sciendo

Interpretive Context, Counterpart Theory and Fictional Realism without Contradictions

Raphael Morris
The Australian National University

Abstract
Models for truth in fiction must be able to account for differing versions and interpretations of a given fiction in such a way that prevents contradictions from arising. I propose an analysis of truth in fiction designed to accommodate this. I examine both the interpretation of claims about truth in fiction (the 'Interpretation Problem') and the metaphysical nature of fictional worlds and entities (the 'Metaphysical Problem'). My reply to the Interpretation Problem is a semantic contextualism influenced by Cameron (2012), while my reply to the Metaphysical Problem involves an extension and generalisation of the counterpart-theoretic analysis put forth by Lewis (1978). The proposed analysis considers interpretive context as a counterpart relation corresponding to a set of worlds, W, and states that a sentence φ is true in interpretive context W iff φ is true at every world ($w \in W$). I consider the implications of this analysis for singular terms in fiction, concluding that their extensions are the members of sets of counterparts. In the case of pre-existing singular terms in fiction, familiar properties of the corresponding actual-world entities are salient in restricting the counterpart relation. I also explore interpretations of sentences concerning multiple fictions and those concerning both fictional and actual entities. This account tolerates a plurality of interpretive approaches, avoiding contradictions.

Keywords
Fiction, singular terms, counterpart theory, fictional realism, semantic contextualism

1 Introduction

We can truly say that Chewbacca predeceased Han Solo: Han

watched as Chewbacca sacrificed himself on the planet of Sernpidal to save Han's youngest son, Anakin Solo (Salvatore 1999). We can also truly say that Han Solo predeceased Chewbacca: Chewie saw Han stabbed through the heart by his son Kylo Ren, formerly Ben Solo, on Starkiller Base (Abrams dir. 2015).

Taken together, and given reasonable assumptions about linear time and the permanence of death, these two statements appear to entail a contradiction. Apparently, Chewbacca was alive after Han died—the same Chewbacca who predeceased Han. Yet we cannot truly say that Chewbacca came back from the dead.

What will we relinquish to save these fictions from absurdity? One response is that of the fictional anti-realist, who denies the truth of the premises (Everett 2005). The anti-realist retorts that we cannot truly say that Chewbacca predeceased Han Solo, nor vice versa, since neither Han nor Chewie ever actually existed. It will suffice to note that this position is at odds with how we normally talk about fiction. We may readily admit that Luke Skywalker, Harry Potter, and Frodo Baggins do not actually exist, but we do not take ourselves to be merely pretending that Luke is a Jedi, Harry is a wizard, or Frodo is a hobbit. These are properties of the characters, not reflections of the reader's suspension of disbelief—indeed, one might conclude that a certain tale is a fiction specifically because the tale ascribes to entities properties that the reader believes are possessed by no actual entity.

I propose an alternative account to anti-realism to explain the lack of contradiction. My account involves the claim that the Chewbacca who predeceased Han Solo is *not* the same Wookiee as the Chewbacca who watched Han die. Such a proposal has two parts, corresponding to the two questions that it endeavours to answer. The first of these is the question of how we are to interpret sentences about fiction, such as 'Chewbacca predeceased Han Solo'; I call this the Interpretation Problem. The second is the question of the metaphysical nature of these entities, and what the metaphysical relationship between the different versions of Chewbacca is; this is the Metaphysical Problem.

My project in this paper is to address both these problems. I propose a unified account that explains both how sentences about fiction express propositions, and what the referents of the terms in those sentences are, in such a way that both the above claims about

Chewbacca are true without entailing either contradiction or anti-realism. I will also discuss the implications of this account for the interpretation of singular terms in fiction. Finally, I will make some suggestions for interpreting claims relating entities in different fictions and relating fictional and actual entities.

2 The interpretation problem

How can 'Chewbacca predeceased Han Solo' be both true and false without there being a contradiction? Is the proposition that is affirmed when the sentence is true the same as the proposition that is negated when the sentence is false? What is the mechanism by which these sorts of sentences express propositions? These are the questions that I group under the banner of the Interpretation Problem.

Andrew McGonigal (2013) proposes a solution: the sentence does indeed express a single proposition, with a single referent. However, McGonigal believes that the truth value of this proposition depends on other matters—for instance, which works of fiction one takes into consideration when evaluating the proposition. Thus, if the fiction that includes the events on Sernpidal is salient, then 'Chewbacca predeceased Han Solo' is true. The truth values of propositions—but not propositions themselves—are relative to the context of assessment in which the propositions are evaluated. There is only a contradiction if one fails to appropriately limit the relevant circumstances of evaluation.

Like McGonigal, Caplan's (2014) 'work contextualism' holds the propositions expressed by sentences about works of fiction do not change depending on context. However, on Caplan's view, the propositions expressed by the works of fiction themselves do change depending on context, which is what causes the changes in truth value of the propositions expressed by sentences about fiction.

The motivation behind both McGonigal's relativism and Caplan's work contextualism is to avoid contradictions between works of fiction and their sequels, and to make sense of the way in those who have engaged with the sequels can correctly judge utterances that seemed true before the release of sequels to be false now. In this pursuit they are successful. The sentence 'Sherlock Holmes fell to his death at the Reichenbach Falls' can be taken to express a single proposition.

