

# Modal Meinongianism and Fiction: the Best of Three Worlds<sup>1</sup>

Francesco Berto

## Abstract

We outline a neo-Meinongian framework labeled as *Modal Meinongian Metaphysics* (MMM) to account for the ontology and semantics of fictional discourse.

Several competing accounts of fictional objects are originated by the fact that our talking of them mirrors incoherent intuitions: mainstream theories of fiction privilege some such intuitions, but are forced to account for others via complicated paraphrases of the relevant sentences.

An ideal theory should resort to as few paraphrases as possible. In § 1, we make this explicit via two methodological principles, called the Minimal Revision and the Acceptability Constraint. In § 2, we introduce the standard distinction between *internal* and *external* fictional discourse. In §§ 3 to 5, we discuss the approaches of (traditional) Meinongianism, Fictionalism, and Realism – and their main troubles.

In § 6 we propose our MMM approach. This is based upon (1) a modal semantics including *impossible* worlds (§ 6.1); (2) a qualified Comprehension Principle for objects (§ 6.2); (3) a notion of *existence-entailment* for properties (§ 6.3). In § 7 we present a formal semantics for MMM based upon a representation operator. And in § 8 we have a look at how MMM solves the problems of the three aforementioned theories.

## 1. The methodology of fiction: Minimal Revision and the Acceptability Constraint

Fictional objects are objects mentioned and described in operas, tales, stories, myths, etc. Fictional objects in this broad sense may well be imported from the actual, historical world: Napoleon in *War and Peace*, Virgil in Dante's *Divina commedia*. However, many such objects are native to the respective works of fiction: Sherlock Holmes, Anna Karenina, or Mr. Pickwick. We will call them *purely fictional* objects. These will be the main focus of this paper; but we shall have something to say also on non-purely-fictional objects in due course.

Several competing accounts of fictional objects have been on the philosophical market for a long time. This depends on the fact that, *prima facie*, our thinking and talking of such objects mirrors incoherent intuitions. We claim, on the one hand, that Sherlock Holmes does not exist, its fictional status consisting precisely in this; but on the other hand, we also claim that there is such a fictional character as Sherlock Holmes. We want to take the proposition that Holmes is a detective as in some sense true, and we appreciate its

---

<sup>1</sup> Thanks to Friederike Moltmann, Robert Stalnaker, Shahid Rahman, Alberto Voltolini, Benoît Conti, and an anonymous referee, for comments and helpful suggestions on various versions of this paper.

difference in truth value with the proposition that Holmes is a teddy bear. But when asked why Holmes cannot help the police to solve crimes, we are likely to admit that, well, he is a detective but a fictional one; he cannot do such things. Mainstream theories of fictional objects privilege some of these intuitions and provide straightforward accounts for them. But given the *prima facie* incoherencies at issue, all theories are then forced to dismiss other intuitions, or to account for them only via complicated paraphrases of the sentences involving them.

We agree that any consistent theory of fictional objects must give up appearances somewhere. But we think that the *best* theory of fiction is the one that resorts to as few paraphrases of our ordinary discourse on such objects as possible. More specifically: we adopt as a general methodological maxim that a theory of fictional objects should depart from the literal reading of fictional discourse only to the extent that such reading appears to be absurd, blatantly false, or incoherent when taken at face value. This we call *Minimal Revision* – a kind of Ockham’s razor for theories of fiction: “Do not revise fictional discourse beyond necessity”.

Another methodological maxim we follow is that, when revisions are needed, competent speakers should be capable of *accepting* the proposed paraphrase of their literal discourse, once the accompanying theoretical explanation of the paraphrase has been provided. Intuitions can be massaged, but only to some extent. This we call the *Acceptability Constraint*.

It is widely agreed that the standard Russell-Quine paraphrase strategy in order to dispense with reference to, and quantification on, fictional objects, massively violates these methodological aims.<sup>2</sup> The other three main (clusters of) approaches to fictional entities are (1) Meinongianism (Routley [1980], Parsons [1980], Jacquette [1993]), (2) Fictionalism (Currie [1990], Walton [1990]), and (3) Realism (Wolterstorff [1980], Salmon [1998], Thomasson [1999], [2001], van Inwagen [1977], [2003]). Meinongianism, Fictionalism, and Realism may be summarized in a single sentence by considering what each theory claims of the name “Sherlock Holmes”. According to Meinongianism, the name denotes a nonexistent object. According to Realism, it denotes an abstract, existent object. According to Fictionalism, it denotes nothing at all.

In the following Sections we will briefly rehearse the respects in which each of the three approaches fares well, and the main, generally recognized troubles. We have to say at the outset that our preferences are for Meinongianism. However, we are aware of the many unsympathetic reactions to traditional Meinongian theories. In fact, the aim of this paper is to propose a revised Meinongianism, which we deem capable of picking the best part of

---

<sup>2</sup> First, proper names such as “Sherlock Holmes” and “Pegasus” cannot be understood as (abbreviations of) definite descriptions, as Russell and, more explicitly, Quine had hoped. Second, no one has an idea of how systematic paraphrases of fictional discourse could be carried out. To pick an example from van Inwagen [2003]: 136, it is difficult to get rid of such complex quantificational structures as: “There is a fictional character who, for every novel, either appears in that novel or is a model for a character who does”. Third, the Russell-Quine paraphrases, even when they can actually be carried out, seem unable to preserve the right truth values and entailments. We want to appreciate the intuitive difference in truth value between “Anna Karenina is a woman” and “Anna Karenina is a teddy bear”, but the standard Russell-Quine analysis makes all such claims indiscriminately false.

each of the three mainstream theories, while at the same time avoiding the respective troubles.

## 2. Internal and external fictional discourse

We begin by introducing a distinction commonly made in the philosophical literature on fictional objects: the one between the so-called *internal* and *external* fictional discourse. Sentences describing a character within works of fiction, as well as statements made by readers (by spectators, and audience in general) that report the content of such works, belong to the former kind: “Holmes is a detective”, “Holmes lives in 221b Baker Street”, “Holmes and Watson were close friends”, etc. This is how Holmes is characterized by Doyle in his stories, either explicitly, or implicitly, given a suitable notion of entailment, plus background information (more on this later). And this is how we talk about Holmes when we are reporting and describing the contents of Doyle’s stories.

On the other hand, when readers and critics talk about fictional objects *as* fictional characters, and discuss, e.g., the historical role of a certain character in the development of a literary tradition, or the circumstances of its introduction in a story, etc., these claims belong to external discourse: “Holmes is a fictional character”, “Holmes is more famous than any real detective”, “Holmes is an epitome of the values of the Victorian age”, etc.

The distinction between internal and external fictional discourse mirrors the fact that fictional objects lead, so to speak, a double life. In Kit Fine’s words:

On the one hand, they have certain properties within the contexts in which they appear; they love and hate, thrive and fail, and live their varied lives. On the other hand, they also relate to the real world; they are created by authors, read by readers, and compared, for better or worse, with one another and with what is real. (Fine [1982]: 97).

Having registered the distinction, let us move to a brief account of the three main (clusters of) theories of fiction, beginning with Meinongianism.

## 3. Naïve and nuclear Meinongianism

First, Meinongians distinguish the *Sein* of objects – their existential status – from their *Sosein*: their having (certain) features or properties. And Meinongians claim that an object can have a set of features even if it doesn’t exist. This is the so-called Principle of Independence. According to Meinongians, we can refer to, and quantify on, nonexistent objects. “Holmes is a detective” can be taken at face value, treating the singular term occurring as the grammatical subject as an authentic, referring name. And the sentence is true because Holmes actually and literally *is* a detective, albeit a nonexistent one.

