Organon F 24 (3) 2017: 339-365

The Truth about Sherlock Holmes

FREDRIK HARALDSEN[1]

ABSTRACT: According to *possibilism*, or non-actualism, fictional characters are possible individuals. Possibilist accounts of fiction do not only assign the intuitively correct truth-conditions to sentences in a fiction, but has the potential to provide powerful explanatory models for a wide range of phenomena associated with fiction (though these two aspects of possibilism are, I argue, crucially distinct). Apart from the classic defense by David Lewis the idea of modeling fiction in terms of possible worlds have been widely criticized. In this article, I provide a defense of a possibilist account against some lines of criticism. To do so, I assume that names for fictional characters are directly referential and a possible-worlds model that accommodates transworld identity. On this background, I argue, it is possible to construct an elegant model of fictional discourse using familiar models of information exchange in ordinary discourse, and I sketch how this model can be used to i) make a natural distinction between fictional and counterfactual discourse, ii) account for creativity, and iii) sustain a natural definition of truth-in-fiction that avoids certain familiar objections to possibilism. Though I set aside questions about the metaphysical commitments of a possible-world interpretation here, there is accordingly reason to think that the battle over possibilist treatments of fiction will have to be fought over metaphysical foundations rather than technical shortcomings.

KEYWORDS: Creativity – fiction – non-actualism – possibilism – truth-in-fiction.

[1] Received: 20 April 2017 / Accepted: 20 July 2017

✉ Fredrik Haraldsen
SASPRO/Marie Curie Fellow
Institute of Philosophy
Slovak Academy of Sciences
Klemensova 19, Bratislava 831 64, Slovakia
e-mail: fredrikhar@yahoo.no

© 2017 The Author. Journal compilation © 2017 Institute of Philosophy SAS

1. Introduction

Realists about fictional characters are committed to an ontology of fictional characters. A benefit of realism is the ability to make relatively straightforward sense of talk about such characters; though 'Trump is human' and 'Sherlock Holmes is human' differ in truth-value, the sentences seem *prima facie* to have the same kind of semantic content, and if 'Trump' refers to Trump and 'Trump is tall' expresses a singular proposition whose truth-conditions depend on how things are with Trump, one may expect the same to apply to 'Holmes'. *Prima facie*, then, 'Holmes' refers to Holmes, and sentences involving Holmes express singular propositions that are assigned truth conditions the same way sentences containing 'Trump' are.

Realists disagree about the nature of these characters, however. According to *possibilism*, or non-actualism, fictional characters are *possible objects*. Although Holmes does not actually exist, he could have, and does exist in some possible world different from the actual one. Moreover, we understand sentences containing 'Holmes' at least in part by grasping what things are like in worlds relative to which those sentences are true. In the present article I provide a partial defense of possibilism. However, to do so we should distinguish two distinct goals a possibilist analysis could aim to achieve. First, a possibilist analysis might aim to provide a semantic theory for sentences in fiction, one that assigns intuitively correct truth-conditions to such sentences. Now, I will assume that sentences containing 'Holmes' express singular propositions and that 'Holmes' directly refers to the merely possible object Holmes, who exists in multiple worlds where he instantiates different properties. A challenge to this view that I briefly discuss below is to explain how we can refer to the right individual (see Kripke 2011; Thomasson 1999). To provide a semantic theory that identifies truth-conditions for each sentence in a fiction, however, we need not worry for instance about what counts as true in a fiction or whether there is a set of relevant worlds that can serve as truth-makers for the *whole* fiction.

A second goal is to develop a model of fiction as a type of *discourse*, that can help explain how we interpret, and share interpretations of, fictional narratives. Doing so requires for instance accounts of what counts as being the case in a story, what distinguishes storytelling from counterfactual discourse and how possible worlds can be used to reflect the *creative*

aspect of storytelling. The primary aim of this article is to sketch a framework that can support the second goal. I do not intended to offer a comprehensive theory of literary interpretation, of course; questions about the role of the narrator or rhetorical devices such as irony or metaphor, for instance, are beyond the scope of this article and probably beyond what we can expect to model with possible worlds alone. Rather, my goal is to identify some basic tools that can later be supplemented in various ways.

To provide such a framework we need an account of what it takes for a claim to count as *true* in a fiction. A first stab may be:

(1) *In fiction f, 's' is true* iff *in the (relevant) possible worlds reflecting f, s*.

Two points are worth making here. First, we need to specify what it takes for worlds to be *relevant*. For the moment, think of the relevant worlds as the *closest* ones in which the events described in *f* take place. I return to this issue below. Second, (1) assumes that utterances of sentences such as 'Holmes is a detective' are (in the relevant contexts) elliptical for 'in the fiction, Holmes is a detective'. We could, however, interpret such sentences as non-elliptical and literally false but rather understand fiction talk for instance as engendering a shift in the *context of assessment* (following e.g. Kölbel 2002; MacFarlane 2005). Either option will in principle work for my purposes.

Possibilism does not only predict intuitively correct truth-conditions for sentences containing names for fictional characters, but in many cases also their intuitively correct *truth-values*, and does so in many ways better than other realist theories – Meinongian views, Platonism or abstract artifact theories – or irrealist views. Possibilism predicts that 'Holmes does not exist' and 'there is no Holmes' are true, insofar as existence ascriptions involve implicit restriction to the actual world. 'Holmes is a detective' has both a true and a false reading, depending on the presence or scope of an 'in-the-fiction' operator. Moreover, possibilism seems able to account for issues such as embedded fictions, intertextual comparisons ('Holmes is smarter than Poirot'), meta-fictional claims ('Holmes is a fictional character') and even make sense of speculations about characters' motivations or psychology, also when these are not explicitly detailed, insofar as there may well be a fact of the matter in the closest worlds where a character acts

as described in the fiction. Possibilism even gives us a basis for explaining how it is possible to *learn* from fiction – not only facts about the actual world related in the fiction, but about morality or the human condition. Of course, I will not have space to discuss all these features in detail here. But even the apparent ease with which possibilism can accommodate the intuitive data is some evidence in its favor.

There are, however, important questions I will set aside here; I mention some of them in Section 2. In Section 3 I sketch a possibilist model of fiction and explain how it accommodates i) creativity and ii) the distinction between fictional and counterfactual talk. In Section 4 I sketch and defend a promising theory of *truth-in-fiction*.

2. Referring to Holmes

There are several issues I do not have room to discuss. I merely note some of them here:

1) To provide a theory of truth for fiction it is natural to assume that we need possible worlds to be fairly metaphysically robust. To what extent we are committed to *modal realism* is a question I leave open.

