Celli, French Trans.). Paris: Seuil. (Original work published 1927).
- Woolley, B. (1992). *Virtual worlds*. Oxford: Blackwell.
- Yablo, S. (1998). Does ontology rest on a mistake? *Proceedings of the Aristotelian Society, Supplementary Volumes*, 72, 229–261.
- (2001). Go figure: A path through fictionalism. *Midwest Studies in Philosophy*, 25, 72–102.
- (2005). The myth of the seven. In Kalderon (2005), 88–115.
- Yagisawa, T. (2001). Against creationism in fiction. *Philosophical Perspectives*, 15, 153–172.
- Zalta, E. N. (1983). *Abstract objects*. Dordrecht: Reidel.
- (1988). *Intensional logic and the metaphysics of intentionality*. Cambridge: MIT Press.

- (2000). The road between pretense theory and abstract object theory. In Everett & Hofweber (2000), 117–147.
- (2003). Referring to fictional characters. *Dialectica*, 57, 243–254.
- Zhai, P. (1998). *Get real: A philosophical adventure in virtual reality*. Lanham: Rowman & Littlefield.

# Notes on Contributors

FRED ADAMS is Professor and Chair of Linguistics & Cognitive Science at the University of Delaware. He is also Professor of Philosophy, and Executive Board Member of the program of Science, Ethics, and Public Policy at the University of Delaware. He has published over 100 articles, chapters, and reviews in the areas of action theory, epistemology, philosophy of language, and philosophy of science, including his 2008 book (with Ken Aizawa), *The Bounds of Cognition*.

ROBERT HOWELL is Professor of philosophy at the University at Albany, SUNY (State University of New York at Albany), and a graduate of the University of Michigan (Ph.D.) and Kenyon College (A.B.). He has taught also at the University of Illinois (Urbana), Stanford University, and, as a visitor, at The Johns Hopkins University and Moscow State University. He has published essays on the metaphysics of fiction and of art. In addition, he has written on Kant's theoretical philosophy (*Kant's Transcendental Deduction*, 1992). He hopes eventually to explore the relation of recent accounts of fiction to Kant's work.

FREDERICK KROON teaches Philosophy at the University of Auckland. He has written extensively on realism, rationality, the theory of reference, and the ontology of fiction, and is on the editorial board of the *Australasian Journal of Philosophy* and the *Stanford Encyclopedia of Philosophy*.

FRANCK LIHOREAU is a Postdoctoral Research Fellow for the Portuguese Science and Technology Foundation (FCT) at the Philosophy of Language Institute (IFL), New University of Lisbon. His main research interests

include epistemology, logic and philosophy of language, metaphysics, and the philosophy of fiction. He is the editor of a book on *Knowledge and Questions* (2008).

BRENDAN MURDAY is Assistant Professor of Philosophy at Ithaca College, New York. His main areas of research are metaphysics, philosophy of language, and epistemology.

TERENCE PARSONS is Professor of Philosophy and Linguistics at the University of California at Los Angeles. His research interests in the philosophy of language and metaphysics include theories of non-existent objects and the notion of indeterminate identity, while his work in linguistics concentrates on event-based semantics. His recent work is mainly in the history of logic, especially medieval logic.

GRAHAM PRIEST is Boyce Gibson Professor at the University of Melbourne, a Distinguished Professor at the CUNY Graduate Center, and an Arché Professorial Fellow at the University of St Andrews. His books include: *In Contradiction*, *Beyond the Limits of Thought*, *Introduction to Non-Classical Logic*, *Towards Non-Being*, *Doubt Truth to be a Liar*. His research interests include logic, metaphysics, history of philosophy, and Buddhist philosophy.

ERICH RAST is a Postdoctoral Research Fellow at the Philosophy of Language Institute (IFL), New University of Lisbon. He is the author of a book on *Reference and Indexicality* (Logos, 2007) and has written various articles on topics in the philosophy of language and logic. His current research concerns linguistic context dependence, the semantics-pragmatics interface, theories of reference, formal ontology, and applications of belief and preference upgrade in pragmatics.