Naïve Meinongianism endorses what we may call, following Parsons [1979b], [1980], and by analogy with naïve set theory, an “Unrestricted Comprehension Principle” for objects:

(UCP) For any condition  $\alpha[x]$  with free variable  $x$ , some object satisfies  $\alpha[x]$ .<sup>3</sup>

The intuition is that we specify an object via a given set of properties ascribed to it, such as *x is a detective, x lives in 221b Baker Street, x has amazing powers of observation and deduction...* Suppose  $\alpha[x]$  is the conjunction of the relevant predicates. Then, according to the UCP, an object is described by  $\alpha[x]$ . Call the object so described “Holmes”, *b*. Then Holmes has the pertinent set of properties:  $\alpha[b]$ .<sup>4</sup>

However, the UCP in its naïve form leads to unacceptable consequences. The two most famous objections are due to Russell [1905a-b]: first, if one takes into account inconsistent characterizations, the UCP forces us to admit not only mere *possibilia*, but also impossible objects – things that violate the Law of Non-Contradiction by instantiating inconsistent properties. Quine’s [1948] (in)famous example is “the round square cupola of Berkeley College”. But one doesn’t need to coin *ad hoc* descriptions: sometimes fictional objects are characterized inconsistently by the actual stories in which they appear.<sup>5</sup>

Secondly, if the UCP held generally for any condition, we could run a generalized ontological argument to prove the existence of anything we like. For instance, one may pick the condition  $\alpha[x] = “x \text{ is golden} \wedge x \text{ is a mountain} \wedge x \text{ exists}”$ . Now existence is a perfectly normal first-order property for Meinongians, so the UCP would give us, completely *a priori*, an existent golden mountain. This will not do.

As a result, it seems that we have to restrict the class of conditions or properties that can figure in a comprehension principle for objects. Parsons, Routley, and Jacquette introduce a subclass of predicates (called *assumptible, characterizing* or, more often, *nuclear*), and restrict the Comprehension Principle to nuclearity:

(NCP) For any condition  $\alpha[x]$  *on nuclear properties* with free variable  $x$ , some object satisfies  $\alpha[x]$ .

In particular, existence is not a nuclear property. Whereas Holmes does have such nuclear properties as being a detective because he is so characterized by the Doyle stories, he does not have the property of existing, despite the fact that he is certainly characterized as such

---

<sup>3</sup> Actually, Russell’s famous [1905a-b] criticisms of Meinong addressed definite descriptions, so the UCP might accordingly be reformulated as:

(UCP) Any definite description  $\iota x \alpha[x]$  designates an object satisfying the description.

<sup>4</sup> Plus, perhaps, whatever properties follow from them according to a given notion of entailment (more on this later).

<sup>5</sup> Now and then the inconsistency may be unintentional. For instance, in one of Conan Doyle’s stories, *The Sign of the Four*, we are told that Watson limps because of a war wound at his leg; in *A Study in Scarlet*, Watson’s wound is not on his leg, but on his shoulder, and Watson does not limp. But sometimes the inconsistency may be essential to the plot. Suppose we write a novel, and in its first chapter we have the Mad Mathematician produce a round square. If the intentional inconsistency is excised, the fact that mathematicians all over the world are amazed by this result in the second chapter becomes unexplainable.

(at least implicitly) by those stories: in Doyle's novels, Holmes is a very existing guy, but existence is an extranuclear property.

How to distinguish between nuclear and extranuclear properties in a principled way is a traditional problem of nuclear Meinongianism.<sup>6</sup> But the worst difficulties come from the fact that both naïve and nuclear Meinongianism are committed to what Kit Fine [1982], [1984] has called *literalism*: the doctrine that fictional objects literally have the features (at least the nuclear ones) ascribed to them in the works of fiction. Fine opposes literalism to *contextualism*, that is, the view that Holmes does not literally have those properties: he has them only in the appropriate stories (the "contexts").<sup>7</sup>

Now, when we are dealing with *internal* discourse on fictional entities, contextualism seems to be more reasonable a position than literalism. For one thing, literalism seems to fall foul of what we take to be facts of the matter: not only it is literally false that Holmes exists, but also that he lives in Baker Street (221b Baker Street in London used to host a bank), even though living in Baker Street looks very much like a nuclear property. Besides, if Holmes were a man, and a detective, and lived somewhere, etc., all these being (nuclear) properties Holmes is supposed to literally have, he would be a concrete, spatiotemporally located entity: *being a detective*, after all, intuitively entails such features as having a body, being located in space and time, being capable of causally interacting with the environment, etc. But Holmes, as far as we know, is nowhere to be found, even if we were freed from our practical limitations in travelling through space and time. In Doyle's stories, Holmes is also said to sense and reason about the world, but how can this be literally true if he is a nonexistent object? Hence, Holmes cannot literally have the features which would be apparently *entailed* by his being literally a detective, or a man.

#### 4. Fictionalism

Fictionalism on fictional objects is an appealing strategy because it aims at dispensing with literalism completely. Discourse on fictional objects is taken as essentially involving pretence and make-believe. According to Currie [1990], authors of fiction produce sentences with the aim that the audience will make believe the propositions expressed by such sentences. In Walton's [1990] general theory of representation and fiction, individuals play collective games of make-believe in which it is specified that so-and-so is to be imagined or represented. Roughly, the account works by means of the notion of something being "true in a fiction", or "true within a pretense". But fictional truths are not *truths* – not even a special kind of truths. That it is fictionally true that Holmes is a detective means that it is fictional (it is a component of the game of make-believe, and is to be imagined within

---

<sup>6</sup> Here are a few examples, taken from Terence Parsons' classic works [1979a], [1980]. Nuclear predicates: "is blue", "is tall", "kicked Socrates", "was kicked by Socrates", "kicked somebody", "is golden", "is a mountain". Extranuclear predicates: *Ontological*: "exists", "is mythical", "is fictional"; *Modal*: "is possible", "is impossible"; *Intentional*: "is thought about by Meinong", "is worshipped by someone"; *Technical*: "is complete", "is consistent".

<sup>7</sup> Fine [1982], p. 97. See also Fine [1984], pp. 135ff., for an explanation of the close connections between Parsons' nuclear Meinongianism and literalism.

the game) that Holmes is a detective. In Walton's words, "Propositions that are 'true in a fictional world,' or fictional, are propositions that, in a given social context, are to be imagined as true." (Walton [1991], p. 380.)

Fictionalism aims at accounting for representation without objects of representation. The intuition is supported by observing that, when Doyle wrote that Holmes lived in Baker Street, he didn't actually assert the sentence expressing this proposition. Doyle just *pretended*, in a non-deceptive way, to perform this illocutionary act, so he was not committed to the content of his claim's being actually the case.

The main problem with Fictionalism is that it seems to fall foul of our maxim of Minimal Revision. Some sentences involving fictional objects are literal truths (or falsities), seriously asserted without any explicit or implicit pretence. This apparently holds for external fictional discourse. Nobody in her mind would criticize Doyle for having written the literal falsity that Holmes lived in 221b Baker Street. That's true in the story, not in the real world. But someone engaged in literary criticism may be rightly criticized (or approved) for having written that Holmes is an epitome of the values of the Victorian age. People making such claims appear to aim at saying literal truths about the actual world; they appear to make genuine assertions and to commit themselves with no pretence. The fictionalist revision of external discourse cannot be motivated by the need to deal with inconsistencies or blatant falsities inhabiting it. Besides, Fictionalism may also contravene the Acceptability Constraint: it would be hard to convince ordinary speakers that they are *pretending*, not only when they claim that Holmes lives in Baker Street, but also when they say that Holmes is Doyle's main literary character.<sup>8</sup>

## 5. Realism, *de re* and *de dicto* pretence

Realist theories of fictional objects share two main theses: (a) fictional objects are existent, abstract entities, and (b) the apparatus of predication and ascription of properties to fictional objects is ambiguous. Typically, realist theories take external discourse on fictional objects literally. Internal discourse, on the contrary, enjoys a special status to be generically labeled as fictional, in the sense that it is not taken as literally true. Realist theories of fiction differ in the kind of strategy adopted in the analysis of internal discourse.

In Thomasson's [1999] account, fictional objects are abstract artifacts created by the storytelling activities. Internal sentences have to be understood as implicitly prefixed with a non-factive "according to the story" operator. Whereas it is literally true that Holmes is the main character of Doyle's stories, it is only true in Doyle's stories that Holmes is a detective – and rightly so: in the Realist account, Holmes is an abstract object, and abstract objects cannot, of course, literally be detectives.