2) I will assume that names are directly referential. A commitment to direct reference sits uncomfortably with a counterpart theory of modality, which appears to require some form of descriptivism (see Sullivan 2005, for discussion). A counterfactual claim about Aristotle must, on a direct reference view, be a claim about *Aristotle*, and not his potentially descriptively indistinguishable counterparts (on pain of a change in *semantic content*). To determine whether 'Aristotle could have failed to teach Alexander' is true, then, we don't search the space of worlds to discover one in which an identifiable Aristotle fails to teach Alexander. Rather, we consider what the case is with *Aristotle*, and evaluate the counterfactual claim by determining whether *he* is part of a world where he doesn't teach Alexander. On direct reference views, reference precedes the identification of relevant worlds. Accordingly, Aristotle must exist in multiple worlds, and the relevant worlds must contain the numerically same individual Aristotle. And since we want the semantic content of names in fiction to be of the same type as ordinary names, 'Holmes' should be directly referential as well. As a consequence, we reject counterpart theories of modality in our

model of fiction. Possibilism is not wedded to a counterpart theory, however. If we need robust modal realism, then transworld identity of individuals may be ensured by overlapping worlds or by adopting some form of modal dimensionalism.

3) For possibilism model fiction, there must be possible worlds containing magic, ghosts, dragons and people waking up as beetles or cockroaches. Widespread *essentialism* or proliferation of metaphysical necessities would be a problem for such accounts of fiction by limiting the kinds of stories authors could coherently tell even among stories that appear *prima facie* to be coherent. I will assume that if a story can be apparently coherently told in which an object of type *A* is described as being *not-B*, this is at least evidence that there are worlds where *A*s are *not-B*s.

The extent to which even widespread essentialism is an obstacle may be a function of what commitments our account makes to the metaphysics of modality. A constructivist or ersatzist approach *might* allow us to sidestep essentialist worries even consistently with accepting a wide range of metaphysical necessities: to provide the truth-conditions for sentences in fiction, we need the space of worlds to contain the logically (and analytically) possible worlds, but just as fictions are free to change laws of nature, fictions are free to suspend *metaphysical* necessities that are not also logical or analytical necessities and still be able to model the semantic contents of sentences correctly. If x is essentially human, x is human in all metaphysically possible worlds, but 'x is not human' is not *logically* or analytically impossible; we can construct or describe logically and analytically consistent worlds in which x is not human, and stipulate that these are as relevant for our purposes as the metaphysically possible ones.

4) We may, as mentioned above, legitimately worry about how we can *refer* to merely possible objects. To defend possibilism as a model of fictional discourse, I need to say something about reference, and in the rest of this section I sketch a response to this worry (a complete answer will have to be postponed).

Whenever I utter a sentence containing a name I succeed in referring presumably in virtue of standing in the right sort of relation – perhaps a causal-historical relation – to the introduction of the name, and this relation is plausibly the same for my use of 'Holmes' as for my use of 'Caligula'. But whereas we can tell an obvious story about how 'Caligula' got attached to Caligula in the first place, it is less clear how Doyle attached 'Holmes'

to the merely possible object Holmes when introducing the name. Doyle cannot have established the reference of 'Holmes' by a baptismal act where the name was bestowed upon a demonstratively available object, since merely possible objects are not demonstratively accessible. Now, if demonstration were necessary for fixing reference, then a number of names for objects also outside of fiction would fail to refer, including names for abstract objects, individuals hypothesized to exist, or unobserved astronomical objects whose existence is inferred by calculation. In these cases, however, we can individuate referents as the individuals that uniquely satisfy certain descriptions. But descriptive individuation also seems unavailable in the Holmes case. Many merely possible objects – perhaps even some (or all) actual ones – satisfy the descriptions associated with 'Holmes' in the stories in different worlds. But insofar as 'Holmes' is directly referential – and thus *rigid*, picking out the same object in every world in which it exists – it must pick out *one* of them on pain of contradiction.[2]

Note that for descriptivists about the semantic contribution of proper names this kind of worry would not arise; for them, the contribution of 'Holmes' is a descriptively individuated *role* that can be satisfied by numerically different objects in different worlds without affecting the semantic content of sentences containing the name. But since we reject descriptivism, we need to know i) which object, among all suitable candidates, is Holmes; and ii) how Doyle managed to refer to it. I think the correct answer to the first question is that *it doesn't matter*. To be sure, we need 'Holmes' to refer to a single one among the merely possible referents, but it makes no difference *which one*. As for the second question, I suggest that when Doyle started writing about Holmes (or dreamt up the character), his implicit decision to describe a fictional character sufficed to ensure that some merely possible object was assigned, arbitrarily selected from among those that satisfy whatever properties – perhaps none apart from the name – Doyle associated with it at the time. On the surface, the stipulative introduction of 'Holmes' resembles a baptism, but instead of having a referent

---

[2] Suppose Holmes = α in world w_1 and Holmes = β in w_2 and $\alpha \neq \beta$. Since 'Holmes', by assumption, is rigid, 'Holmes' refers to Holmes in both w_1 and w_2, and the transitivity of identity gives us a contradiction.

demonstratively available, Doyle assigned 'Holmes' to *some (possible) individual or other* – the selection domain being restricted perhaps (only) by the condition that it be non-actual – that he could later develop to serve whatever purpose he wanted in the stories.

It is sometimes thought to be required for reference that there is some property that *o* must have for a name *n* to refer to *o* that can, in a non-circular manner, be invoked to distinguish *o* from other potential referents – either through a universally applicable condition such as *acquaintance* or a *causal relation* between name tokening and referent, or that *some* such condition must be satisfied in any particular instance, even if different relations – satisfying descriptive conditions or causal links – may do the trick in different cases. Of course, according to my suggestion there *is* a property Holmes has – *being arbitrarily selected as the referent of 'Holmes'* – in virtue of which the relation holds, but this property cannot be specified without invoking 'reference' (or a notion depending on *reference*, such as 'pick out' or 'select'). So, there is no non-circularly specifiable property in virtue of which 'Holmes' refers to a particular individual. Does it matter? There are at least two (compatible) lines of response available to the worry that it does.

First, one may deny that such a (non-circular) property is necessary for reference. A no-criterion theory of reference has been defended for names in general (cf. Breckenridge & Magidor 2012), and albeit controversial, I think that approach is more obviously acceptable in the special case of names for fictional characters. After all, given that the properties Doyle associated with Holmes at the introduction weren't contradictory, he was guaranteed that there is an appropriate individual, that the selected referent wasn't the *wrong* one, and that he would not go on to describe it incorrectly. There was no *unique* candidate for 'Holmes', but it is not entirely obvious that *unique satisfier* (across all worlds) is a better criterion for individuating referents than *arbitrary selected individual among numerous satisfiers* in the context of fiction. Second, what ultimately matters is not whether the relationship between 'Holmes' and the arbitrarily selected object is *a reference relation* or something else – perhaps *schmeference*, a relation between name and referent characteristic for individuating characters in fiction. What matters is that we can associate singular propositions with sentences containing 'Holmes' that have the truth-conditions they are supposed to have and play the role in our

theories of meaning and intentionality[3] they are supposed to play. This is something singular propositions containing arbitrarily selected, possible individuals seem able to do even if the relation between name and object is schmeference rather than reference.