MANUEL REBUSCHI is an Associate Professor at the University Nancy 2 and a researcher at the Poincaré Archives (CNRS, UMR 7117). His research interests include philosophical logic, epistemology, the philosophy of language, and the philosophy of mind.

MARION RENAULD is completing a Ph.D. in Philosophy at the Poincaré Archives (CNRS, UMR 7117) at the University Nancy 2. Specializing in the philosophy of fiction, her main research interests include metaphysics, philosophy of language, and aesthetics.

R. M. SAINSBURY is Professor of Philosophy at the University of Texas at Austin. Before that, he was Susan Stebbing Professor at King's College London. He was editor of *Mind* for the last decade of the last century. Other publications include: *Russell* (1979), *Logical Forms* (2000, 2nd ed.), *Paradoxes* (2009, 3rd ed.), *Departing from Frege* (2002), *Reference without Referents* (2005) and *Fiction and Fictionalism* (2009).

GRANT TAVINOR is a Lecturer in Philosophy at Lincoln University in New Zealand. His research interests are in the philosophy of the arts, especially fiction, the arts and morality, the popular arts, and videogames. His book *The Art of Videogames* was published with Wiley-Blackwell in their New Directions in Aesthetics series in 2009.

ALBERTO VOLTOLINI (Ph.D. Scuola Normale Superiore, Pisa, 1989) is a philosopher of language and mind whose work focuses mainly on fiction, intentionality, and Wittgenstein. He is currently an Associate Professor in Philosophy of Mind at the University of Turin. He has got scholarships at the Universities of Geneva and Sussex, been a visiting professor at the Universities of California, Riverside (1998), Auckland and Australian National University, Canberra (2007), and a member of the Steering Committee of the European Society for Analytic Philosophy (2002–2008). He is now a member of the Board of the European Society for Philosophy and Psychology. His publications include *How Ficta Follow Fiction* (Springer, 2006).





# Index

- Aarseth, E., 224–228, 242  
abstract object, 15, 16, 32, 43, 47–50, 56, 64, 71, 83, 84, 101, 117, 128, 155–157, 161, 162, 204–206, 208, 210, 211, 215–217  
acceptance, 20, 21, 132, 134, 137–152, *see also* belief, assumption-relativity  
actualism, 14, 56, 64, 70, 78, 79, 155, 158, 160, 163–165, 176, 177, 179, 188, 202, *see also* possibilism  
actuality, 10, 13–15, 18, 19, 25, 28, 37, 43, 46, 48, 52–63, 65, 67–70, 73–84, 87–89, 105, 109, 111, 114, 117, 120, 123, 127, 128, 130, 134, 155–158, 161–165, 168–179, 182–184, 186–194, 198–201, 204, 231, 236, 238, 270, 284  
Adams, F., 18–20, 153  
ambiguity (of fictional names), 33–34, 85  
Anderson, A. R., Belnap, N. D. & Dunn, J. M., 81  
anti-realism, 86, 88, 104, 105, 107, 108, 113–118, 173, 201, 202, 221, 226, 275, *see also* supervenience  
artefact, 15, 23, 24, 32, 94, 120, 127, 128, 176, 204, 207, 208, 215, 217, 223, 224, 227, 229, 231, 235, 238, 239, 241, 242, *see also* abstract object  
Asher, N., 260  
assertion, 7, 21, 46, 68, 73, 74, 122, 124, 125, 128–131, 139, 152, 158, 159, 173, 177, 179, 193, 195–201, 212  
assumption-relativity, 21–22, 182, 183, 185–197, 199–202  
Bach, K., 81  
Bakhtine, M., 284  
Balzac, H. de, 278  
baptism, 17, 49, 113, 114  
Barbero, C., 103  
Barboza, D., 228  
Barthes, R., 278  
belief, 15, 20, 21, 81, 85, 86, 88, 96, 97, 102, 117, 121, 122, 126, 137–152, 156, 159, 164, 210, 231, 252, 253, *see also* doxastic object  
Benjamin, W., 267  
Bloom, H., 218  
Bonomi, A., 8  
Bouveresse, J., 266, 279, 280, 283  
Bratman, M. E., 140  
Braun, D., 121, 153, 205, 211  
Brock, S., 209, 221, 226  
Calvino, I., 269, 273  
Caplan, B., 203, 211  
Carnap, R., 251  
Carroll, N., 231, 239, 242  
Castronova, E., 224, 228  
characterisation, 111, 116, 118  
characterizing/non-characterizing property, *see* nuclear/extranuclear property  
Cohen, L. J., 150  
Colyvan, M., 263  
constant-/variable-domain semantics, 17, 80, 105, 109, 112, 118  
context, 20, 25, 33, 53, 60–62, 78, 80–83, 85, 90, 91, 100, 101, 103, 104, 140, 145, 147, 150–152, 247, 248, 254, 255, 258–261, 273, 278, 285  
counterfactual, 53–55, 59, 67, 83, 112  
creationism, 15–17, 19, 22, 23, 93, 101–106, 203–213, 215–221, *see also* artefact