Also van Inwagen takes internal fictional discourse as involving pretense. According to him, at best we can say that fictional characters *bold* the properties ascribed to them by

---

<sup>8</sup> As an anonymous referee has pointed out to me, pretence theories also have the problem of giving a semantics for claims of the form "It is pretended that Holmes lived in Baker Street", without lapsing into some form of Meinongian or Platonic realism.

internal fictional discourse (van Inwagen [2003]: 146), or better, *are ascribed* such properties, ascription being a three-place relation: fictional character  $x$  is ascribed property  $p$  in (some locus of) work of fiction  $w$ .<sup>9</sup>

The main pro of Realism consists in its taking many sentences (those belonging to the external discourse) at face value, thereby respecting the maxim of Minimal Revision more than Fictionalism in this respect. Realism can afford the ontological commitment entailed by these sentences: it *is* committed to fictional objects, admits them in the domain of quantification, and accounts for the peculiarities of their status (for instance, their being nowhere to be found in the physical world) by drawing on their abstractness.

Hence comes also a main, widely recognized con of Realism, having to do with negative existentials – a point on which, on the contrary, both the vexed Meinongianism and Fictionalism have an evident pro. Adult people deny in all seriousness that there exist such things as Holmes, Santa, or the Big Bad Wolf. People generally accept that negative existentials concerning fictional objects are literally true. And these sentences count as external, not internal discourse: in Doyle’s stories, Holmes is very much existent. It seems that if “Holmes is a fictional character” should be taken as a literal truth, then by the same token also “Holmes does not exist” should be taken the same way. It is precisely because Holmes is a purely fictional character that its existence is denied. But Realist theories are forced to break this intuitive uniformity: Amie Thomasson finds herself admitting that negative existentials concerning fictional characters are literally false (see Thomasson [1999], p. 111). In a footnote to *Creatures of Fiction*, van Inwagen admits that, given his theory, what to do with “Mr. Pickwick does not exist” is “a very complicated question”.<sup>10</sup> The Acceptability Constraint is at stake.

A proposal often made in these cases is to assume that such denials of existence work as cases of implicitly restricted quantification. What we are claiming is not that there is no such thing as Mr. Pickwick at all, but that Mr. Pickwick is not to be found in the domain of concrete entities, for it’s an abstract object (see Thomasson [1999]: 112). However, this is unconvincing. Such nonexistence claims are naturally interpreted as totally *unrestricted*. Implicitly restricted quantification can be detected when we use “there is” (“There are no dolphins”, uttered at the Taronga Zoo, means: no dolphins in the zoo – not that dolphins are extinct), not when we use “(there) exists”. When I say “There exist no dolphins”, or “Dolphins do not exist”, as Walton [2003] has observed, I am actually making the (luckily false) claim that dolphins are extinct, full stop.

A second problem of Realism comes from its resorting to pretence and in-the-story operators. Since Holmes is an abstract object, it can be literally true that Holmes is a fictional character, but it cannot be literally true that Holmes is a detective. Internal fictional

---

<sup>9</sup> This solves the problem of inconsistent fiction: fictional characters can be ascribed inconsistent properties in internal fictional discourse, for they do not actually instantiate them. Notice that the *holding* and *being ascribed* relations do not constitute, according to van Inwagen, special kinds of predication such as Zalta’s encoding of properties by abstract objects that do not exemplify them (see Zalta [1983]). It is not that fictional objects somehow have the properties ascribed to them within fictional works, but do not exemplify them. Internal fictional discourse is taken by van Inwagen precisely as a kind of pretence (see van Inwagen [2003]: 150, fn. 18).

<sup>10</sup> Van Inwagen [1977], p. 308 fn. 11.

discourse cannot be taken at face value and involves some kind of pretence. Which kind? Two options are available, depending on whether we take the pretence to be *de re* or *de dicto* (Thomasson [2003]: 210); and neither appears to be very palatable.

When an in-the-story operator is prefixed to internal discourse, we obtain such things as:

(1) In Doyle's stories, Holmes is a detective.

If we read (1) *de re*, this means that we claim, of Holmes, that he has the property of being-a-detective-in-Doyle's-stories. If we read it *de dicto*, we claim that 'in Doyle's stories [Holmes is a detective]', this meaning something like: Doyle pretends that there was some man called "Holmes" who was a detective.

The *de re* option was not open to full-blooded Fictionalism: it entails that "Holmes" does refer to something, the element of pretence or make-believe being confined within the predicate. But for Fictionalists there is no such thing. According to Realism, though, there *is* such a(n abstract) thing as Holmes. So the *de re* reading is an option for them, and is actually embraced in Thomasson [1999].

Taking the pretence involved in internal fictional discourse to be *de re* brings some advantages when one deals with sentences involving non-purely-fictional objects. When Doyle claims that Holmes lived in London, 221b Baker Street, the name "London" refers *de re* back to the real London, with Doyle making believe, of London, only that it had the property of being inhabited by Holmes at 221b of its Baker Street. However, as many have emphasized, the *de re* version of pretence has the unpalatable consequence that works of fiction pretend that *abstract* objects have such properties as being a detective, living in Baker Street and solving crimes. Again, the Acceptability Constraint is at stake: should we accept that, when we or Doyle claim of Holmes that he was a brilliant detective, we are pretending to make (actually false) claims on an abstract entity?

What about the *de dicto* reading of (1), then? First, this kind of pretence raises a problem if one has a broadly Millian conception of names,<sup>11</sup> for then any sentence with an empty name doesn't express a complete proposition. If so, then also any sentence in which the former is embedded gets infected by this failure of reference, and does not express a complete proposition either. Embedding a sentence which does not express a proposition in the scope of a non-factive sentential operator is of no help. If "Holmes is a detective" doesn't express a complete proposition, (1) does not express a complete proposition either. *Cross-fictional* sentences like "Sherlock Holmes is more famous than Conan Doyle", or worse, "Sherlock Holmes is taller than Bilbo Baggins", are even more puzzling and intractable.

Second, no uniform account is available anymore when works of fiction also include names of non-purely-fictional objects: for then, whereas within internal fictional discourse "Holmes" does not refer, "London" should plausibly still refer *de re* back to London, the

---

<sup>11</sup> Millianism being the view according to which the semantic content of such singular terms as names and demonstratives is exhausted by their referents, which enter as constituents of the propositions expressed by the sentences including such singular terms.



pretence being limited to there being someone called “Holmes” living in it during the Victorian age. Van Inwagen [1977]: 51 is forced to claim that “Napoleon” is ambiguous. Normally, it is used to denote the real man. But when critics discuss *War and Peace* in the context of external discourse, “Napoleon” denotes a completely different object, that is, an abstract entity. And no uniform account can be expected also for the names of purely fictional objects. “Sherlock Holmes”, when used by Conan Doyle in writing the fiction, merely pretends to refer to something; whereas, when used in external fictional discourse as a name for the character created by Doyle, it refers to an abstract existent.

A final difficulty comes from objects that do not belong to fiction *stricto sensu*, but are postulated within false non-literary theories or religious beliefs: typical cases often mentioned are Zeus, Vulcan (the planet, not the god), or phlogiston. Could these entities be actual, existent artifacts? The XIX Century astronomers Jaques Babinet and Urbain Le Verrier postulated the existence of Vulcan as an intra-Mercurial planet, which was supposed to explain an irregularity in the perihelion of Mercury. They got it wrong, and their failure is explained by Einstein’s relativity. Were our world Newtonian, they might have guessed well. In this case, to be sure, they wouldn’t have discovered an abstract object, but a concrete planet. (The astronomers who postulated Uranus, on the other hand, did get it right; and what they discovered, of course, was not an abstract object either.) So we have that a name such as “Vulcan”, used by some scientists in order to refer to a concrete object that turned out not to exist, ended up referring to an abstract, existent one –this *despite* the fact that the very creators of the theory did not intend to refer to an abstract object at all. The Acceptability Constraint is under pressure again. And the problem is especially pressing for Thomasson [1999], since she categorizes myths and (mistaken) theories together with literary works.