The idea, then, is that when introducing the Holmes, Doyle selected a merely possible object to serve as referent for 'Holmes' among numerous potential candidates. Given transworld identity, the arbitrarily selected referent exists in multiple worlds where he instantiates different properties. Presumably, he is a barrel maker in some of these worlds, a detective in others, and a future space-traveler in yet others. When making the initial selection, Doyle didn't select any particular world; if he had not yet even decided that Holmes was going to be a detective when he started writing about him, Doyle had not yet ruled out the barrel-maker worlds. As long as the selected individual satisfies the properties Doyle ultimately came to associate with it in the stories in *some* world in which it exists, Doyle selected the right one.

To *guarantee* that Doyle selected the right one, then, we need to ensure that the selected referent *could* play the role of Holmes. I suspect that *any* merely possible object can do so, in which case we get the guarantee for free. This claim may be challenged. In particular, if many properties are essential properties of any individual that instantiates them, the guarantee would be harder to obtain. If *being human* is an essential property of any human, then Doyle must have decided to select a referent among possible individuals that are human. Having decided on species membership is perhaps not an unreasonable condition for making a correct selection, but any further properties taken to be essential makes the guarantee consecutively harder to obtain; if, say, *height* were an essential property, my account would be in trouble if Doyle initially selected someone shorter or taller than he eventually described Holmes as being. As mentioned above possibilism

---

[3] Questions of reference have, of course, both a semantic side and a cognitive side. I am here concerned with the former and have little to say about the latter. I assume, however, that *singular thoughts* are singular in virtue of their *form*, and hence that one can entertain a singular thought even if there is no referent the thought is, in fact, about. If this is correct, then there should be no in-principle obstacle to having a singular thought about an object to which the relation is schmeference rather than reference.

sits awkwardly with a proliferation of metaphysical necessities, and the account becomes less elegant if many properties are essential properties. We can, however, account – at some cost – for moderate degrees of essentialism; I return to this below. It is worth mentioning that full-blown *haecceitism* would *not* be a challenge to my account, however. If fictional characters possess individual essences, like *being Holmes*, then the problem of securing reference would not arise – in that case Holmes just is the unique individual in modal space that satisfies the description 'the individual that *just is* Holmes'.

3. The Model

According to a common model of information exchange (especially Stalnaker 1978) the content of a discourse contribution removes worlds incompatible with the content of that contribution from a *context set* of worlds reflecting the information presupposed or shared by the discourse participants (that is, the content intersects with the context set). This model can be adapted to fiction. (I leave it open whether we understand it as a model of the unfolding narrative or of Doyle's construction of the story – these may look rather different.) Now, Doyle, having decided to write about Holmes, can be taken to have individuated a context set containing all worlds in which the individual selected to be Holmes exists. Assuming that Doyle had not yet decided that Holmes was a detective or determined his height, the set contains worlds in which Holmes is not a detective or has different heights. Presumably, by Doyle's intention to engage in fiction and introducing Holmes as a fictional character, the *actual* world is not a member of the context set.

Having determined a context set, Doyle proceeded, over the course of the stories, as in ordinary discourse, to eliminate worlds from that set, starting with worlds where Holmes is not a detective, didn't live in London, and so on.[4] When introducing new characters, new reference selections are needed. Eligible candidates are those that populate scenarios in which Holmes could have taken part, and must thus inhabit worlds in the initial

[4] My presentation, to make the exposition easier to follow, does not reflect the actual Holmes storyline.

context set. So, for any eligible Moriarty candidate there must, at the outset, be a world in which Holmes interacts with that candidate as he satisfies the descriptions associated with Moriarty. The candidates are potentially any possible individual (at least among those not yet selected for other roles), and there is accordingly, in the initial context set, for any such individual *x* a world in which Holmes interacts with *x* playing the role of Moriarty. As for Holmes, when Doyle started writing Holmes may have been maximally incomplete, and every (contingent) property such that Holmes instantiates that property in some world. The initial selection, however, ensures that it is the numerically same referent that has these different properties in different worlds. Moreover, Holmes was never completely described; Doyle never specified Holmes's blood type, for instance. Incompleteness means that we are left with several worlds in which the same individual has different blood types, not different worlds with different referents.

As in ordinary discourse, certain updates of the context set are more difficult to accommodate. For instance, assume for the sake of argument that when starting to write *Fight Club*, Chuck Palahniuk had not decided that Tyler Durden and the narrator (the referent of 'I') were the same. At the outset, then, he selected different referents for the expressions, and, given the necessity of identity, there is no world in the context set in which they are the same. When subsequently deciding that Durden and I are, indeed, identical, Palahniuk needed to make a new selection. He would accordingly have to reinterpret everything in the narrative thus far relative to a new context set defined by the new referent. This consequence does not strike me as a failure of the model but rather as reflecting what actually happens when authors make momentous plot decisions (deciding that names for lesser characters in fact co-refer requires less backtracking). However, as the case suggests, widespread essentialism would make the model somewhat less elegant by making backtracking and reselection far more frequent and often required in cases where the author would be unaware that it is needed. As mentioned above, the plausibility of a possibilist analysis of fiction does to some extent hinge on apparently coherent scenarios described in fiction being genuinely *possible*.

In the next sub-sections I describe, first, how the model facilitates *creativity* and metafictional talk, and second how it captures the distinction between fictional and counterfactual talk. In Section 4 I provide an updated definition of *truth-in-fiction*.

3.1. Creating and describing characters

A common complaint raised against possibilist analyses is that they fail to capture the creative aspect of fiction making, turning it rather into a matter of *discovery*. Of course, when discovery is *discovery of possibilities*, the difference between discovery and creation is less than obvious. Consider the claim attributed – probably apocryphally – to Michelangelo that the sculpture was already complete within the marble block, and that he just had to chisel away the superfluous material. The description seems correct, if idiosyncratic, insofar as the resulting statue was, indeed, an already possible result of the process. Analogously, developing a story can be modeled as shrinking the context set, and developing a character as removing worlds by deciding which properties the character should have among those it could have had, to individuate a scenario that was, after all, already a possible outcome of the creative process.