- creationist locutions, 15–17, 22, 23, 108, 114, 127, 219
- Crimmins, M., 153, 221
- cross-fiction comparisons, 71–72
- Currie, G., 8, 12, 100, 153, 231
- Davies, S., 241
- de re/de dicto, 99, 100, 104, 130, 156, 159, 175, 183, 186, 193, 194
- Dennett, D. C., 246
- Descombes, V., 246, 268, 271, 282
- description, 13, 14, 19, 50, 52, 53, 55, 59, 60, 80, 81, 85, 86, 100, 111, 114, 122, 123, 125, 128, 132, 142, 154, 189, 190, 226, 229, 230, 237, 247, 252, 254, 255, 258
- descriptive/interpretative use of language, 26, 246, 247, 253, 263–266, 269–275, 278, 281, 284–286
- descriptivism (Russellianism about names), 12, 14, 15, 43, 48, 49, 77, 83, 85, 86, 88, 100, 101, 104, 178, 251, 254, *see also* description
- Deutsch, H., 94, 108
- diaphanous entity, 209–212, 221
- direct reference (Millianism about names), 12, 13, 15, 18, 44, 48–50, 59–61, 85, 99–101, 119–121, 135, 178, 283, 284
- Discourse Representation Theory (DRT), 257–261
- Donnellan, K., 97, 98, 252
- doxastic objects, 85–87
- empty name, 13, 18–20, 43, 45, 46, 49, 50, 63, 67, 68, 71, 97–101, 103, 120–123, 125, 126, 130, 131, 133, 157, 178, 183, 186, 199, 249, 251, 253, 261
- ersatz world, 56, 57, 66
- Evans, G., 98, 164, 171, 252
- Everett, A., 45, 153, 156, 160, 161, 163–167, 170, 172–176, 193, 201
- existence predicate, 14, 29, 30, 79–81, 89, 110, 113
- existentialism (Plantingan), 50–51
- existentially creative/conservative games, 98–102
- fictional operator, 12–14, 43–46, 51, 56, 58, 60, 61, 67, 71, 72, 156, 199, 200, 221, *see also* two-dimensional semantics
- fictionalism, 20, 22, 23, 205–206, 220, 221, *see also* make-believe
- fictive/transfictive/metafictive discourse, *see* textual/paratextual/metatextual use
- Fine, K., 18, 204
- first-order logic, 79, 80, 246, 247, 249, 254–257, 260, 261, 285
- Fraisse, E. & Muralis, B., 274, 277
- free logic, 78, 253
- Frege, G., 60–62, 79, 245, 248, 249, 253, 255, 257, 265, 285
- Generating Principle, 10, 11, 28, 29, 35, 36, 39, 41
- Genette, G., 274, 276
- Goodman, N., 272
- Heim, M., 224, 225, 229
- Hintikka, J., 249, 257
- Hofweber, T., 254, 259
- Howell, R., 12, 18, 21–22, 205
- identity, 11, 17, 18, 28, 29, 31, 36–38, 40, 68, 73–76, 78, 82, 86, 87, 112, 115–116, 130, 155, 161–165, 167–176, 201, 209, *see also* indeterminacy, indiscernibility, individuation
- imaginary companion, 213–221
- impossible fiction, 63, 66, 67, 72
- impossible object, 28, 34, 64
- impossible world, 66, 67, 89, 91, 109, 111, 118, 284
- incomplete fiction, 63, 65–67, 69–70, 72, 95, 109, 272, 273, 281
- incomplete object, 28, 30, 34, 36, 38, 39, 80, 81, 168, 175, 188