## 6. A compressed presentation of MMM

The examination of traditional Meinongianism, Fictionalism, and Realism, has left us with a set of open problems and detectable violations of the maxim of Minimal Revision, of the Acceptability Constraint, or of both. We will now devote this Section and the following one to an exposition of our revised Meinongianism.

Our theory is based upon three pillars: (1) a world semantics including *impossible* worlds, besides possible ones; (2) the admission of a Comprehension Principle for objects in unrestricted, but qualified form, which follows suggestions by Daniel Nolan [1998], Nick Griffin [1998], and above all Graham Priest [2005]; (3) an intuitive notion of *existence-entailment* for properties. We will have an informal look at the three ideas in the three following sub-paragraphs. Their combination produces a new form of *Modal Meinongian Metaphysics* (MMM). In Section 7, then, we will build a formal semantics that draws on these ideas. In Section 8, we shall finally show that MMM deals with each of the problems manifested by the three competing theories in a more satisfactory way.

## 6.1. Impossible worlds

What are impossible worlds? One can detect four main definitions in the literature. Ordering them from the more to the less general, they go as follows.

(1) First, dual to the identification of possible worlds with ways things could have been comes the identification of impossible worlds with ways things could not have been. Not everything is possible, that is, some things just can't happen. Anything that just can't happen must be an impossibility; and these ways the world could not be are impossible worlds (see e.g. Restall [1997], Barwise [1997], Beall & van Fraassen [2003], Ch. 6).

(2) A second definition has it that impossible worlds are worlds where the laws of logic are different (see e.g. Priest [1992]). This is logic-relative: given some logic  $L$ , an impossible world is one in which the set of truths is not one that holds in any acceptable interpretation of  $L$ .

(3) A more restrictive definition claims that impossible worlds are worlds where the set of things that hold is not the set of things that hold in any *classical* interpretation (see Priest [1997]: a classical logician can consider a world where the Law of Excluded Middle (LEM) fails as a logically impossible world, since she takes classical logic as the correct logic).

(4) A still more specific definition claims that an impossible world is a world where some contradictions are true, that is, where sentences of the form  $\alpha$  and  $\neg\alpha$  hold, against the Law of Non-Contradiction (LNC) (see Lycan [1994]).

Why should we admit such turbulent guys as impossible worlds in modal semantics and ontology? Well, for instance, because we are capable of considering logically impossible situations, and of making discriminations about what goes on at them. Worlds semantics for minimal logic includes non-normal worlds in which the LEM and *ex falso quodlibet* (that is, the law according to which a contradiction entails everything) fail. The former also fails in standard Kripke semantics for intuitionist logic. Now, it seems that we refer to these worlds when we evaluate such conditionals as “If intuitionist logic were the correct logic, then the LEM would fail” (true); and “If intuitionist logic were correct logic, then *ex falso* would fail” (false). Anyone who understands intuitionism, or minimal logic, or quantum logic, etc., knows how things would be if one of these logics were correct (assuming they are not). Discourse on ways things couldn't be has its own logic in a broad sense: some reasoning in it is correct, some is not. And impossible worlds are nowadays proposed by various authors as a natural extension of possible worlds theories, having useful applications in the study of the notions of propositional content, intentional state, belief management, etc.<sup>12</sup>

---

<sup>12</sup> For a comprehensive perspective on impossible worlds, see Berto [2009]. On the controversial metaphysical status of such worlds, see Berto [forthcoming].

## 6.2. Qualified Comprehension Principle

The second move towards MMM we have advertised above is due mainly to Priest [2005]. It consists in admitting a Comprehension Principle for objects in unrestricted form, by *parameterising* it to worlds: given any condition  $\alpha[x]$ , some object is described by it. However, it has its characterizing properties, not necessarily at this world, but at *others* – at the worlds that make the characterization true; This we call the Qualified Comprehension Principle:

(QCP) For any condition  $\alpha[x]$  with free variable  $x$ , some object satisfies  $\alpha[x]$  *at some world*.

Explicit reference to worlds in the QCP is what motivates calling this kind of Meinongianism “modal”. This move can help with the ontology of fictional objects. After all, fictional objects are the target of intentional states and cognitive representations. Hence comes Priest’s justification of the strategy:

Cognitive agents represent the world to themselves in certain ways. These may not, in fact, be accurate representations of this world, but they may, none the less, be accurate representations of a *different* world. For example, if I imagine Sherlock Holmes, I represent the situation much as Victorian London (so, in particular, for example, there are no airplanes); but where there is a detective that lives in Baker St, and so on. The way I represent the world to be is not an accurate representation of our world. But our world could have been like that; there *is* a world that is like that (Priest [2005]: 84).<sup>13</sup>

Objects picked by a description, therefore, may *always* satisfy the relevant condition  $\alpha[x]$ . Unlike nuclear Meinongianism, we don’t need to isolate a subset of “nuclear” or “characterizing” properties. Also {goldenness, mountainhood, existence} works fine now, for we need not assume that an object so characterized, that is, an existent golden mountain, has its characterizing properties at the actual world. As far as we know, no golden mountain inhabits the actual world, but golden mountains are available at the worlds at which the stories we can tell on existent golden mountains hold.

Worlds work in the theory in a way similar to Kit Fine’s *contexts* in his [1982] discussion of nonexistent objects. Fine rejects the identification of worlds with contexts, or the possibility of representing context by means of a worlds semantic machinery, on the ground that “one of the peculiarity of contexts is that the propositions true in them need not form either a consistent or a complete set” (Fine [1982]: 101). But of course this worry vanishes once one has countenanced impossible worlds: these are precisely things in which propositions can be true, and such that the propositions true in them need not form either a consistent or a complete set. Fine’s contexts are representations of (occasionally

---

<sup>13</sup> To be more precise, there are many such worlds, since representations are incomplete with respect to many details. For instance, is the world of Sherlock Holmes a world where Holmes has an even or an odd number of hairs on his head when he first meets Watson? (The example comes from Lewis [1978]: 270.) Doyle tells us nothing about it. One should assume that there are various “Sherlock Holmes worlds”, and that the question has different answers at different worlds.

inconsistent and incomplete) situations in which fictional objects occur, and they can be encoded by worlds in the MMM theory.

### 6.3. Existence-entailments

In our Meinongian framework fictional objects should also have properties *entailed* by those they are explicitly characterized as having, given some suitable notion of entailment. We will say more on the kind of entailment at issue in the next Section. But we begin by remarking that, since in some forms of Meinongianism (such as Routley's [1980], [1982]) to exist is to have causal properties, and/or to be located in space and time, one may take some properties or relations involving causal features to be existence-entailing in one or more arguments (*which* properties or relations are existence-entailing may be, of course, a matter of ontological debate).

This idea, as far as we know, is due mainly to Linsky and Zalta [1994].<sup>14</sup> For instance, if  $x$  kisses  $y$ , then both  $x$  and  $y$  must exist, and if  $x$  thinks about  $y$ ,  $x$  must exist although  $y$  need not. (At least at possible worlds: one may probably fantasize on impossible nonexistent kissers and thinkers! – More on this in the following Section.) This accounts for the intuitive idea that Sherlock Holmes, being a nonexistent object *at* this world, cannot actually kiss anyone here, nor can he entertain any thought here, although he can be thought of (by Doyle, or by any reader of the Doyle stories). However, if some Doyle novel features Holmes kissing Watson, then Holmes does kiss Watson at the worlds at which the characterization is true and, at those worlds (or at least, at the possible ones among them), Holmes does exist.

This position entails a partial rejection of the Meinongian Principle of Independence: not every *Sosein* is independent from *Sein*, for many properties do entail existence. However, MMM is still a full-fledged form of Meinongianism, in that it denies that *all* properties are like that. As such, MMM is a rejection of so-called Serious Actualism, that is, of the claim that any object must exist in every circumstance in which it has any property – the thesis that predication, or the having of properties as such, entails existence.