Of course, Michelangelo's description neglects the crucial point that he *chose* which of many possible statues he was going to carve out, which is presumably essential to the process being *a creative process*. Similarly, freedom of choice is a distinctive feature of fictional storytelling, and Doyle was crucially free to narrow down the context set any way he wanted: As opposed to ordinary factual discourse, where the target is the actual world, Doyle freely chose his target world(s) among those available at any given point, and freely chose which steps to take to individuate it/them. Of course, possibilists will have to grant that in some sense Holmes was already *there*, and thus contradict the creationist intuition that Doyle really *brought Holmes into being*. But our intuitions that Holmes really is *brought into being* are shaky at best, and it is reasonable to wonder whether intuitions here are driven by prior metaphysical commitments rather than vice versa. At least I think the onus is on the critic to explain precisely what's missing from the possibilist account – keeping in mind it seems hard to deny that the Holmes stories are stories that could possibly have been told also prior to Doyle telling them.

Now, Doyle literally brought things like the narrative, name and features of the *representation* of Holmes into (actual) being, and thereby affected the causal structure of the actual world. The distinction between representations and scenarios represented must be managed with some care, however. The *stories* – at least the words and sentences – exist in the actual

world. Claims like 'in *Wozzeck* the eponymous protagonist is described as driven to murder by the oppressive mechanisms in his society' or (arguably) 'the plot of *L'Étoile* is very complicated' are claims at least partially about the actual world and aspects of our representations rather than the possible objects represented.

Keeping this distinction in mind, our account does not only account for in-fiction claims, but also common meta-fictional or critical claims. Consider the following (adapted from Brock 2002, 4-5):

(2) Holmes does not exist.
(3) Holmes is a fictional character.
(4) Anna Karenina is less neurotic than is Katerina Ivanovna (Howell 1979).
(5) Holmes would not have needed tapes to get the goods on Nixon (Lewis 1978).
(6) The character Odysseus who occurs in the *Odyssey* is identical to the character who occurs in *Inferno*, Canto 26, under the name 'Ulysses' (Howell 1979).
(7) There are characters in some nineteenth-century novels that are presented with a greater wealth of physical detail than is any character in any eighteenth-century novels (van Inwagen 1977).

(2) is straightforwardly true, as is (3) – *being a fictional character* entails *being a merely possible (non-actual) individual*, so (3) truthfully says (at least) that Holmes is a merely possible individual. (4), too, is true assuming that we measure their neuroticism based on the characters' appearances in the relevant stories. For (5), start with the worlds that serve as truth-makers for the stories. In the closest possible worlds – or at least some close worlds – relative to those (which are not necessarily worlds that define the *context set* for the stories, but counterfactual scenarios *accessible* from worlds in that set) in which Holmes got the goods on Nixon, he didn't use tapes. (6) and (7) require some care, since they talk both about representations and scenarios represented. (6) is true iff 'Odysseus' and 'Ulysses' co-refer, which is certainly likely on our account (though it would, of course, depend on factors related to the introductions of the characters); the two stories mentioned may, of course, not describe the same worlds of fiction, but that is – given transworld identity – not necessary for the *referent*

to be the same. For (7), it is perhaps best to interpret ‘characters’ as denoting possible objects and the sentence as making a claim about how *those* are being presented in the novels.

Other cases may be trickier. ‘Holmes is famous’ arguably seems to predicate a property Holmes cannot have in a world in which he doesn’t exist, yet is ostensibly a claim about the actual world. However, the sentence probably requires paraphrase on any view; on abstract artifact theories, for instance, Holmes is an existing abstract object, and abstract objects are not detectives and thus intuitively the wrong targets for ordinary ascriptions of *fame*. I think a case can be made even at an intuitive level that the ascription requires paraphrase – Holmes does not *really* have the property of fame in the same way, say, Caligula does. A possibilist could at least say that he is much talked and thought about, so a suitable paraphrase doesn’t seem unachievable.

3.2. Fiction and counterfactuals

Since fictional characters are possible individuals, and fictional stories represent series of events in possible worlds, the difference between fictional and counterfactual discourse is not a matter of semantic content or the ontological status of the truth-makers for the discourse. There is, of course, intuitive support for this. Although conversational markers like ‘consider the following possibility’ and ‘once upon a time’ conventionally serve to signal whether what follows is a counterfactual claim or a piece of fiction, respectively, it seems surprising – to say the least – that this would affect the nature of the proposition expressed by or metaphysical commitments of a subsequent utterance of ‘there used to be life on Mars’.

That doesn’t mean that there is *no* difference. Counterfactual talk and fictional storytelling are distinguished by felicity constraints on utterances and rules for updating or managing one’s beliefs relative to new information. Counterfactual talk should be constrained by the requirement that the worlds of the context set be potential candidates for being the *actual* world; that is, counterfactual talk is ultimately about the actual world, and utterances of counterfactuals are felicitous only to the extent that they inform agents of *actual* states of affairs. Fiction is not so constrained.

The following suggestion makes the constraint more precise: Let the context set be a set W of worlds that contains (or is supposed to contain) the actual world. Contributions to a discourse are felicitous only if they narrow down the information consistent with the initial presuppositions, modeled as removing worlds from W. A counterfactual claim is accordingly felicitous only if it removes at least one world from W. How? Suppose p and q are true in all worlds $w \in W$. A counterfactual claim u to the effect that 'if p weren't the case, q would not have been' asks us to consider what the case might be in the closest worlds w^* accessible from the various worlds w where p is not the case, and u removes from W all worlds w from which the closest accessible worlds w^* in which although p does not hold, q still does. For fictional discourse, on the other hand, contributions proceed in the manner of normal, non-counterfactual information updates (though it is of course possible to incorporate counterfactual talk in fiction, as when fictional detectives speculate, counterfactually, about possible explanations for observations at a crime scene).

Insofar as the difference is a matter of pragmatic felicity constraints defined by discourse goals I predict that the distinction between counterfactual and fictional discourse might sometimes be hard to draw – especially in the absence of explicit conversational markers. This seems correct. For a relatively mundane illustration, consider *examples* in (say) a law textbook. Such examples may involve relatively elaborate stage setting and even names for characters. Insofar as their purpose is to illustrate facts about the actual world they are reasonably interpreted as describing counterfactual scenarios, yet they often look and behave like brief fictions. According to my proposal the correct classification depends on what role the story in the example is supposed to play, not on its truth-conditional content.

4. Truth in fiction

In light of the model we can offer a definition of truth-in-fiction. Now, the proposition expressed by a sentence in a fiction (or its truth-conditions) is of course determined by the (unrestricted) complete set of possible worlds. What counts as *true-in-fiction*, however, is the subset that reflects the information the fiction conveys, and which is relevant to interpreting

it, for instance with regard to understanding (non-explicit) character motivations. In 4.1 I introduce a definition of *truth-in-fiction*, and in the following sections I consider some upshots.