The same proposition is found in the Sherlock Homes stories as of 1893, when 'The Final Problem' (Doyle 1893) was published, which makes the sentence true in that context. In 1903, when the publication of 'The Adventure of the Empty House' (Doyle 1903) changed the canon of the stories so that the proposition's negation was found in the stories, the very same sentence is false. Both Caplan's and McGonigal's accounts can successfully explain the discrepancy here.

The problem with these accounts arises when we are not dealing with sequels but with alternative continuities side by side. On April 25 2014, Disney announced that almost all the *Star Wars* novels, comics, and various spin-offs produced before this date would be moved into a separate continuity (dubbed *Legends*) from the works produced after that date (Lucasfilm Ltd. 2014). This was done to create room for new stories, so that future works (including *The Force Awakens*) would not have to be concerned with consistency with the myriad of earlier works. However, while stories such as *Vector Prime* (in which Chewbacca dies and Han Solo survives) are now no longer official canon, they still have a secured place within the *Legends* continuity.

The retconning (a portmanteau of 'retroactive continuity') of *Vector Prime* changed the context of the novel, but this change in work context did not directly change the truth value of the claim 'Chewbacca predeceased Han Solo'. Rather, it changed the way that conversational context determines the propositional content of the claim, differentiating the *Legends* version of Chewbacca from the canonical version of Chewbacca. Once again, it appears that the context in which a sentence about fiction is uttered affects the propositional content of the sentence. Caplan's account can only change the meaning of the sentence by changing the context of the work, but here the context of conversation—whether one is talking about the current canon or the *Legends* stories—is what determines the truth value of 'Chewbacca predeceased Han Solo', not the context of the work itself. Similarly, McGonigal's relativist account holds that if I have just watched *The Force Awakens* and seen Chewbacca outlive Han Solo, and my friend has just read *Vector Prime*, I should take my friend's utterance of 'Chewbacca predeceased Han Solo' to be false—it is my context of assessment that matters, not the context of my friend's utterance, according to McGonigal. Work contextualism and semantic relativism are inadequate for explaining this, so we

must turn to another theory.

Ross P. Cameron (2012) has developed a semantic contextualist model for the interpretation of fiction that includes this feature. Cameron's solution, which I favour, holds that there is not one Chewbacca, interpreted in different contexts, but multiple Chewbaccas. One of these, Chewbacca-Living, saw his partner Han Solo die on Starkiller Base; the other, Chewbacca-Dead, perished on Sernpidal. The apparent contradiction contained by 'Chewbacca predeceased Han Solo' being both true and false is resolved in that the sentence expresses different propositions in different contexts—and one of these propositions is true, while the other is false. Using <S> to denote the propositional content of a sentence 'S', <Chewbacca-Dead predeceased Han Solo> is true, and this is the proposition expressed by the claim 'Chewbacca predeceased Han Solo' in the context of the *Legends* continuity. <Chewbacca-Living predeceased Han Solo>, on the other hand, is false, and this is the proposition expressed by 'Chewbacca predeceased Han Solo' in the context of a discussion of the current official *Star Wars* canon.

Cameron's solution is able to account for a wide variety of character-variants. There are the variants who differ based on context that is changed by the publication of a sequel, such as Holmes-1893, who perished at the Reichenbach Falls, and Holmes-1903, who survived. There are the variants who are separated by different continuities, such as Chewbacca-Living and Chewbacca-Dead. Beyond these, there are variants separated by other facets of interpretive context. Discussions of Albus Dumbledore that admit statements by the author as true are discussions of one Dumbledore, while discussions of Dumbledore from an interpretive perspective that upholds death of the author are about a different Dumbledore; it is true to say that the former Dumbledore is gay (Irwin 2015), but not true to make the same claim about the latter Dumbledore.

Finally, there are facts about the world that are not specified in the fiction, but may be assumed by the interpretive context. David Lewis (1978) asks whether we can truly say that Sherlock Holmes lives closer to Paddington station than to Waterloo, based on reasoning from our knowledge of Holmes' address according to the fiction and knowledge of real-world London geography. Following Cameron, I respond that the answer depends on whether the geography of

real-world London is part of our interpretive context.

This is my answer to the Interpretation Problem: The subjects of sentences about fiction depend on the context in which those sentences are interpreted. This context has two parts. The first is the corpus that is taken to comprise the fiction: which continuity is used, whether sequels are included, the admissibility of authorial declarations, and so on. The second part of context is the set of background assumptions that are deemed admissible in the interpretation, which may include or exclude the laws of physics, all or part of the geography of real-world London, or the claim that the story's narrator is reliable. Taken together with the sentence in question, both elements of context determine the specific proposition expressed, which can then be evaluated as true or false objectively.

3 The metaphysical problem

The answer to the Interpretation Problem says that the term 'Chewbacca' has different referents in different contexts. The Metaphysical Problem concerns the metaphysical nature of these referents: real or unreal, concrete or abstract, located in the actual world or in some other world. It also covers the question of what, if anything, is the metaphysical relationship between different referents.

Cameron's response to the Metaphysical Problem is an unusual combination of nominalism and fictional realism. On Cameron's (2012) view, fictional entities exist, and such entities are abstract. However, Cameron does not include such entities as ontological commitments. His ability to do this relies on a meta-ontology that includes the following two claims (Cameron 2012: 181):

(1) "An existential claim can be made true by something other than what it says exists."

(2) "The ontological commitments of a claim are those entities that must be invoked as *truth makers* for the claim if it is to be held true."