However, we agree with Nathan Salmon in claiming that Serious Actualism is “a confused and misguided prejudice” (Salmon [1998]: 290), which can be exposed also independently of our favourite metaphysics of fiction. For many non-fictional objects that do not exist still display properties. Past existents, for instance. Walter Scott is both long gone nowadays, but he still bears many properties: he has just been mentioned by us; he has the property of having written *Waverley*. Of course, when Scott wrote *Waverley* he did exist, but writing something and having written it are different properties. Similarly, Heraclitus is still misunderstood, Caesar was betrayed by Brutus, and John Lennon is thought about by his fans.

---

<sup>14</sup> Actually, they talk of *concreteness*-entailing properties, but it is easy to provide a rough-and-ready translation from their modal theory to MMM.

The QCP tells us that for any condition,  $\alpha[x]$ , such a condition is satisfied by some object at some world, although it need not be the actual world. By admitting impossible worlds, we can account for inconsistent objects, such as Quine’s round square cupola of Berkeley College, that is, an object which, given the assumption that if something is a square then that thing is not round, is round and not round,  $Rx \wedge \neg Rx$ . We simply admit inconsistent worlds that realize contradictions. Notice that we do not need to admit *true* contradictions, or even possibly true ones. We can stick to the intuition that Quine’s cupola is an impossible, therefore not possibly existing, object. If we tell a story about some rock climber who climbs on the round square cupola of Berkeley College, supposing that “ $x$  climbs on  $y$ ” is a predicate which is existence-entailing in both arguments, then the round square cupola of Berkeley College is round, square, and existent – not at this world, nor at any possible world, but at those impossible worlds described by our story, and which realize the characterization “(existing) round square cupola of Berkeley College”.

## 7. Formal MMM semantics

After presenting the three pillars of MMM informally, we are now going to make things more precise via a little bit of formal machinery.<sup>15</sup> First, we take the usual first-order language with a set of individual constants,  $n$ -place predicates (with a distinguished one-place predicate,  $E!$ ), individual variables, connectives  $\neg, \wedge, \vee, \rightarrow$ , two quantifiers,  $\Lambda$  and  $\Sigma$  (so written for reasons to be explained soon), and the usual rules for well-formedness. Next, we add to the standard language an intensional *representation* operator,  $\Psi$ : if  $\alpha$  is any well-formed formula, ‘ $\Psi\alpha$ ’ is a well-formed formula, to be read: “It is represented in such-and-such story that  $\alpha$ ”.

The primitive, Meinongian quantifiers  $\Lambda$  and  $\Sigma$  (to be read as “for all” and “for some”) are taken as existentially neutral: one can quantify on, and talk in general of, nonexistents. Existence is taken as a perfectly normal first-order property expressed by the predicate  $E!$ , employed to provide explicit existential commitment and to define the ordinary, existentially loaded quantifiers. “All existing things are such that...” is:

$$\forall x\alpha[x] =_{\text{df}} \Lambda x(E!x \rightarrow \alpha[x]);$$

and “There exists something such that...” is:

$$\exists x\alpha[x] =_{\text{df}} \Sigma x(E!x \wedge \alpha[x]).$$

---

<sup>15</sup> The following formal semantics comes basically from Priest’s [2005] account, but with a few adjustments.

Meinong claimed “There are objects of which it is true that there are no such objects” (Meinong [1969]: 490). But this is not a flat contradiction at all, when one has distinguished the two couples of quantifiers – on the contrary, it is mirrored in everyday talk:

There is something which has been sought by many, namely the site of Atlantis, but it does not exist (Wolterstorff [1961]).

I thought of something I would like to give you as a Christmas gift, but I couldn’t buy it for you because it doesn’t exist (Priest [2005]).

An interpretation of the language is a sextuple  $\langle P, I, @, D, R, \nu \rangle$ , where  $P$  is the set of possible worlds and  $I$  is the set of impossible worlds.  $P$  and  $I$  are disjoint and  $W = P \cup I$  is the totality of worlds.  $@ \in P$  is the distinguished actual world. (This would not be needed if all we wanted from the semantics were accounts of validity and logical consequence; but  $@$  is to do other jobs, as we shall see.)  $D$  is a non-empty set of objects,  $R$  is a binary relation on the whole set  $W$ . The semantics is framed as a constant domain structure to keep things simple.<sup>16</sup> Finally,  $\nu$  is the interpretation function assigning denotations to the non-logical symbols as follows:

If  $c$  is an individual constant, then  $\nu(c) \in D$ ;

If  $P$  is an  $n$ -place predicate and  $w \in W$ , then  $\nu(P, w)$  is a pair  $\langle \nu+(P, w), \nu-(P, w) \rangle$ , with  $\nu+(P, w) \subseteq D^n$ ,  $\nu-(P, w) \subseteq D^n$ .

If  $P$  is an  $n$ -place predicate,  $\nu$  assigns to it an extension  $\nu+(P, w)$  and an anti-extension  $\nu-(P, w)$  relative to worlds. Intuitively, the extension of  $P$  at  $w$  is the set of  $n$ -tuples of which  $P$  is true there, and the anti-extension is the set of  $n$ -tuples of which it is false. For possible worlds, we require that the two be exclusive and exhaustive for any  $P$  – let us call this the Classicality Condition:

$$\begin{aligned} \text{(CC)} \quad \text{If } w \in P, \text{ then } \nu+(P, w) \cap \nu-(P, w) &= \emptyset \\ \nu+(P, w) \cup \nu-(P, w) &= D^n \end{aligned}$$

This reflects the idea that possible worlds must be consistent and maximal: for any predicate  $P$ , if  $w$  is a possible world,  $P$  is either true or false of the relevant object(s) at  $w$ , but not both. Nevertheless, truth and falsity conditions are spelt separately (in a way familiar

---

<sup>16</sup> This is quite natural for Meinongians. People assume variable domains in ordinary modal semantics to account for the idea that different things may exist at different worlds. But in a Meinongian framework the domain of each world is simply the *totality* of objects: that some object  $o$  exists at world  $w_1$ , but not at world  $w_2$ , is accounted for by having  $o$  satisfy the existence predicate  $E!$  at  $w_1$  and not at  $w_2$ , and all the epicycles of Kripke variable domain semantics are left behind. *Simplex sigillum veri*.

from various kinds of non-bivalent semantics), for things may go differently at impossible worlds.

In order to evaluate quantified sentences we need assignments of denotations to the variables, the usual way: if  $a$  is an assignment (a map from the variables to  $D$ ), then  $v_a$  is the suitably parameterized interpretation:

If  $c$  is an individual constant, then  $v_a(c) = v(c)$ ;

If  $x$  is a variable, then  $v_a(x) = a(x)$ .

Next, we read “ $w \models_a^+ \alpha$ ” as “ $\alpha$  is true at world  $w$  with respect to assignment  $a$ ”, and “ $w \models_a^- \alpha$ ” as “ $\alpha$  is false at world  $w$  with respect to assignment  $a$ ” (we will omit the subscript when dealing with closed formulas, for which different assignments, as usual, make no difference). Then we have, for atomic formulas:

$w \models_a^+ P t_1 \dots t_n$  iff  $\langle v_a(t_1), \dots, v_a(t_n) \rangle \in v^+(P, w)$

$w \models_a^- P t_1 \dots t_n$  iff  $\langle v_a(t_1), \dots, v_a(t_n) \rangle \in v^-(P, w)$

For negation:

$w \models_a^+ \neg \alpha$  iff  $w \models_a^- \alpha$

$w \models_a^- \neg \alpha$  iff  $w \models_a^+ \alpha$

Since extensions and anti-extensions are exclusive and exhaustive at possible worlds, if  $w \in P$  we have that  $w \models_a^+ \neg \alpha$  iff it is not the case that  $w \models_a^+ \alpha$ . So thanks to the CC at possible worlds negation works “homophonically”, the classical way. And since  $@ \in P$ , that is, the actual world is possible, there are no truth value gluts or gaps at it: truth *simpliciter*, truth at the actual world, behaves in an orthodox way with respect to negation.