4.1. Truth in fiction

Above we assumed that *p* is *true-in-fiction-f* iff *p* is true in the closest possible worlds in which *f* holds. I return to *closeness* below, and will for the moment just assume it. Two other points are worth making. First, the formulation differs from Lewis (1978); according to Lewis, *p* is true in worlds where *f* is *told as known fact*. Lewis introduced the 'told as known fact' qualifier to rule out the possibility that, by cosmic coincidence, the Holmes stories report actual events, thus making the actual world a potential truth-maker. The qualifier would complicate out reference fixing story by relativizing reference to a merely possible storyteller. Since we already rule out the possibility of the actual world being a truth-maker by Doyle's intention to refer to a merely fictional character, however, we do not need the qualifier.

Second, fictional characters are incomplete. Many properties, such as a Holmes's blood type, are never settled in the stories in which those characters feature. Possible individuals are complete. In every world where Holmes exists he has a determinate blood type. It should, however, not be *true-in-the-stories* that Holmes has, say, blood type A rather than B, even if he does in some worlds consistent with the stories. One option we may dismiss is that Holmes is literally an incomplete individual inhabiting an incomplete world. Although fictional characters are under-described, they are often also attributed traits that entail that they are complete. A character described as 'a normal person' has a particular blood type since normal persons have determinate blood types. Incompleteness is not *metaphysical incompleteness*. Accommodating incompleteness is straightforward on our account. If we follow Lewis and deny that there must be a uniquely closest possible world, then *incompleteness* just means that different worlds consistent with the story are equally good truth-maker candidates, including worlds that differ only with respect to Holmes's blood type. These worlds also agree on everything explicitly stated in or entailed by the story, including Holmes being a detective and not eight feet tall, and also on a lot of stuff not explicitly asserted by (but not incompatible with) the

stories, such as the usual laws of nature and that Edinburgh is to the north of London.

Let *W* be the set of equidistant worlds representing fiction *f*. We define *truth-in-f* by quantifying over the members of *W* as follows:

> *p* is true-in-*f* iff *p* is true in all worlds $w \in W$
>
> *p* is false-in-*f* iff *p* is true in *none* of the worlds $w \in W$.
>
> *p* is neither-true-nor-false-in-*f* iff *p* is true in some worlds $w \in W$ but not others.

'Holmes is a detective' is thus true-in-the-fiction, 'Holmes is five feet tall' is false, and 'Holmes has blood type A' is neither, unless of course it is, unknown to us at present, entailed by some property actually ascribed to Holmes in worlds where the usual facts about biology hold.

The possibilist definition of truth-in-fiction has sometimes been criticized (cf. Sainsbury 2010; Wright 2014) for helping itself to controversial resources, such as *supervaluationism* (see Stone 2010).[5] Supervaluationism would be needed only if we let every world in which some possible individual – a different one in different worlds – satisfied the descriptions associated with the Holmes stories be a truth-maker for those stories; in that case, sentences of the form 'Holmes is *F*' would be associated with

---

[5] Stone and Wright both notice similarities between the problem of selecting a referent and *the problem of the many*, which can be illustrated as follows (cf. Lewis 1993): A cloud is an aggregate of droplets. At the outskirts of the cloud the density of droplets gradually falls off, but since the boundaries are vague it is impossible to tell where the boundaries of the cloud actually are. As a consequence, many different aggregates are equally good candidates to be the cloud, and we seem to have no means of identifying the cloud with any one of these rather than another. But all of the aggregates are clouds. And then we have many clouds, even though there is, of course, only one. The suggestive parallel to the selection problem breaks down quickly, however. Even if we developed a strategy for picking out a particular aggregate – perhaps an arbitrary one – as the value of 'that cloud' it would not solve the problem that the other aggregates are, by definition, clouds and that the world ends up containing vastly many more clouds than it should. The *problem of the many* is the ontological problem that the number of clouds is too large, not a problem of how to pick out particular aggregates. Though we need, as discussed in Section 2, to select a single referent for 'Holmes' among multiple equally good candidates, the existence of multiple candidates is not itself a problem for us as long as we can select one of them, as I argued we can.

multiple singular propositions containing different referents, and supervaluationism or similar means would be needed to evaluate these sentences. On our account, although there are multiple worlds among the truth-makers for sentences involving 'Holmes' where Holmes has somewhat different properties, it is *Holmes* who has these properties in the relevant worlds, not his surrogates, counterparts or doppelgangers. Worlds where surrogates, counterparts or doppelgangers have these properties are not truth-makers for the Holmes stories. We do not need supervaluationism.

Indeed, *truth-in-fiction* is, formally, just a standard box operator from modal logic ranging over the restricted set of closest worlds where the events described in the fiction take place, and 'p is *false-in-fiction-f*' is equivalent to 'it is true-in-fiction-f that not-p.' Since 'Holmes has blood type A' is true in some but not all of these worlds it is not true-in-the-fiction, but it is *consistent with the fiction*: We could define an operator 'consistent-with-the-fiction' from 'true-in-the-fiction' the way we define the possibility operator from the necessity operator and negation.[6] The tautology 'Holmes either has blood type A or he doesn't,' however, is true-in-the-fiction since it is true in all the truth-making worlds. The consequence that all tautologies are true in any fiction has been criticized (cf. Proudfoot 2006, 11). However, given that the connectives mean the same in fiction as in ordinary speech all tautologies *should* be true in all fictions. We would have a problem if *true-in-fiction-f* distributed over the disjunction – that is, if *true-in-fiction* ($p \vee \neg p$) entailed *true-in-fiction-p* $\vee$ *true-in-fiction-*$\neg p$ – but, insofar as *true-in-fiction* is treated as a restricted box operator, it doesn't. $\Box p \vee \Box q$ is not a consequence of $\Box(p \vee q)$ in any normal modal logic.

An advantage of our account is that what is *true-in-fiction* evolves over the course of creating a story. When the author adds information to the stories the set of worlds that reflect the stories – the *context set* of the stories – shrinks, and information previously consistent-with-the-fiction becomes

[6] Alternatively, assuming *closeness*, we could treat truth-in-fiction as a version of counterfactual reasoning. Let Γ be the set of claims asserted in the fiction. p is true according to the fiction iff $\Gamma \mathrel{\Box\!\!\to} p$. Given the incompleteness of fictional characters we would then have to deny Conditional Excluded Middle, $(\Gamma \mathrel{\Box\!\!\to} p) \vee (\Gamma \mathrel{\Box\!\!\to} \neg p)$; claims about Holmes's blood type, for instance, will not have a determinate truth-value.

true-in-the-fiction. By the same token, our account has no problem explaining or predicting continuity across books involving the same universe, or other authors picking up where Doyle left off, inventing new stories involving the (numerically) same characters and settings. The recent TV series locating Holmes in modern-day London, for instance, might be thought of as counterfactual Holmes stories involving the (numerically) same characters. Similarly, Alban Berg's Wozzeck, Manfred Gurlitt's Wozzeck, and Georg Büchner's Woyzeck inhabit slightly different (sets of) worlds that are truth-makers for slightly different stories, but may still concern the (numerically) same individual located in those different worlds. (How to determine whether a character portrayed in one story is the same one portrayed in another will presumably depend either on certain psychological facts about the author, or causal links between an author's portrayal and previous introductions of characters.)