On Cameron's account, the truth makers for claims about the existence of fictional entities are concrete events, namely the acts of

interpretation involved in reading a fiction. The abstract fictional entities exist, but they are not ontological commitments, because what makes it true that they exist is the concrete act of interpretation.

If we accept both of Cameron's meta-ontological claims, then we should laud his account of fiction as both successful and impressively ontologically minimalist. However, there is good reason for rejecting the second claim, a blow which would prove fatal for Cameron's metaphysics of fiction.

Cameron's view that the existence of fictional entities is entailed by our interpretive practices in reading and discussing fiction derives from Amie L. Thomasson's (1999, 2003) account of fiction. Indeed, Cameron owes a great deal to Thomasson and her 'easy ontology'. Thomasson (2014: 129) argues that that "trivial inferences from uncontroversial premises" suffice to resolve a great many ontological debates. She proposes that we reconceive ontology as being primarily about conceptual analysis, beyond which non-philosophical empirical work is sufficient to settle most existence questions—including those about fictional entities.

To a large extent, Cameron is on board with Thomasson's project. He agrees that many—perhaps most—existence questions can be settled by the one-two punch of philosophical conceptual analysis and subsequent non-philosophical empirical investigation. But Cameron (2012: 184–5) sees a hierarchy of existence conditions. Perhaps what it takes for there to be tables are molecules arranged table-wise, and what it takes for there to be molecules are atoms arranged molecule-wise, and so on. According to Cameron, this process must come to an end. At the end of the line, he argues, are the truth makers for the existence claims concerning everything in the chain. And we need only incorporate these ground-level truth makers into our ontology in order to get the existence of everything else for free.

The problem is with the hierarchy. Cameron needs it to justify his assertion that the only ontological commitments of an existence claim are its necessary truth makers. But as Thomasson (ms.) points out, language is *ontologically flexible*. Instead of the truth of 'tables exist' depending on the truth of 'there are molecules arranged table-wise', we could just as easily say that it depends on the truth of a whole host of alternative phrasings, each with different ontological commitments. Perhaps the truth maker is the property, *being a table*,

instantiated by the object. Perhaps it is the state of affairs of there being tables, obtaining. In fact, it seems quite plausible that tables are involved in making true 'there are molecules arranged table-wise'; certainly, we seem to need a conception of tables in order to grasp how molecules might be arranged table-wise (Thomasson ms.)

Cameron presumes that any given claim has some entities that *must* be invoked as truth makers. But the apparent ontological flexibility of language seems to demonstrate that many, if not all, claims have no such determinate entities—and therefore, no ontological commitments at all. In seeking to minimise ontology, Cameron's approach accidentally eliminates it entirely—a rather unpalatable conclusion for those of us who think our ontologies aren't completely empty.

In response to this criticism, Cameron (ms.) argues that the ontological flexibility of language is no obstacle to constructing a hierarchy of truthmaking, it simply shows that looking at what entails what is no help to figuring out the hierarchy. This may well be the case, but if it is, then it severely curtails the usefulness of truthmaking until we have some way of finding out what this hierarchy looks like, which Cameron admits is "very, very hard." In particular, I see no convincing reason (except perhaps a suspiciously convenient parsimony) to claim that Sherlock Holmes is *not* a truth maker for 'Sherlock Holmes exists.'

Of course, one can still adopt Cameron's theory without his meta-ontology and accept an ontology that includes abstract objects. I don't doubt that his response to the Interpretation Problem can be integrated with a more standard artifactualist metaphysics, of the kind given by Thomasson (1999, 2003). However, for anyone who wishes to eschew abstract fictional objects, I propose the following alternative:

David Lewis (1978) offers a solution that avoids the trouble of integrating abstract objects into our ontology, while providing a clear relationship between the many variants of any given fictional character. Lewis maintains that fictional characters are not singular entities but rather correspond to sets of counterparts. On Lewis' view, 'Chewbacca' and the names of other fictional characters, objects and events map to sets of concrete, fully-determined entities who exist in other possible worlds, but not the actual world. A property is only true of the object or character if it is true for all the various entities under our consideration. These entities, while not identical, are related to

each other by a counterpart relation. Depending on the chosen counterpart relation, two entities may or may not be counterparts.

I propose that it is our interpretive context, as described in the previous section, that determines the scope of the counterpart relation. For example, if our interpretive context is that of the current *Star Wars* canon, then 'Chewbacca' means 'Chewbacca-Living', and only Chewbacca-entities who saw Han Solo die on Starkiller Base are included in this reference. These different entities may differ in a variety of ways, from the length of their fur to major life experiences, but all of them have all the properties attributed to Chewbacca within the current *Star Wars* canon, which makes them Chewbacca-counterparts per this specific counterpart relation. In contrast, within an interpretive context of the *Legends* continuity, 'Chewbacca' refers to only those entities with all the properties Chewbacca has in the *Legends* continuity, including dying in an act of heroic sacrifice on Sernpidal.

Lewis' metaphysical perspective also explains the sense in which Chewbacca-Living and Chewbacca-Dead are both Chewbacca. If one uses a broader counterpart relation that includes all Chewbacca-entities consistent with the depiction of Chewbacca in the first six *Star Wars* films, then all of the Chewbacca-entities included in the previous two counterpart relations are valid Chewbacca-candidates. And one can go even further—Lewis' metaphysics can also explain how non-canonical versions of a character, such as those existing in fan fiction or early drafts of screenplays, are still that character. This is another advantage of a counterpart-theoretic metaphysics—with a sufficiently inclusive counterpart relation, there are Chewbacca-candidates who are not even Wookiees. In the fan fiction *Luke and the Beanstalk* (starwenn 2016), for instance, Chewbacca is a human.