Conjunction, disjunction, the quantifiers and the (strict) conditional get the usual clauses at *possible* worlds. For all  $w \in P$ ,

$w \models_a^+ \alpha \wedge \beta$  iff  $w \models_a^+ \alpha$  and  $w \models_a^+ \beta$

$w \models_a^- \alpha \wedge \beta$  iff  $w \models_a^- \alpha$  or  $w \models_a^- \beta$

$w \models_a^+ \alpha \vee \beta$  iff  $w \models_a^+ \alpha$  or  $w \models_a^+ \beta$

$w \models_a^- \alpha \vee \beta$  iff  $w \models_a^- \alpha$  and  $w \models_a^- \beta$

$w \models_a^+ \alpha \rightarrow \beta$  iff for all worlds  $w_1$  (if  $w_1 \models_a^+ \alpha$  then  $w_1 \models_a^+ \beta$ ).

$w \models_a^- \alpha \rightarrow \beta$  iff for some world  $w_1$  ( $w_1 \models_a^+ \alpha$  and  $w_1 \models_a^- \beta$ ).

$w \models_a^+ \Lambda x \alpha$  iff for all  $d \in D$ ,  $w \models_a^+_{a(x/d)} \alpha$

$w \models_a^- \Lambda x \alpha$  iff for some  $d \in D$ ,  $w \models_a^-_{a(x/d)} \alpha$

$w \models_a^+ \Sigma x \alpha$  iff for some  $d \in D$ ,  $w \models_{a(x/d)}^+ \alpha$

$w \models_a^- \Sigma x \alpha$  iff for all  $d \in D$ ,  $w \models_{a(x/d)}^- \alpha$

(Where “ $a(x/d)$ ” denotes the assignment which is the same as  $a$ , except that it assigns to  $x$  the value  $d$ .) This means that the standard logical operators behave in an orthodox fashion (except for the fact that they are given truth and falsity conditions separately) at possible worlds. Instead, they behave anarchically at impossible worlds: the interpretation function  $v$  treats formulas formed by means of them as *atomic*, assigning extensions and anti-extensions directly, that is, at points in  $I$  the truth values of conjunctions, disjunctions, and quantifiers, are not assigned recursively, but directly determined by  $v$ .<sup>17</sup>

Now, the existence-entailing features of some properties mentioned in the Section above can be accounted for in the machinery by adding some formal constraints. We will assume that, if an  $n$ -place predicate  $P$  is existence-entailing in its  $i$ th place, it is such at all possible worlds:

If  $w \in P$ , then if  $\langle d_1, \dots, d_i, \dots, d_n \rangle \in v+(P, w)$ , then  $d_i \in v+(E!, w)$ .

Existence-entailments can be regarded as something similar to meaning postulates, fixing the semantics of some predicates and, specifically, their internal connections to the predicate “exists”. Meaning postulates are usually taken as necessary truths, holding at all possible worlds. So if Santa Claus thinks about Pegasus at  $w$ , then Santa Claus exists at  $w$  though Pegasus need not; if Holmes kicks Moriarty at  $w$ , then both exist at  $w$ ; and this holds whenever  $w$  is a possible world. What happens at impossible worlds, of course, is another story; nonexistent things, for instance, may think or kick at some impossible world.

Finally, worlds are allowed to access impossible worlds when the truth conditions for the representation operator  $\Psi$  are at issue. For any  $w, w_1 \in W$ :

$w \models_a^+ \Psi \alpha$  iff for all  $w_1 \in W$  such that  $wRw_1$ ,  $w_1 \models_a^+ \alpha$

$w \models_a^- \Psi \alpha$  iff for some  $w_1 \in W$  such that  $wRw_1$ ,  $w_1 \models_a^- \alpha$

These are the key clauses, and their intuitive explanation goes as follows. The semantics of the representation operator is just a restatement of the usual (restricted) binary accessibility semantics for modal operators of ordinary modal logics. The idea is that  $wRw_1$ , that is, there is a *representational accessibility* (R-accessibility) from  $w$  to  $w_1$ , if and only if, at  $w_1$ , things are as they are represented to be (within such-and-such a story, tale, fictional work, etc.) at  $w$ ; or, equivalently:  $w_1$  is R-accessible from  $w$ , just in case  $w_1$  realizes the way things are characterized, or described – that is, represented – to be (within such-and-such a story) at  $w$ . For instance, take the worst nightmares one may have (at the actual world  $@$ ); then a world  $w$  such that  $@Rw$  is a nasty world at which those worst nightmares come true.

---

<sup>17</sup> The idea was introduced by Rantala [1982] in order to provide a semantics for intentional operators making logical omniscience fail.



Now the reason why impossible worlds are accessible when one evaluates formulas of the form ‘ $\Psi\alpha$ ’, is precisely that we can form inconsistent descriptions more or less of any kind:  $\alpha[x]$  may be “ $x$  is a round square”, “ $x$  is round and not round”, “ $x$  is red or blue but it is not the case that  $x$  is red and it is not the case that  $x$  is blue”, etc. We may represent (imagine, think of, tell stories about, etc.) objects having any (set of) properties, however extravagant. And our QCP tells us that for any condition  $\alpha[x]$ , there will be worlds at which some object satisfies  $\alpha[x]$ .

Next, we need a definition of logical validity and logical consequence. If  $S$  is a set of formulas:

$S \models \alpha$  iff for every interpretation  $\langle P, I, @, D, R, \nu \rangle$ , and assignment  $a$ , if  $@ \models_a^+ \beta$  for all  $\beta \in S$ , then  $@ \models_a^+ \alpha$ .

For logical validity, we have just that  $\models \alpha$  is  $\emptyset \models \alpha$ . So logical consequence is truth preservation at the base world,  $@$ , in all interpretations, and logical validity is truth at  $@$  in all interpretations.<sup>18</sup>

The representation operator, despite having its semantics provided by an accessibility relation involving such anarchic guys as impossible worlds, must somehow be regimented. How to spell out the regimentation in detail is a subtle issue (as shown by the classic discussion in Lewis [1978]) and, luckily, one we need not get into here. We will limit ourselves to some general remarks. Fictional representations induced by fictional discourse must be closed under some notion of logical consequence. This is shown by the fact that we can argue about how things are in the relevant situation(s) and, in the process of arguing, we draw inferences. Some reasoning on what follows from what within a fiction is correct, some is not, also when the fiction at issue involves inconsistencies, even intentional ones.

The premises of such reasoning sometimes are not explicitly stated within the fiction, but are imported as default assumptions from actuality, background information, shared knowledge, etc. For instance, if within a Doyle story it is said that Holmes was in London at a certain time, we can infer that he couldn’t have been in Australia the day before, even when this isn’t explicitly stated in the story. This is correct because Holmes’ London is Victorian London; lacking information to the contrary, we can therefore import into our reasoning the actual truth that during the Victorian age there were no aeroplanes or other means that could allow one to move from Australia to London in one day. The set of worlds accessible by means of the representational relation must be somehow be appropriately constrained on the basis of the background information against which we read the fiction.

How does the QCP square with the formal machinery? In naïve and nuclear Meinongianism any condition (or any nuclear condition) characterized some object, but

---

<sup>18</sup> In fact, one may also define logical consequence as truth preservation at all possible worlds in all interpretations: the semantics sketched contains nothing to differentiate  $@$  from any other world  $w$  in this respect, insofar as  $w \in P$ , that is,  $w$  is possible (we are into what follows logically from what, therefore, is what follows from what at the worlds where logic is *not* different).

with no proviso on worlds at which the object had the relevant properties. In particular, for any (nuclear) condition  $\alpha[x]$ , calling  $o$  an object characterized by the condition (at the actual world), we could have (using our newly introduced notation)  $@ \models^+ \alpha[o]$ , and therefore,  $@ \models^+ \Sigma x \alpha[x]$ . Precisely this kind of literalism spelled troubles for traditional Meinongians, as we have seen above.