Of course, discussions may be had about what counts as indeterminate according to the stories, and the distance between the truth-making worlds for a fiction and the actual world may be different than the author intended if the author relies on false beliefs about, say, biology or geography, or if the story contains unintended plot holes or anachronisms. This is not by itself a problem for our account. We should, however, note that when thinking about fictions containing discontinuities or anachronisms, we are often inclined to engage in repair and, if possible, disregard such errors. I have no detailed story to tell about what (pragmatic) principles we rely on in such cases, but our model is able to accommodate various suggestions.

4.2. Closeness and indeterminacy

The *context set* for the stories, as discussed above, contains all worlds consistent with everything stated in a fiction, and is not limited by *closeness* – it should be entirely up to an author to suddenly and coherently decide that Greenland doesn't exist in her fiction, for instance, or that different laws of nature hold. Yet when interpreting a story we often assume that – unless contradicted by the story – ordinary facts hold, such as basic geography or laws of nature. I take it that such assumptions play an important role in making sense of and enjoying fiction – at least speculations about what characters could or should have done rely on importing unstated assumptions about the real world such as standard laws or

generalizations about psychology. The worlds that determine *truth-in-fiction* should be a subset of the context set rather than the whole, identified by the proper importation of facts about the real world, according to some parameter.

Closeness is an obvious candidate for adding the appropriate restrictions; that is, the worlds that determine *truth-in-fiction-f* are those consistent with what is stated in or entailed by *f* and are, on balance, *closest* to the actual world (or at least do not gratuitously depart from it). Assuming *closeness* would mean that speculations about details of relevant events not explicitly detailed in the stories may have determinate correctness conditions. And intuitively, we treat them as if they do; even speculations for instance about Holmes's sexual orientation or whether he had a neurodevelopmental disorder often seem to give rise to genuinely factual disputes, and insofar as what is *true-in-the-stories* is determined by *closeness*, claims about these matters may indeed (but may of course not) have a determinate truth value, depending on whether enough information is given in the stories, in combination with facts imported from the actual world, to settle the issue.

It is also worth pointing out that closeness may be incorporated in different ways. Lewis (1978) suggests relativizing *truth-in-fiction* to *the beliefs prevalent in an author's community* rather than to the actual world.[7] One possible upshot of doing so is expansion of the set of truth-making worlds. If the author's community does not have a prevalent belief about whether some fact p holds or not of the actual world, the truth-making set would have to include both p and $\neg p$ worlds, thus introducing more indeterminacy.

Now, possibilism is not committed to *closeness*. In principle, a possibilist might define truth-in-F over the whole context set for F – what's explicitly stated (and entailed) in the fiction and nothing else – or use any other means to restrict the set of truth-making worlds. One could for instance imagine restricting the context set to the worlds that are most *aesthetically pleasing*, although it may in that case of course be difficult to formulate conditions for whether a particular fact ought to be imported.

---

7 The question of whether the set of truth-making worlds should be relativized this way reflects prominent and – I take it – open debates among literary theorists. It is not particularly difficult to adjust the definition of *truth-in-fiction* to either option.

Indeed, given that i) providing truth-conditions for individual sentences does not depend on issues related to interpreting the fiction as a whole, and ii) our model of fiction as a type of discourse does not essentially rely on identifying the relevant truth-making subset, formulating the conditions for which facts to import is not obviously a greater or more pressing challenge for me than for other theories that appeal to *truth-in-fiction*. That said, I think many objections to *closeness* are unpersuasive. In 4.2.1 and 4.2.2 I sketch how we may answer some common complaints.

4.2.1. Accidental imports

It has been argued that *closeness* automatically imports facts we do not want to import (see Proudfoot 2006). One may for instance wonder whether it ought to be *true-in-Hamlet* that Sneferu built pyramids, which it probably is if *truth-in-Hamlet* is defined in terms of the closest worlds in which the events described in *Hamlet* take place. But although that consequence may be surprising, it is even in principle not potentially *problematic* – if there were good reasons to think it shouldn't be true-in-*Hamlet* that Sneferu built pyramids, then those reasons would also block the importation of that fact. The mere intuition stemming from the observation that Egyptian pharaohs are *irrelevant* to *Hamlet* is, I think, moot; it *is* irrelevant (otherwise it wouldn't get imported by default), but I see no good reason why irrelevance should imply truth-value gaps. Indeed, if pyramids ever became a question in discussions of *Hamlet* – for whatever reason – it seems legitimate to point out that *Hamlet* is supposed to take place in the actual world (that is, the worlds closest to the actual), and in the actual world Sneferu built pyramids.

Proudfoot (2006) sharpens the worry with examples of *bounded fictions*. Suppose we have a fiction *G* depicting a fictional war in Europe. Though the fiction does not concern England per se, it includes a map of part of downtown Cambridge, with the marked location of a number *n* of enemy tanks. It is accordingly *true-in-G* that there are *n* enemy tanks in the depicted part of Cambridge. Proudfoot's worry, though, is that it is *also* true-in-*G*, if truth-in-*G* is defined as the closest worlds consistent with the fiction, that there are no tanks in downtown Cambridge *outside* of the depicted area, ostensibly since worlds where as much as possible is kept the same as in the actual world except for the *n* depicted tanks are worlds where

there are no tanks outside of the bounds of the given map ('no-tanks worlds').

This result seems wrong.[8] But it is not clearly an objection to using *closeness* as a measure for truth-in-fiction rather than to using a particular *measure* for closeness. I suspect we have the intuition that it shouldn't be true-in-*G* that there are no tanks outside the depicted area because, intuitively and independently of fiction, we *don't* think that the *closest* worlds where there are *n* tanks in the depicted area are no-tanks worlds; hence, a measure of closeness that has the no-tanks result is simply the wrong measure.[9] We cannot just *keep any truth about the actual world we can* when introducing the tanks; we also need to supplement that change for instance with plausible stories of how the tanks got there. That 'there are no tanks outside the depicted area' is true of the actual world does not mean that no-tanks worlds are closer when we also make the background changes needed to get them there – in the closest worlds in which England is invaded, for instance, the invaders presumably employ more than *n* tanks. Since there seems to be many equally plausible background stories, it is indeterminate what is the case outside of the depicted area.