4 Articulation of the proposed account

Combining the responses to the Interpretation Problem and the Metaphysical Problem, we may articulate a comprehensive account of the circumstances in which sentences about fiction are true. I propose the following analysis: *A sentence φ is true in interpretive context* W *iff φ is true at every world $w \in$ W.*

Under this analysis, the interpretive context corresponds to a set of worlds, W, and a counterpart relation between entities at different

worlds in W. Are we interested in all the worlds with a character called 'Chewbacca', or all the worlds consistent with the first six *Star Wars* films, or all the worlds consistent with a specific continuity? Are we only interested in worlds that fit certain other parameters—where the second law of thermodynamics applies, or where Albus Dumbledore experienced sexual attraction towards Gellert Grindelwald?

Furthermore—though this is often less controversial—in the worlds under our consideration, which entities are counterparts of which others? Consider the differences between Ares and Mars—the former often feared, the latter widely venerated. Without changing the set W, we might debate whether or not two entities are so essentially different as to preclude their being metaphysical counterparts, and use different counterpart relations for different purposes.

Using this analysis to frame discussions about fiction, it is apparent that much of the debate about fictional events, entities, and objects amounts to disagreement about the domain of W, and perhaps the choice of counterpart relation. Debate about the legitimacy of reasoning about fiction from authorial intent amounts to debate about whether the author's declarations must be true at every world $w \in W$. Debates about the legitimacy of using psychology to make inferences about characters' motivations, or physics to make inferences about the collateral damage from a fictional battle, or medicine to evaluate the severity of characters' injuries are reducible to disagreement over whether the principles of these sciences must hold true at all worlds $w \in W$. This does not diminish the value of such debate; it merely helps elucidate the questions faced by those who study literature.

Finally, note that there is no mention of fiction specifically within my analysis. This is because whether a sentence concerns fictional worlds forms part of the interpretive context. 'Sherlock Holmes is a detective' is true when W includes the worlds of Conan Doyle's Sherlock Holmes stories, but false when W only includes the actual world. In contrast, 'The Sherlock Holmes stories were written by Sir Arthur Conan Doyle' is true when W only includes the actual world, but not when it includes the worlds of the fiction. One of the responsibilities of a competent interpreter is to distinguish claims about the actual world from claims within a fiction.

5 Some implications for the analysis of singular terms in fiction

5.1 Singular terms in counterpart theory

Before we can begin to explore the implications of this analysis for our conception of singular terms in fiction (pre-existing or otherwise), we must clarify what it is that we mean by a 'singular term'. The naïve conception of a singular term as picking out a single entity in a single world will not do; the above analysis makes it clear that when we talk about a given fictional entity, we are really talking about entities in a multitude of worlds. Unless we want to forego all talk of singular terms in fiction, our notion of a 'singular term' must include a mechanism for picking out the intended entities within whichever worlds they exist. What follows is a blueprint and defence of counterpart theory as this mechanism.

Many of the same considerations that led us to adopt a counterpart-theoretic ontology in the analysis will here lead us to interpret singular terms in fiction as shorthand for sentences that quantify over a set of counterparts, as in the analysis. The terms 'Chewbacca', 'Minas Tirith', and 'Excalibur' do not refer to singular entities, but are used in sentences that quantify over a great many different entities. What unites these disparate entities is that all are counterparts of any 'typical' Chewbacca-entity, Minas Tirith-entity and Excalibur-entity, respectively. (The restriction 'typical' ensures that in more permissive interpretive contexts, the more outlandish counterparts are not used to circumscribe the domain of quantification. If we are discussing *Star Wars* fan fiction, all our Chewbaccas should be counterparts of one from the films, not one from an arbitrary fan work.) The precise domain of quantification thus depends on interpretive context (choice of counterpart relation and set of worlds), as outlined in the analysis.

In the same manner, this analysis captures the meanings of pre-existing singular terms as they appear within fictions. Philip Roth's (2004) novel *The Plot Against America* depicts Charles Lindbergh winning the presidential election of 1940. 'Charles Lindbergh' here refers not to the historical Lindbergh, but is used in a sentence that quantifies over his many counterparts as they appear in the worlds of

the fiction. Similarly, 'Richard Nixon' in the graphic novel *Watchmen* (Moore and Gibbons 1987) refers not to the actual Nixon, but is used to help quantify over his counterparts in the fiction.

However, singular terms in most discussions of history and geography *do* refer to singular entities (at least as they are generally understood). 'Charles Lindbergh', 'Richard Nixon' and 'New York' are generally held to each have a unique referent: the corresponding entity in the actual world. In part out of a desire to preserve this uniqueness of reference, various philosophers have formulated criteria of transworld identity. Of particular note is Saul Kripke's (1980) conception of the singular term as a 'rigid designator', which refers to the same entity across many worlds (rather than an entity and its counterparts).

Kripke and the others of his persuasion would have us forego all talk of counterparts within our analysis, replacing it with a concept of identity across worlds. According to this view, Chewbacca-Living and Chewbacca-Dead are the same entity, and not merely counterparts. Similarly, Charles Lindbergh in the actual world is identical with President Lindbergh in Roth's book. Singular terms are no longer shorthand to guide quantification, but once again are genuinely singular.