- incomplete proposition, 13, 19, 121–124, 126, 134
- inconsistency, 31, 63, 65–67, 109, 157, 163, 165–167, 175, 176, 188, 201
- indeterminacy, 11, 12, 18, 27, 34–41, 65, 66, 68, 69, 156, 161, 163–166, 169–175, 188, 201, 229
- indiscernibility, 17, 28, 36, 115
- indispensability, 25–26, 245–247, 263–267, 269, 275, 285, 286
- individuation, 28, 36, 40, 68
- interactivity, 24, 224–227, 230, 232–244
- irrealism, 9–10, 18–22, 138, 147–149, *see also* fictionalism
- Kamp, H., 257
- Kaplan, D., 14, 53, 55, 80, 90, 100, 189, 190
- Kripke, S., 49, 64, 81, 84, 85, 93, 96–98, 102–104, 125, 153, 177, 188–190, 203, 205, 210, 211, 252, 253
- Kroon, F., 18, 22–23, 153, 226
- Kundera, M., 268, 270
- Lamarque, P., 231, 242, 266, 276
- Leibniz, G. W., 17, 28, 115
- Lewis, D. K., 12, 13, 56, 63, 65, 78, 79, 82, 83, 87, 89, 104–105, 141, 188, 286
- Linsky, B. & Zalta, E. N., 79, 89
- literalism, 18, 20, 22, 204
- Lopes, D. M., 239, 240
- make-believe (pretense), 16, 22, 23, 64, 65, 96–103, 128, 153, 154, 157–161, 165, 176, 177, 179, 185–187, 189, 191, 194–202, 205, 208, 209, 220, 221, 228, 231, 242
- Martinich, A. P., & Stroll, A., 19, 127, 131–135
- Meinong, A., 28–29, 77, 79, 226
- Meinongianism, 10–11, 17, 27–29, 89, 128, 157, 204–205, *see also* noneism
- Millian theory of names, *see* direct reference
- Morizot, J., 272
- Murday, B., 10, 13–14
- Murray, J., 225
- Nagera, H., 213
- native/immigrant character, 30–32, 41
- negation, 15, 30, 44, 81, 109–110, 112
- negative existential, 10, 19, 72–76, 78, 120, 122, 124–127, 147–148
- Newman, J., 235
- non-existent objects, *see* Meinongianism
- non-standard predication, 80–82
- noneism, 17, 89, 107–109
- nuclear/extranuclear property, 10–11, 29–31, 37, 38, 204, 205
- ontological commitment, 25, 28, 50, 88, 154–156, 160, 193, 206, 207, 212, 245, 247–249, 251, 255, 260–264, 285, 286
- Parsons, T., 10–12, 89, 105, 128, 155, 204
- Peregrin, J., 258, 259
- Plantinga, A., 51, 189, 190
- possibilism, 14, 70, 77–79, 87–89
- Pouillon, J., 282
- Pouivet, R., 268
- Predelli, S., 14, 100, 101, 205
- presupposition, 20, 21, 137, 144, 146, 147, 151, 221, 257, 258, 261
- pretending/characterizing/hypostatizing use, 16–17, 93, 102–104
- pretense, *see* make-believe
- Priest, G., 10, 11, 17–18, 66, 77–79, 81, 89, 91, 105, 204
- prop, 22–24, 231, 233, 240–244
- propositional attitude, 20, 60–62, 81, 85, 137–142, 144
- proxy, 224, 236–238, 242
- Putnam, H., 25, 245, 263, 283
- quantification, 23, 79, 80, 108, 110, 112–114, 117, 156, 176, 187, 195, 206, 207, 209–212, 214, 215, 247, 251, 253–257, 259, 260, 263
- Quine, W. V. O., 25, 80, 99, 245–251, 253, 255, 261–263, 269, 285