One upshot is that it will be difficult to provide a formal measure of closeness; a proper definition that takes into account *plausible background histories* (or consequences) would presumably require a formalization of *relevance* and probably a solution to some version of the *frame problem*. Fortunately, for the purposes for which we need *truth-in-fiction* – to interpret and fill in blanks in a reasonably systematic way – we don't need a formal criterion, but can be rather pragmatic about the matter. The less precise the measure for closeness is, the more claims will end up neither true nor false in the story, but it is not clear why that would be a *bad* thing.

---

[8] Proudfoot's primary point, though, is that possibilism cannot respect the *boundedness* of such fictions; i.e. exclude from being true-in-the-fiction any fact beyond the boundaries. She is right that possibilists cannot (easily) recognize such bounds, but as argued above this is not something that should worry them.

[9] The discussion of similarity relations is vast, and I cannot go into details (see Lewis 1979 for some considerations). Using *closeness* for truth-in-fiction does of course assume that some ordering can be achieved, but unlike other issues for which closeness matters (counterfactuals, perhaps verisimilitude) we don't need it to be very fine-grained.

Closeness still yields precision enough in the vast majority of cases to use it as a tool for interpreting fictions.

4.2.2. Competing models

The challenges to realism raised by Everett (2005) require further refinements. Consider fictions where it is unclear whether certain characters actually occur, such as Tatyana Tolstaya's *The Slynx*, in which "in the end, I think, it is pretty much left open whether or not there really is a Slynx" (Everett 2005, 630). Now, the set of possible worlds identified by Tolstaya does indeed contain both worlds where there is and worlds where there isn't a Slynx,[10] but I do not think we should conclude that it is neither *true-in-The-Slynx* nor *false-in-The-Slynx* that there is one; at least that is not the complete story. Holmes's blood type is indeterminate because it doesn't matter to the story; the existence or not of the Slynx *does* matter to Tolstaya's story and will matter significantly to how we understand other parts of it. Moreover, whereas worlds where Holmes have different blood types are equally close to the actual world, the non-Slynx worlds seem *prima facie* closer than the Slynx worlds on the (crucial) assumption that they are equally consistent with the rest of the story.

I think the correct thing to say is rather that the set of truth-making worlds are either the set of Slynx-containing worlds, or the set of those that do not contain a Slynx, but not both. Rather, the story leaves it unclear or indeterminate which of these sets contains worlds that are overall i) closer, and ii) cohere better with what is stated in the fiction. Hence, we get *competing models* of the fiction, and perhaps no clear means for deciding that one of them is the right one. Such indeterminacy is relatively common. In many cases, one may perhaps say that a text can give us multiple different fictions with very different answers to what is true-in-the-fiction. Another example is Patrick Ness's *A Monster Calls*, which may (legitimately) be read as a fantasy story and as a metaphor-loaded description of the main character's psychological processing of his life situation.

Competing models are also needed for Everett's *Frackworld* example, in which "[n]o one was absolutely sure whether Frick and Frack were really

---

[10] Thus, for the question of whether The Slynx is a fictional character the answer is 'yes'; the question is whether this character in fact appears in the story(-worlds).

the same person or not" (Everett 2005, 629). Now, insofar as Frick = Frack is either necessarily true or necessarily false, the competing interpretations require disjoint context sets. The interesting wrinkle in this case is that on the interpretation on which 'Frick' and 'Frack' refer to a single individual, the names refer to a different individual than they do on the interpretation on which they refer to different ones. Accordingly, relative to each interpretation a sentence 'Frick is *F*' will express different propositions, and the *meaning* of the sentences in the story will depend on what interpretation we assume. But this upshot is as expected if we assume that names are directly referential; an author of a story like this would indeed write sentences that are genuinely ambiguous, without having a determinate intention with regard to which interpretation is correct. Accordingly, I do not find the example to be a particularly worrisome.

4.3. Contradictions and surrealism

Sometimes storytellers make mistakes and locate the truth-making worlds for a fiction further away from the actual world than intended. But sometimes fictions also contain outright contradictions and impossibilities, true in *no* world. Inconsistent fictions and inconsistent character developments are commonly raised objections to possibilism. Now, the problem is at least somewhat less pressing on the account developed here. If the reference of a name for a fictional character were fixed by description, then ascriptions of contradictory properties would entail that no referent is picked out. However, insofar as reference is determined by stipulation, independently of the properties eventually ascribed, ascription of inconsistent properties is no obstacle to *referring* to that character. Instead, inconsistent ascriptions become a problem for defining truth-in-the-fiction and identifying an appropriate context set.

Though contradictory fictions are tricky, the problem should not be overstated. First, the problem arises from *genuine* contradictions. Watson's infamous war wound, which appears to move around over the course of the stories, does not generate a genuine contradiction – there *are* worlds where war wounds move around; they're just further away than we generally assume the Holmes worlds to be. We often respond to such cases with local repairs, but that is because we (for pragmatic reasons) choose to – the repairs are not *forced* upon us by the fiction. Certain genuine contradictions

may also be solved – as Lewis (1983) suggests – by compartmentalization, defining truth for *parts* of the fiction rather than the whole, for instance. (A further resource is appeal to competing models, as suggested above.) Intuitively, compartmentalization and local repairs reflect ordinary practice when facing continuity problems or contradictions, and having to resort to such measures is at least no objection to the possibilist account of how claims made in fiction can be meaningful and have truth-conditions.

Nor is it an obvious objection to possibilism that it has nothing to say about how to understand surreal stories or events. In the case of apparent violations of the laws of logic, for instance, we *should* be inclined to say that it becomes unclear what should be considered true-in-the-fiction and even what propositions the text is expressing. *Through the Looking Glass*, for instance, contains passages where it is less than clear that we have a firm grasp of the truth-conditional content. The unsurprising lesson to draw from such examples is that fiction serves purposes and exhibits qualities not tied to assertions (or pretense assertions) of matters of fact – its propositional, truth-conditional content – that may even *depend on* us being unable to pinpoint what claim it in fact makes. Similarly, that possible worlds can be used to define truth-conditions for claims in fiction and model important features of narrative fictions, does not mean that they are particularly useful for capturing the value of, say, poetry or humor. Our model is intended as a basic model of fictional storytelling as a type of discourse, not a template for analyzing *literary appreciation* or explaining all literary tools at an author's disposal.