Though such a revision of the analysis might seem appealing, I believe that we have good reason for rejecting it and insisting that we are indeed speaking of counterparts. One reason becomes apparent when we consider the case in which a historical or otherwise pre-existing entity appears within a work of fiction, without being named as such. The late Sir Terry Pratchett was a master of this technique, with his Discworld books filled with homages and parodies of various historical, mythic, and literary entities. Examples include Leonard of Quirm (Pratchett 1997): the bald, bearded genius and artist who drew up unprecedented designs for flying machines and submarines and was horrified at the idea that his inventions might be used for war; and Djelibeybi (Pratchett 1989): the country sustained by a great crocodile-infested river, renowned for its pyramids and where cats are considered a sacred animal.

These entities match definite descriptions of Leonardo da Vinci and Ancient Egypt, respectively. But as Kripke (1980) correctly observes, definite descriptions are non-rigid designators. The fact that Leonardo da Vinci and Leonard of Quirm share a definite description fails to entail any sort of transworld identity of one with

the other. The proposed revision of the analysis is unable to grasp a metaphysical relation between the two beyond a similarity of properties. The similarity is quaint, but metaphysically irrelevant; it does not explain how Leonard of Quirm is a *version* of Leonardo da Vinci.

In contrast, our counterpart-theoretical analysis of fiction succeeds at capturing this relation. Leonardo da Vinci and Leonard of Quirm are different entities (technically, as a fictional character, the latter captures many distinct entities). Nonetheless, the two are counterparts, and they are connected by a counterpart relation that is delineated not by any name or singular term, but rather by a certain definite description. Counterpart theory enables the analysis to account for tropes such as parody, homage, and allusion, whose subtlety escapes a strict criterion of transworld identity.

Another reason occurs when we consider fan fictions like *Luke and the Beanstalk*. The version of Darth Vader in this story is both a version of Darth Vader from *Star Wars* and a version of the giant from the fairy tale Jack and the Beanstalk. Because transworld identity is an identity relation, it is transitive and symmetric. Therefore, if Darth Vader from the fan fiction is identical to Darth Vader from *Star Wars*, then the giant in Jack and the Beanstalk is also identical to Darth Vader from *Star Wars*, which seems implausible. The counterpart relation carries no such requirements of transitivity or symmetry, and can thereby avoid these problems.

Are singular terms truly singular? According to our analysis, the singular term 'Charles Lindbergh' refers to all Lindbergh-counterparts within the scope of a contextually-defined counterpart relation. Why call such a term 'singular' at all, given the plurality of its referents? The answer lies in the context of the counterpart relation. Discussions about fiction lend themselves to broad counterpart relations with domains spanning many worlds. But in everyday conversation, we are frequently operating under an exceptionally narrow counterpart relation: one which considers only an entity's counterparts within the actual world. When I say: 'Charles Lindbergh made the first solo transatlantic flight', by default I am taken to be speaking about this world's Charles Lindbergh, and no one else. The illusion of singular terms as possessing unique referents stems from the ubiquity of the actual-world counterpart relation.

It appears that the reference of fictional names can contextually

shift somewhat more easily than the reference of ordinary names, especially within a sentence. Consider the following true (if misleading) sentence: 'In 1540, Cromwell was beheaded on the King's orders, but in 1649, the King was beheaded on Cromwell's orders.' Here the first Cromwell is Thomas Cromwell (beheaded on the orders of Henry VIII) and the second Cromwell is Oliver Cromwell (who ordered the beheading of Charles I). This reference shift makes 'Cromwell was beheaded' true in contexts where 16th Century history is salient, and false in contexts where 17th Century history is salient. Furthermore, if Ann says 'Cromwell was beheaded' in the first context and Mary says 'Cromwell was beheaded' in the second, I can't say 'Both Ann and Mary claim that Cromwell was beheaded'; there is no one Cromwell they are both talking about!

We observe a similar reference shift in the true sentence 'In *The Force Awakens*, Chewbacca outlived Han Solo, but in *Legends*, Chewbacca dies before Han Solo.' Again, if Ann claims 'Chewbacca outlived Han Solo' in one context, and Mary makes the same utterance in the other, there is no one Chewbacca about which they are both talking.

The difference between the two cases is that anaphora and ellipsis seem far more permissible in the fictional case than the ordinary one. It seems I can truly say 'In *The Force Awakens,* Chewbacca outlived Han Solo, but in *Legends*, he didn't'. But I probably can't truly say 'In 1549, Cromwell was beheaded, but in 1640, he ordered the King's beheading'. At present I do not know how to explain this difference.

5.2 Interpreting pre-existing singular terms in fiction

Charles Lindbergh's counterparts in the worlds of *The Plot Against America* do not necessarily share all the properties of the historical Lindbergh. They even possess some properties, like the property of being president in 1941, that the historical Lindbergh explicitly lacks. Other those specifically mentioned in the novel, does Lindbergh necessarily share any properties with his fictional counterparts?

Once again, the answer depends on our counterpart relation. One could conceivably read Roth's novel through a fantastically lenient interpretive lens, allowing for worlds with Lindbergh-counterparts who differ entirely from the historical Lindbergh in all ways except for stipulated similarities. In some of these worlds, Lindbergh

is secretly an android, werewolf, or extra-terrestrial. Such a reader might even be unaware that the historical Lindbergh existed at all.