In our modally qualified framework, on the contrary, if an object  $o$  is characterized by  $\alpha[x]$  (at the actual world), then in general we have that  $@ \models^+ \Psi \alpha[o]$ . Given the semantics of the representation operator, this entails only that, for all worlds  $w$  that realize the way things are described as being at  $@$ , that is, for all  $w$  such that  $@Rw$ ,  $w \models^+ \alpha[o]$ . As prescribed by the QCP, objects do have the features they are actually described as having, but not *at*  $@$ : they have them only at the  $R$ -accessible worlds (which may be possible or impossible ones) that realize such descriptions. Sherlock Holmes is described (at  $@$ ), by Doyle and by the Sherlock Holmes fans and Doyle readers engaged in internal fictional discourse, as being a detective and (let us assume) kicking Moriarty. He does not have these properties at  $@$ , though – in particular, because being a detective and kicking someone appear to be existence-entailing properties, whereas, at  $@$ , Holmes does not exist. He has them, though, at the worlds that realize the Doyle stories.

## 8. The niceties of MMM

Let us see, now, how MMM fares with respect to the issues raised by the examination of the mainstream theories of fictional objects above.

First, the issue of *negative existentials* has a totally straightforward treatment, just as in traditional Meinongianism: “Holmes does not exist” expresses the literally true statement that Holmes is (at the actual world  $@$ ) a nonexistent object.

Second, *all* sentences concerning Holmes, whether they belong to external or to internal fictional discourse, refer to one and the same thing: Holmes. “Sherlock Holmes” is a rigid designator across worlds. Apparently true sentences belonging to external discourse, such as “Holmes is a fictional character”, or “Holmes is a typical epitome of the Victorian age”, need no paraphrase at all: they can be taken at face value, and they are literally true. They can be taken as such, for being a fictional character, or being an epitome, are not existence-entailing properties: Holmes, as a nonexistent object, can actually have them at this world.

Objects that do not exist (at  $@$ ) may bear (at  $@$ ) several other properties that do not entail existence. Holmes actually has such logical properties as the one of being self-identical; he has converse intentional properties such as the one of being thought about by many fans; and still other intentional properties, such as the one of being more famous than any actually existing detective.

The situation is different for internal fictional discourse, which has to be (minimally) paraphrased. Sentences belonging to internal discourse are taken as true or false *according to* the fiction. The problems of literalism connected to ordinary Meinongianism are thus

avoided. The QCP is formulated with reference to the contexts, that is, the worlds, at which objects have the properties they are characterized as having by the stories describing them.

In this respect, MMM comes closer to Fictionalism. Just as happens both with Realism and with Fictionalism, our intuition that such sentences as “Sherlock Holmes is a detective” are *in some sense* true is explained by claiming that they are true according to such and such fiction (Doyle’s stories), that is, at the worlds that make the fiction true. “It is represented in such-and-such story that  $\alpha$ ” is true at  $@$  iff  $\alpha$  is true at the worlds  $w$  that realize the story, that is, at the worlds  $w$  such that  $@Rw$ , where  $R$  is the appropriate  $R$ -accessibility relation. By following this path, MMM can also do justice to the remark that sentences of internal fictional discourse are not actually asserted, e.g., by the authors: for to assert  $P$  is to posit  $P$  as true at the *actual* world. Doyle didn’t want to claim that Holmes actually is a detective; neither do we, when we are reporting the content of his stories.

However, notice the differences with respect to both Realism and Fictionalism. First, although  $\Psi$  has been introduced as a sentential operator, we assume that the inference *de dicto - de re* is always licensed. When one says “It is represented in Doyle’s stories that Sherlock Holmes is a detective”, we are always enabled to infer that one is actually (and truly) claiming, *of* the nonexistent (at  $@$ ) Holmes, that he has the property of being represented in Doyle’s stories as a detective. And Holmes *is* a detective – at the worlds that realize Doyle’s stories. This may be settled formally, for instance, by adding to the semantics some clause for  $\lambda$ -abstraction: from  $\Psi D(b)$ , infer that  $[\lambda x \Psi D(x)]b$ . The element of pretence is present in the form of Holmes’ not actually having the property of being a detective, that is, not being a detective at  $@$ . However, this is *real*, not fictional, *reference* to the object Holmes: it is of the nonexistent Holmes that one claims that, at the worlds or in the situations that realize Doyle’s stories, he is a detective. This solves the Millian problem: the proposition actually embeds Holmes as a constituent; the sentence expressing the proposition contains no terms lacking reference, and sentences in which it is embedded express complete propositions in their turn.

Next, “Sherlock Holmes” denotes the same object in both internal and external fictional discourse. It is the very same object which is the bearer of the name “Sherlock Holmes”. That unique object, as represented in Doyle’s stories, is a brilliant detective with amazing powers of observation and deduction, living in 221b Baker Street. At  $@$ , it is a fictional character and a nonexistent object, more famous than any real detective. Holmes has the former features at the worlds that make Doyle’s stories true, the latter ones at the actual world.

And no cross-category ontological problems arise because of this. We don’t have the unwelcome consequence of Realism plus *de re* pretence for internal discourse, namely that internal discourse pretends to ascribe properties of typically concrete objects, such as being a detective and living in Baker Street, to fictional characters that actually are abstract objects. *The Hound of the Baskervilles* is not a sequence of propositions mostly about abstract entities; it is a sequence of propositions mostly about objects that have the relevant properties at the relevant worlds.

But even non-merely-fictional objects that figure in fictional stories receive an intuitive account in the MMM framework. There is no need to treat “Napoleon” as

ambiguous – (a) usually referring to the concrete, real man, but also (b) referring to an abstract entity completely different from him when it occurs in external fictional discourse, and (c) (perhaps) referring to nothing at all when it occurs in internal fictional discourse. According to MMM, “Napoleon” always denotes, and denotes the very same object in all these contexts. This is simplicity itself. For (a) when we find in a book of modern history that Napoleon was defeated at Waterloo, this is literally true of the unique Napoleon – true in the actual world. (b) When critics, or people engaged in external discourse, say that *War and Peace*’s Napoleon is a fictional character, they are referring to the same individual, and they claim something which is literally true, true of the actual world. Of course, this does not entail that Napoleon is a purely fictional, nonexistent character by itself: *being a fictional character* is not an existence-entailing property, but Napoleon may well exist or have existed (at the actual world) for different reasons. And he does: he’s an actual, historically existed man. (c) When “Napoleon” occurs in internal discourse, such as the one constituted by the sentences composing *War and Peace*, the properties ascribed to Napoleon are instantiated by him *at* the worlds that realize Tolstoj’s description. Some of those properties, e.g., that of being the self-proclaimed emperor of France, he may also have, or have had, at the actual world. Some others, such as (let us assume) being pompous, he may have only at those worlds, not at the actual one.<sup>19</sup>

Or take Virgil, who is both a historical, real poet, and a fictional character in Dante’s *Comedy*. It is still the same Virgil that we are talking about: at the actual world, he has (had) the proprieties ascribed to him by the historians (to be sure, insofar as they got it right); he (the very same Virgil) has different properties at worlds different from the actual: at the worlds that make the story told in Dante’s poem true, he has *those* properties and features.

Or take another difficult case, involving cross-contextual ascriptions:

- (1) Shakespeare’s Caesar is more vigorous than the historical Caesar, and less ridiculous than Asterix’s Caesar.

Easily enough, “Caesar” in (1) denotes the same individual in all its three occurrences. And (1) is true if and only if Caesar (*that* unique guy) is overall more vigorous at the worlds that make Shakespeare’s novel true than at the actual world and, at those worlds, is less ridiculous than at the worlds that make true Asterix’s comics.

---

<sup>19</sup> Does the distinction between existence-entailing properties and not existence-entailing ones perfectly match with the one between properties ascribed, respectively, in internal and external fictional discourse? Well, not quite. Internal fictional discourse can ascribe to (purely) fictional characters properties that are not existence-entailing; for instance, Doyle may have ascribed to Holmes, in one of his stories, the property of being thought about by Moriarty. But also, internal fictional discourse can ascribe properties that do not entail existence to characters that do not exist even at the worlds that realize the story. For instance, we may write a novel about a girl, Ricky, who falls in love with a fictional character, Prince Roland, described in a book she is reading. Not only Prince Roland does not exist at the actual world (as well as the fictional book Ricky is reading); but Roland does not exist at the worlds that make our story true either (whereas the book Ricky is reading does). At those worlds, Roland is a nonexistent, purely fictional object Ricky falls in love with. So several properties that do not entail existence, such as the one of being a fictional character, can be ascribed to him in the context of the internal fictional discourse.