One source of ostensibly problematic examples come from *embedded fictions*. Now, embedded fictions (modeled using worlds accessible from the worlds of the first-order fiction) are themselves not particularly problematic for possibilism, but sometimes fictions blur the distinction and let characters that are real according to the fiction interact with characters that are fictional according to the fiction, or enter the fiction-within-the-fiction. Insofar as 'fictional' entails 'non-actual', such scenarios are contradictory; in fact, they cannot even be meaningfully described, any more than a fiction using the phrase 'married bachelor' in a fiction describes a married bachelor. However, nothing in possibilism rules out worlds where individuals have different relationships with the scenarios they describe in their stories; the stories-within-the-fictions would not be *fictions* (for the participants in the first-order fiction) in *our* sense, and their characters not strictly

speaking *fictional characters* (insofar as they are available for causal interaction, which genuinely fictional characters, by definition, are not).[11] Fictions that describe such encounters or interactions do, by the meaning of 'fiction' for *us*, accordingly not describe interactions with fictions, but perhaps with fictions* – contents of stories that are not separated from life in the world of the fiction by being non-actual relative to that world, but by some other barrier.

That leaves us with examples of blatantly contradictory fictions such as *Sylvan's Box* (cf. Priest 2005), which purports to describe an object that is both there and isn't there at the same time, or Everett's (2005) examples where e.g. the symmetry of identity does not hold. In these examples the contradictions are explicit and central elements of the story, and compartmentalization or local repairs do not help. In such rare cases I am inclined to bite the bullet and deny that the stories are, in fact, best interpreted as completely meaningfully describing a contradictory state of affairs (see Hanley 2004).[12] *Truth-in-fiction* is a species of *truth*, and unless we are dialetheists 'it is both there and not there the same time' isn't in any sense *true*, not even true-in-fiction, even if the author claims otherwise.

5. Conclusion

I have endeavored to provide a partial defense of a possibilist approach to fiction against some common objections, and to (cursorily) elaborate on some potential advantages. Major obstacles remain. I have not discussed possibilism's potential commitment to modal realism. Nor have I developed my response to the selection problem in detail. But insofar as these questions can be answered, possibilism has some attractive features. First,

---

[11] That is: There is nothing contradictory about me interacting with Holmes, and there are worlds where I do. But it *is* a contradiction in terms to say that I might (causally) interact with a fictional character as such, and should be on any acceptable analysis of 'fiction'.

[12] Indeed, Everett's examples rather call for repair. A non-symmetric relation is by definition not an identity relation, so an author purporting to describe non-symmetric identity is using the word 'identity' wrong; what he or she describes is in fact a different relation, for instance a supervenience relation.

it provides a straightforward account of the *meaning* of claims made in fiction, which are assigned the same kind of semantic content as claims outside of fiction but, of course, supposed to be interpreted under assumptions that the interpreter ordinarily ought to be aware do not *actually* hold. Second, possibilism underpins a fruitful model for fictional discourse. Of course, further challenges remain, especially since fictional narratives are often far more complex than the toy examples discussed here. We have had nothing to say for instance about the role of the narrator (which presumably requires at least an extra level in the model), how to understand, say, fourth wall violations, or importation of assumptions related to genre conventions.

Finally, I have made no attempt to compare the possibilist analysis with its competitors. My goal has accordingly been rather modest, and given the outstanding challenges to possibilism – in particular concerning metaphysical commitments – this article is perhaps best understood as an attempt to help clarify where the real battle should take place.

Acknowledgments

The research behind this article was made possible in part through a grant from the SASPRO/FP7-Marie Curie Actions-COFUND scheme [grant number 0086/01/03/-b] as well as VEGA grant 2/0049/16: Fictionalism in philosophy and science (primary investigator Marián Zouhar). I am grateful to participants at the Philosophy of Language (1): Semantics of Fictional Discourse conference in October, 2016, as well as anonymous referees, for comments on earlier versions of this paper.

References

BRECKENRIDGE, W. & MAGIDOR, O. (2012): Arbitrary Reference. *Philosophical Studies* 158, No. 3, 377-400.

BROCK, S. (2002): Fictionalism about Fictional Characters. *Noûs* 36, 1-21.

EVERETT, A. (2005). Against Fictional Realism. *The Journal of Philosophy* 102, No. 12, 624-649.

HANLEY, R. (2004): As Good As It Gets: Lewis on Truth in Fiction. *Australasian Journal of Philosophy* 82, 112-128.

HOWELL, R. (1979): Fictional Objects: How They Are and How They Aren't. *Poetics* 8, 129-177.

KÖLBEL, M. (2002): *Truth without Objectivity*. London & New York: Routledge.

KRIPKE, S. (2011): Vacuous Names and Fictional Entities. In: *Philosophical Troubles: Collected Papers Volume 1*. Oxford: Oxford University Press, 52-75.

LEWIS, D. (1978): Truth in Fiction. *American Philosophical Quarterly* 15, 37-46.

LEWIS, D. (1979): Counterfactual Dependence and Time's Arrow. *Noûs* 13, 455-476.

LEWIS, D. (1983): Postscripts to Truth in Fiction. In: *Philosophical Papers*. Oxford: Oxford University Press, 276-280.

LEWIS, D. (1993): Many, But Almost One. In: Bacon, J., Campbell, K. & Reinhardt, L. (eds.): *Ontology, Causality and Mind*. Cambridge: Cambridge University Press, 23-42.

MACFARLANE, J. (2005): Making Sense of Relative Truth. *Proceedings of the Aristotelian Society* 105, 321-339.

PRIEST, G. (2005): *Towards Non-Being. The Logic and Metaphysics of Intentionality*. Oxford: Clarendon Press.

PROUDFOOT, D. (2006): Possible Worlds Semantics and Fiction. *Journal of Philosophical Logic* 35, 9-40.

SAINSBURY, M. (2010): *Fiction and Fictionalism*. London: Routledge.

STALNAKER, R. (1978): Assertion. In: Cole, P. (ed.): *Syntax and Semantics, Vol 9*. New York: Academic Press, 315-332.

STONE, J. (2010): Harry Potter and the Spectre of Imprecision. *Analysis* 70, 4, 638-644.

SULLIVAN, A. (2005): Rigid Designation, Direct Reference, and Modal Metaphysics. *Pacific Philosophical Quarterly* 86, 577-599.

THOMASSON, A. (1999): *Fiction and Metaphysics*. Cambridge: Cambridge University Press.

VAN INWAGEN, P. (1977): Creatures of Fiction. *American Philosophical Quarterly* 4, 299-308.

WRIGHT, B. (2014): Many, But Almost Holmes? In: García-Carpintero, M. & Martí, G. (eds.): *Empty Representations: Reference and Non-Existence*. Oxford: Oxford University Press, 290-306.