This reading, while legitimate in principle, seems to miss the point of using pre-existing singular terms in fiction. Writers pick their subjects deliberately. Historical characters, locations, and established figures from literature and mythology are selected for the specific properties each possess. Disregarding this context entirely is an error.

Let us return for a moment to the original example involving Chewbacca. Suppose my friend and I have just watched *The Force Awakens*, in which Han Solo dies and Chewbacca survives. Upon leaving the cinema, I turn to my friend and remark 'Chewbacca predeceased Han Solo.' Strictly speaking, my statement is not false. Relative to the worlds of the *Legends* continuity, it is true that Chewbacca died first. But to judge my claim as true is to ignore a critically salient aspect of the statement's context: that I have just walked out of a film in which Han Solo predeceases Chewbacca.

Similarly, the imaginary reader who, after reading *The Plot Against America* says 'Charles Lindbergh might be an android' has ignored a critically salient aspect of their context—that they have just read a novel based on a historical human being. By ignoring this context, they have failed to restrict the domain of their counterpart relation appropriately, thus leaving themselves guilty of misinterpretation.

Clearly some restriction of the domain is necessary. The exact parameters, however, remain ambiguous. If a fiction is set in present-day London, it is ridiculous to imagine that Elizabeth II (or her counterpart) is not queen—unless of course this has been stipulated within the fiction. But at the other extreme it is unreasonable to assume that each of the guards at Buckingham Palace is a counterpart of an actual member of the Queen's Guard. What counts as an appropriate degree of realism remains a point of contention among both critics and readers.

I decline to take a position on that debate within this paper. Any conditions on the similarity between the referents of pre-existing singular terms among the worlds of a fiction and their actual-world counterparts can be framed in terms of restrictions on the domain of W and the specific counterpart relation. The strength of the proposed account lies in its ability to accommodate varying domains for

W, thereby avoiding contradiction between different interpretations of a fiction.

6 Sentences in transfictive and actual-world contexts

6.1 Truths in fiction and truths about fiction

The analysis presented so far is designed to account for truths *in* fiction—that which we can we truly say to be the case according to a given fiction. However, in talking about fictional entities and events, we also make claims *about* fiction. These relate both actual individuals and fictional entities, as well as fictional entities from disparate fictions. Here are just a few examples from the literature (Lewis 1978, Kroon 1994, Kroon and Volotini 2016):

(A) Smith admires Holmes.

(B) Holmes has acquired a cultish following [in the actual world].

(C) Holmes is cleverer than any actual detective.

(D) Holmes is much cleverer than Poirot.

(E) Smith admires [Jim] Garrison [in Oliver Stone's fiction *JFK*].

(F) Napoleon [in *War and Peace*] was more pompous than Caesar [in *Julius Caesar*].

(G) Holmes would not have needed tapes to get the goods on Nixon.

While a full analysis of these kinds of claims is beyond the scope of this paper, it is immediately apparent that the present account is inadequate. Such claims do not concern goings-on within any one world, so even by quantifying over, the analysis will be ill-suited.

But this is no cause for alarm. The analysis for truth in fiction proposed in this paper is merely a special case of the counterpart semantics. The complete counterpart-theoretic toolbox is well-equipped

to handle sentences that occur in such contexts. Certainly, sentences like 'Smith admires Holmes' or 'Holmes is much cleverer than Poirot' cannot be said to be true at a world—or even at all worlds in a set of worlds. But the very cornerstone of counterpart theory is its ability to relate entities in distinct worlds—for the counterpart relation does exactly that! There is no reason for the counterpart theorist why other relations cannot also take their arguments from distinct worlds.

In declining to explain these sorts of sentences, Lewis underestimates the power of the counterpart semantics. In fact, Kroon (1994) observes that the counterpart theorist avoids many of the problems that plague anti-realist and Meinongian accounts, particularly when it comes to terms like 'London' that may refer to either actual or fictional entities. And yet, the capacity of counterpart theory to explain why the above sentences are true apparently remains in dispute. To assuage this scepticism, I will now formally explicate (A) to (G) in the counterpart-theoretic semantics.

The renderings here are designed to complement the account of fiction presented in this paper. They do not presume some default set of fictional worlds, determined by closeness to the actual world or belief worlds or some such, as is the case for Lewis and Kroon. Rather, they leave membership of sets of fictional worlds to be determined by interpretive context. This is important—many people's admiration for Luke Skywalker (for example) is contingent on whether the interpretive context includes the newer films.

6.2 Formal explication of (A) to (G)

The following examples use some notation from Lewis (1968):

Ixy   (*x* is in possible world *y*)
Ax   (*x* is actual)
Cxy   (*x* is a counterpart of *y*).

(A)   Smith admires Holmes.

$\forall x \, \forall y \, ((Ixy \wedge y \in W_{SH} \wedge Cxa) \rightarrow Rsx)$

Here, a is an arbitrary typical instance[1] of Holmes, s is Smith, and Rxz means 'x admires z'. W_{SH} is the set of Sherlock Holmes-worlds to be considered under a particular interpretive context. As the context changes, the truth of the sentence may also change—Smith may admire the counterparts of Holmes consistent with Doyle's stories, but detest those consistent with his depiction in the contemporary BBC adaptation.