- Rast, E., 10, 14–15  
 realism, 9–10, 154–157, 160, 163–165, 176, 179, 187–191, 202, *see also* Meinongianism, creationism  
 Rebuschi, M. & Renauld, M., 24–26  
 Recanati, F., 100, 253  
 reductionism, 87–88  
 reference-fixer, 13–14, 55, 58–60, 62, 67, 73–75  
 representation, 24, 57, 175, 187, 223–226, 229–233, 235–240, 244, 262, 286  
 representational/procedural semantics, 25, 248–251, 255, 259, 262, *see also* Discourse Representation Theory  
 rigid designation, 21, 22, 46, 52, 157, 177–179, 182–186, 188–192, 194, 195, 200, 251  
 Rivière, J., 271  
 Routley, R., 10, 78, 79, 89, 204  
 Rozakis, D., 270, 279, 282  
 Russell, B., 28, 77, 142, 247–249, 252, 254, 255, 265, 285  
 Russellian proposition, *see* structured proposition  
 Russellian theory of names, *see* descriptivism  
 Ryle, G., 79  
  
 Sade, D. A. F., 278  
 Sainsbury, R. M., 18, 20–21, 100, 193, 221  
 Salmon, N., 78, 79, 83, 84, 101, 153, 171, 203, 205, 209–211  
 Santambrogio, M., 94  
 Schaeffer, J. M., 276, 280, 282  
 Schiffer, S., 101, 103, 104, 135, 153, 208, 209  
 Schnieder, B. & von Solodkoff, T., 162, 169, 175, 176  
 Searle, J., 208  
 Smith, N. J. J., 174  
 Soames, S., 153, 205  
 Stalnaker, R., 52, 55, 87, 137, 145, 146, 150, 151  
 state of affairs, 11, 12, 35, 36, 95, 109, 166  
 stipulation, 16, 84, 85, 87, 93–95, 97, 103, 104, 108  
 Strawson, P. F., 247, 254  
 structured proposition, 13, 18, 22, 44–47, 50–51, 57–59, 62, 66, 68, 70, 71, 73, 119, 121–125, 178, 184, 192  
 supervenience, 17–18, 107–109, 113, 116–118, 131  
  
 Tarski, A., 250–252, 262  
 Tavinor, G., 18, 23–24  
 Taylor, M., 213, 214, 216, 218  
 textual/paratextual/metatextual use, 8–9  
 Thomasson, A. L., 19, 32, 101, 104, 105, 127–132, 135, 153, 160, 165, 169, 170, 174, 176, 181, 203, 205, 209  
 two-dimensional semantics, 14, 51–56, 80, 90  
  
 vagueness, 171–175  
 van Inwagen, P., 47, 64, 135, 153, 155, 160, 171, 174, 175, 177, 206–208  
 Vichianism, 16, 93–101, *see also* stipulation  
 Vico, G. B., 93  
 virtuality, 24, 223–226, 228–230, 235–239, 244  
 Voltolini, A., 10, 15–17, 153, 203, 205, 209  
  
 Walton, K. L., 22–23, 99, 104, 105, 135, 153, 157–160, 162, 176, 185–186, 197–198, 200, 221, 231, 232, 239, 241, 242  
 Williamson, T., 171, 175  
 Wittgenstein, L., 95, 280  
 Wolterstorff, N., 153  
 Woolley, B., 224, 225  
  
 Yablo, S., 153, 160, 198, 200  
 Yagisawa, T., 203, 209, 215, 219  
  
 Zalta, E. N., 105, 128, 204  
 Zhai, P., 224, 225, 229