Furthermore, a name may refer to a nonexistent (at the actual world @) even when the speaker intends to refer to an existent. So even the “Vulcan” case has a straightforwardly simple and intuitive account. Babinet and Leverrier mistakenly believed that, by means of the genuinely singular term “Vulcan”, they were referring to an existent object. Now we do not need to engage in the epicycle of saying that the mistaken Vulcan theory actually originated an abstract object to which, unbeknownst to Babinet and Leverrier, the name “Vulcan” ended up referring. According to MMM, “Vulcan” has no semantic ambiguity. “Vulcan” unambiguously refers to an object that, at the actual world, is nonexistent. That very object does have the properties it is characterized as having by the Babinet-Leverrier theory (by the theory Leverrier and Babinet believed to be true of the actual world), including being a planet and existing. But not at @. The object named “Vulcan” has those properties at the worlds that make that theory true. Babinet and Leverrier believed *of* the (actually nonexistent) Vulcan that it actually existed, and thereby referred to *it*. Their mistake consisted in believing that the actual world was one of those worlds. Objects always have the properties they are described as having – this is settled *a priori* via the QCP. What cannot be known *a priori* is *at which* worlds they have them; for we cannot know which world is ours, the actual one, *a priori*, via a mere conceptual exploration of the realm of possibilities.

What about other troublemakers, such as the cross-fictional sentence “Sherlock Holmes is taller than Bilbo Baggins”? Here one needs to massage our intuitions a bit. But the explanation is, we believe, still quite intuitive, not violating the Acceptability Constraint. Being nonexistent objects at the actual world @, both Sherlock Holmes and Bilbo Baggins lack any height here (otherwise, they would occupy some spatiotemporal region at @, so arguably they would exist here). But one who claims that Holmes is taller than Bilbo is, quite obviously, comparing the heights they have *according to* the respective stories. And the issue whether the claim is true can be settled by means of background information, even though (let us suppose) neither Doyle nor Tolkien provided precise characterizations of their characters’ heights in the respective stories. This is one of the cases in which the relevant representations (and therefore, worlds) import background information from actuality. For instance, one may argue that Holmes is characterized by Doyle as a perfectly normal man in this respect, whereas Bilbo is characterized by Tolkien as a typical Hobbit. Now Hobbits are also characterized in Tolkien’s saga as very short, compared to men, and usually even smaller than dwarfs. This is how we explain our intuition that the claim that Holmes is taller than Bilbo is true.

## 9. Conclusion

MMM is so far just the sketch of a full-fledged theory of fictional objects; the modal semantics employed in its description is non-standard, and its applications are largely untried. Therefore, we would not be surprised to see alleged counterexamples and difficult cases forthcoming soon. However, given the evidence collected, we claim that our revised Meinongian theory may fare better than any other theory of fictional objects proposed so

far with respect to the methodological maxim of Minimal Revision, and the Acceptability Constraint. It agrees with Fictionalism, Realism, and traditional Meinongianism to the extent that they stick to our intuitions and ordinary discourse on fictional objects, while it avoids at least some of the traditional difficulties and counterintuitive consequences of each of the three approaches.

## References

- Barwise J. [1997], "Information and Impossibilities", *Notre Dame Journal of Formal Logic* 38: 488-515.
- Beall J.C. and B. van Fraassen [2003], *Possibilities and Paradox. An Introduction to Modal and Many-Valued Logic*, Oxford UP, Oxford.
- Berto F. [2009], "Impossible Worlds", *The Stanford Encyclopedia of Philosophy*, CSLI, Stanford-CA <http://plato.stanford.edu/entries/impossible-worlds/>.
- Berto F. [forthcoming], "Impossible Worlds and Propositions: A Case Against the Parity Thesis", *The Philosophical Quarterly*.
- Currie G. [1990], *The Nature of Fiction*, Cambridge UP, Cambridge.
- Fine K. [1982], "The Problem of Non-Existents, I: Internalism", *Topoi* 1: 97-140.
- Fine K. [1984], "Critical Review of Parsons' *Nonexistent Objects*", *Philosophical Studies* 45: 94-142.
- Griffin N. [1998], "Problems in Item Theory", talk given at the Australasian Association for Logic.
- Inwagen P. van [1977], "Creatures of Fiction", *American Philosophical Quarterly* 14: 299-308.
- Inwagen P. van [2003], "Existence, ontological Commitment, and Fictional Entities", in Michael Loux and Dean Zimmermann (eds.), *The Oxford Handbook of Metaphysics*, Oxford UP, Oxford.
- Jacquette D. [1996], *Meinongian Logic. The Semantics of Existence and Nonexistence*, De Gruyter, Berlin-New York.
- Lewis D. [1978], "Truth in Fiction", *American Philosophical Quarterly* 15: 37-46.
- Linsky B. and E. Zalta [1994], "In Defense of the Simplest Quantified Modal Logic", *Philosophical Perspectives* 8: 431-58.
- Lycan W. [1994], *Modality and Meaning*, Kluwer, Dordrecht.
- Nolan D. [1998], "An Uneasy Marriage", talk given at the Australasian Association of Philosophy.
- Parsons T. [1979a], "Referring to Nonexistent Objects", *Theory and Decision* 11: 95-110.
- Parsons T. [1979b], "The Methodology of Nonexistence", *Journal of Philosophy* 76: 649-62.
- Parsons T. [1980], *Nonexistent Objects*, Yale University Press, New Haven, Conn.
- Priest G. [1992], "What is a Non-Normal World?", *Logique et Analyse* 35: 291-302.
- Priest G. [1997], "Editor's Introduction", *Notre Dame Journal of Formal Logic* 38: 481-7.
- Priest G. [2005], *Towards non-Being. The Logic and Metaphysics of Intentionality*, Oxford UP, Oxford.

- Rantala V. [1982], "Impossible World Semantics and Logical Omniscience", *Acta Philosophica Fennica* 35: 106-15.
- Restall G. [1997], "Ways Things Can't Be", *Notre Dame Journal of Formal Logic* 38: 583-96.
- Routley R. [1980], *Exploring Meinong's Jungle and Beyond*, Australian National University RSSH, Canberra.
- Routley R. [1982], "On What There Is Not", *Philosophy and Phenomenological Research* 43: 151-78.
- Russell B. [1905a], Review of A. Meinong, *Untersuchungen zur Gegenstandstheorie und Psychologie*, *Mind* 14: 530-8.
- Russell B. [1905b], "On Denoting", *Mind* 14: 479-93.
- Salmon N. [1998], "Nonexistence", *Noûs* 32: 277-319.
- Thomasson A. [1999], *Fiction and Metaphysics*, Cambridge UP, Cambridge.
- Thomasson A. [2001], "Ontological Minimalism", *American Philosophical Quarterly* 38: 319-31.
- Thomasson A. [2003], "Speaking of Fictional Characters", *Dialectica* 57: 205-23.
- Walton K. [1990], *Mimesis as Make-Believe*, Harvard UP, Cambridge, Mass.
- Walton K. [1991], "Précis of Mimesis as Make-Believe", *Philosophy and Phenomenological Research* 51: 379-82.
- Walton K. [2003], "Restricted Quantification, Negative Existentials, and Fiction", *Dialectica* 57: 239-42.
- Wolstertorff N. [1961], "Referring and Existing", *The Philosophical Quarterly* 11: 335-49.
- Wolterstorff N. [1980], *Works and Worlds of Art*, Clarendon Press, Oxford.
- Zalta E. [1983], *Abstract Objects. An Introduction to Axiomatic metaphysics*, Kluwer, Dordrecht.