(B)   Holmes has acquired a cultish following [in the actual world].

$$\forall x \, \forall y \, ((Ixy \land y{\in}W_{SH} \land Cxa) \to \exists z \, (Fz \land Az \land Rxz))$$

Here, a is an arbitrary typical instance of Holmes, Fz means 'z is cultish', and Rxz means 'x has acquired z as a following'. As before, W_{SH} is the set of Sherlock Holmes-worlds in a particular context, and the contextual dependence of its domain is again crucial. While it may be the case that all the counterparts of Holmes consistent with Doyle's stories have earned their followings, the same is unlikely to be true for those Holmes-entities found in the worlds of more obscure adaptations.

The quantifier order is also noteworthy. Different Holmes-counterparts, even within W_{SH}, need not have the *same* following. This becomes relevant when fans are split as to some interpretive question. One might follow the counterparts of Sherlock Holmes who used cocaine often but resisted addiction through sheer mental power, but not the counterparts whose addictions, while serious, did not detract from their deductive powers.

(C)   Holmes is cleverer than any actual detective.

$$\forall x \, \forall y \, \forall z \, (((Ixy \land y{\in}W_{SH} \land Cxa) \land (Dz \land Az)) \to Rxz)$$

Here, a and W_{SH} mean the same as in the previous examples, Dz means 'z is a detective', and Rxz means 'x is cleverer than z'. Note that

---

[1] I suggest that for x to be an 'arbitrary typical instance' of Holmes is to satisfy the following conditions: 1) in a domain consisting only of the world x occupies, x is unambiguously the referent of the name 'Holmes' in this context, and 2) in the relevant context, x is a member of the largest set of entities that both satisfy condition 1) and all have exactly the same set of counterparts. Thanks to an anonymous referee for suggesting that I make this point explicit.

the examples given can all handle what is fictional and what is not without resorting to any prefixed 'according to the fiction' operator.

(D) Holmes is much cleverer than Poirot.

$\forall x \, \forall y \, \forall z \, \forall w \, (((Ixy \land y \in W_{SH} \land Cxa) \land (Izw \land w \in W_{HP} \land Czb)) \to Rxz)$

Here, a and W_{SH} mean the same as in the previous examples, b is an arbitrary typical instance of Poirot, W_{HP} is the set of Hercule Poirot-worlds in a particular context, and Rxz means 'x is much cleverer than z'. The principles of domain-restriction behind this explication demonstrate how counterpart theory can solve what Kroon (1994) calls the "real-fictional problem". Sentences like 'London is prettier than London [in the Holmes stories]' just mean that the actual London is prettier than its counterparts in the worlds of Sherlock Holmes.

(E) Smith admires [Jim] Garrison [in Oliver Stone's fiction *JFK*].

$\forall x \, \forall y \, ((Ixy \land y \in W_{JFK} \land Cxg) \to Rsx)$

Here, g is the actual Jim Garrison, s is Smith, W_{SH} is the set of JFK-worlds in a given context, and Rxz means 'x admires z'. Again, the real-fictional problem presents no difficulty; Smith may well be the actual Garrison and despise himself without loss of coherency.

(F) Napoleon [in *War and Peace*] was more pompous than Caesar [in *Julius Caesar*].

$\forall x \, \forall y \, \forall z \, \forall w \, (((Ixy \land y \in W_{W\&P} \land Cxn) \land (Izw \land w \in W_{JC} \land Czc)) \to Rxz)$

Here, n is the actual Napoleon, c is the actual Caesar, $W_{W\&P}$ is the set of *War and Peace*-worlds in a given context, W_{JC} is the set of *Julius Caesar*-worlds in that context, and Rxz means 'x was more pompous than z'. This is clearly different from comparisons involving the actual Napoleon or the actual Caesar, who reside in the actual world and not the worlds corresponding to the fictions.

(G) Holmes would not have needed tapes to get the goods on Nixon.

$\forall x \, ((Cxa \land \neg Tx) \, \Diamond\!\!\to Rxz)$

Here, a is an arbitrary typical instance of Holmes, n is Richard Nixon, Tx means 'x has tapes', and Rxy means 'x gets the goods on y'. In this explication, I use the equivalence of the English sentence expressed by (G) to the sentence 'Holmes, without tapes, might have got the goods on Nixon'. The operator $\Diamond\!\!\rightarrow$ behaves as defined in Lewis (1973): '$x \Diamond\!\!\rightarrow y$' means 'Were x the case, it might be the case that y.'

Interestingly, in this example, the important contextual variation is not in the domain of a set of worlds, but in the choice of counterpart relation. In other situations, a bumbling and incompetent parody of Holmes might be a counterpart of Holmes. This could not be the case for (G) to be true. Someone who asserts (G) would doubtless be thinking that a much more stringent relation is required for someone to qualify as a counterpart of Holmes.

These examples are far from exhaustive, but I believe they go some way towards demonstrating the benefits of counterpart theory for semantics of fiction. A complete counterpart semantics for fiction would be able to explain also what it is to be a fictional entity—my suspicion is that it would be something like not being in the actual world and being in a world partially described in an actual work of fiction. It would also feature an account of 'counterfictional' sentences, including those comparing entities in fictional worlds with those in counterfictional worlds. One such sentence, from Lewis (1978), is 'Holmes could have solved the ABC murders sooner than Poirot'.

The task of constructing a comprehensive counterpart semantics for fiction is beyond the scope of this paper. My hope is that the translations of (A) to (G) should provide a degree of encouragement for anyone who might attempt such a task.

Conclusion

My project here has been to develop a new counterpart-theoretic account of truth in fiction, to examine the treatment of singular terms within this account, and to sketch a proposal for a counterpart-theoretic treatment of truth about fiction along similar lines. The core account consists simply of the claim: *A sentence φ is true in interpretive context* W *iff φ is true at every world $w \in$ W.*

Much of the philosophical discussion about truth in fiction amounts to attempts to establish restrictions on the domain of W. This is how Lewis (1978) proceeds, and it is especially pertinent in relation to the treatment of pre-existing singular terms like 'Charles Lindbergh' or 'London', which are taken to refer to entities possessing an assortment of known properties.

However, I believe that this is ultimately a non-philosophical concern. I leave arguments over the particular domain of the appropriate interpretive context to the literary critics. There is no need for philosophers to set out alternative analyses for popular interpretive frameworks, as Lewis (1978) does. The philosophical issue is one of the nature of interpretive context in general. My proposal is that interpretive context corresponds to a counterpart relation, spanning various worlds containing counterparts of entities. Furthermore, I contend that the domain of this relation depends not on the context of the work, but on the context of the interpretive claim itself.

Is this fictional realism? That depends on whether one is a realist about these other worlds. I am inclined towards such a realist position, and while a proper discussion of modal realism lies well beyond the scope of this paper, I will offer a tentative argument in its favour.

The Science of Discworld (Pratchett et al. 1999) describes wizards from the fictional Discworld looking with curiosity at Roundworld (i.e. the actual world). The wizards (who in fact have accidentally created Roundworld) do not view it as a mere unreal possibility, but acknowledge its reality. Were they to dismiss our world as an unreal fiction, the reader would doubtless view their mistake with some smugness. If a fictional entity would be in error to claim that our world is unreal, I have little faith in the claim that their worlds are any less real than our own.[2]

Raphael Morris
The Australian National University

Raphael Morris

raphael.morris@anu.edu.au

References

Abrams, Jeffrey Jacob (dir). 2015. *The Force Awakens*, motion picture. Los Angeles: Walt Disney Studios Motion Pictures.

Cameron, Ross. 2008. Truthmakers and ontological commitment: or how to deal with complex objects and mathematical ontology without getting into trouble. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 140(1): 1–18.

Cameron, Ross. 2012. How to be a nominalist and a fictional realist. In *Art and Abstract Objects*, ed. by Christy Mag Uidhir. Oxford: Oxford University Press.

Cameron, Ross. Manuscript. Truthmaking and metametaphysics. In *The Routledge Handbook of Metametaphysics*, ed. by Ricki Bliss and James Miller. Routlege.

Caplan, Ben. 2014. Serial fiction, continued. *British Journal of Aesthetics* 54(1): 65–76.

Doyle, Arthur Conan. 1893. The Final Problem. *The Strand* 3.2.

Doyle, Arthur Conan. 1903. The Adventure of the Empty House. *The Strand* 13.10.

Everett, Anthony. 2005. Against fictional realism. *The Journal of Philosophy* 102(12): 624–49.

Irwin, William. 2015. Authorial declaration and extreme actual intentionalism: is Dumbledore gay? *Journal of Aesthetics and Art Criticism* 73(2): 141–47.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Kroon, Frederick. 1994. Make-believe and fictional reference. *The Journal of Aesthetics and Art Criticism* 52(2): 207–14.

Kroon, Frederick; and Voltolini, Alberto. 2016. Fiction. In *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. by Edward N Zalta, URL = https://plato.stanford.edu/archives/win2016/entries/fiction/.

Lewis, David. 1968. Counterpart theory and quantified modal logic. *The Journal of Philosophy* 65(5): 113–26.

Lewis, David. 1973. *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, David. 1978. Truth in fiction. *American Philosophical Quarterly* 15(1): 37–46.

Lucasfilm Ltd. 2014, *The Legendary Star Wars Expanded Universe Turns a New Page*, viewed 5 January 2017, http://www.starwars.com/news/the-legendary-star-wars-expanded-universe-turns-a-new-page.

McGonigal, Andrew. 2013. Truth, relativism, and serial fiction. *British Journal of Aesthetics* 53(2): 165–79.

Moore, Alan; and Gibbons, Dave. 1987. *Watchmen*. New York: DC Comics Inc.

Pratchett, Terry. 1989. *Pyramids*. London: Victor Gollancz.

Pratchett, Terry. 1997. *Jingo.* London: Victor Gollancz.

Pratchett, Terry; Stewart, Ian; and Cohen, Jack. 1999. *The Science of Discworld*. London: Random House.

Roth, Philip. 2004. *The Plot Against America*. London: Jonathan Cape.

Salvatore, Robert Anthony. 1999. *The New Jedi Order: Vector Prime*. New York: Del Rey.

Starwenn. 2016. Luke and the Beanstalk. *Archive of Our Own*, viewed 5 January 2017, http://archiveofourown.org/works/8899120.

Thomasson, Amie. 1999. *Fiction and Metaphysics*. Cambridge: Cambridge University Press.

Thomasson, Amie. 2003. Fictional characters and literary practices. *British Journal of Aesthetics* 43(2): 138–57.

Thomasson, Amie. 2014. *Ontology Made Easy*. Oxford: Oxford University Press.

Thomasson, Amie. Manuscript. *Truthmakers and the Project of Ontology*.