

Joseph Dartez

Writing Sample for Graduate Admissions

December 7, 2016

A Defense of the Abstract Artifact Theory of Fictional Characters

When we assert a sentence as true, part of the condition for the truth of that sentence is the existence of whatever objects we take the sentence to refer to. By Quine's criterion for ontological commitment, when I assert that "Obama is the President of the United States," I am committed to the existence of both Obama specifically and Presidents of the United States generally.¹ How, then, do we account for our discourse about fictional characters, given the truth of sentences such as "Sherlock Holmes is a fictional character"?

Abstract artifact theory (AAT) is one of the most prominent recent answers to this question. On this account, our discourse about fictional characters *does* incur commitment, specifically to abstract artifacts. An abstract artifact is an entity that (a) does not exist within the space-time continuum and (b) is created by human beings, much like spatio-temporal artifacts such as tools, buildings, books, paintings, and so on. The specific motivation for accepting abstract artifact theory over any other theory is that it more accurately matches our common sense intuitions about fictional entities. Within our non-philosophical discourse about fictional characters, we speak about the relationship between authors and fictional characters as one of creation. Thus, common discourse tends pre-analytically toward ontological commitment to fictional characters as a kind of thing that human beings create. If a theory rejects this relationship between author and character it bears the burden of providing a convincing error-theory to explain why we are mistaken in our thinking.

¹ See Quine (1953), 12.

Recently, two devastating objections to the plausibility of AAT have put it on the defensive. Each of these objections focuses on a distinction central to AAT. The first is an objection to what has come to be known as the *encoding/exemplifying distinction* in property ascription, which allows AAT proponents to distinguish between properties ascribed to fictional character and properties ascribed to human beings. The second is an objection to the *internal/external distinction* in reference, which allows AAT proponents to explain how Holmes is a fictional character but not a human being. In this paper, I accept these objections and re-envision AAT to resolve the standard problems in our discourse about fictional characters without relying on either of these two distinctions. I do so by proposing a new reference distinction that applies not to fictional characters, but to real human beings. Then I defend this new version of the theory against potential objections.

1. Naive Abstract Artifact Theory

The first difficulty that any theory of discourse about fictional characters must face is the apparent contradictions implicit in the different modes of such discourse. This difficulty emerges when we consider the basic categories of sentences that occur in fictional discourse. We can separate them into three basic categories. Paraphrasing from Amie Thomasson's "Speaking about Fictional Characters,"² we have:

1. Discourse internal to works of fiction (sentences that are true about the characters within the story, including sentences in the original text and sentence spoken as if they are in the original text)
2. Discourse external to works of fiction (sentences that are true about relationships between character and author, comparisons between characters from different stories, comparisons between characters and persons, etc.)

² See Thomasson (2003) for a detailed account of the relationships between these three discourses.

3. Nonexistence claims (e.g. “Sherlock Holmes does not exist.”)

I will follow Thomasson in calling these three categories *internal discourse*, *external discourse*, and *nonexistence claims*. All three forms of discourse are *prima facie* at odds with each other, so the task of any theory of fiction is to show how these three forms of discourse are not mutually inconsistent. A brief outline of the conflict: In the internal discourse, Sir Arthur Conan Doyle attributes *being a detective* and *playing the violin* to Sherlock Holmes, as do readers in their discourse with each other about the events in the story. In the external discourse, however, we attribute *being a fictional character* and *being created by Conan Doyle* to Sherlock Holmes. We would not attribute the property *being a detective* to anyone other than a person, and a created fictional character is surely not a person, so we must risk asserting that Holmes both is and is not a fictional character. Additionally, in nonexistence claims, we assert the emptiness of the name “Sherlock Holmes;” which would seem to prevent Holmes from being either a detective or a fictional character.

The AAT proponent navigates these difficulties by establishing the following ambiguity in fictional names. In the internal discourse, we only *pretend* to refer to a real person when we use Holmes’ name, though our pretense does not actually involve reference and so does not incur any ontological commitment. In the external discourse, we *successfully* refer to an abstract artifact when we use Holmes’ name. Finally, in nonexistence claims, we are elliptically asserting that Holmes does not exist *as a person*, but that he does exist *as a fictional character*. I call this the *internal/external reference ambiguity*. It is not merely a satellite element of AAT: the ambiguity features prominently in the theory’s account of the reference of fictional names.

The account runs as follows:

Whenever we speak of Holmes or any other fictional character as a real person, we are engaged in non-committal pretense, or *make-believe*, as Kendall Walton has it.³ Because we do not mean to

³ See Walton (1990).

assert the existence of an entity in this discourse, we do not incur any ontological commitment to an entity.⁴

The nature of the resulting fictional characters has been construed in many different ways: Peter Van Inwagen sees them as theoretical entities in the same class as novels and plots,⁵ Thomasson sees them as social constructs in the same class as marriages and contracts,⁶ and Voltolini sees them as sets of properties contingently associated with a story telling process.⁷ Regardless, what they all have in common is that a fictional character is somehow composed of the properties that have been ascribed to that character in the internal discourse, and that these properties are of a different kind from those we ascribe to them in the external discourse. The act of property ascription that occurs during make-believe literally creates or builds the fictional character. Once has been done, we may engage in external discourse by, e.g., referring to Holmes as a fictional character, relating Holmes to detectives in other stories in which Holmes does not exist, and even comparing Holmes to detectives who really do exist.

Because the AAT proponent's creative mechanism generates an abstract entity through pretend assertion, such a theorist must distinguish between the way we ascribe properties to fictional characters and the way we ascribe properties to human beings. When we assert that "Obama is the President," we are saying that a concrete particular has the property of being President of the United States. This mode of association is sometimes called "exemplification." Conversely, when we *pretend* to assert that "Sherlock Holmes is a detective," we are, according to the theory, not at all saying that a concrete particular has the property of being a detective. Instead, we are adding the property *being a detective* to

⁴ The idea that names in the internal discourse do not refer is common ground for AAT proponents. See, e.g., Searle (1975), Van Inwagen (1977), Schiffer (1996), Thomasson (2003), Voltolini (2006).

⁵ See Van Inwagen (1977), 302-3.

⁶ See Thomasson (1999), 12-3.

⁷ See Voltolini (2006), 65-6.

the list of constituent properties of the abstract artifact Holmes without asserting that the artifact either has or exemplifies the property. This mode of ascription is sometimes called “encoding.”

A fictional character can have properties in two different ways: it can encode properties and it can exemplify properties. The Sherlock Holmes that exists is not the sort of thing that can be a detective: only a person can be a detective. When we speak in the external context about the properties Holmes has in the story, we are only speaking of *encoded* properties. As an abstract artifact, Holmes’s *exemplified* properties are, e.g., *being a fictional character*, *being created by Conan Doyle*. Additionally, since a fictional character is the sort of thing that has properties encoded into it, we can say that Holmes exemplifies such properties as *having the encoded property of being a detective* and *having the encoded property of playing the violin*.

2. Objections to Naive Abstract Artifact Theory

In response to AAT, two major objections have recently been raised which undercut the plausibility of the theory.

2.1 The property ascription objection

The first objection is a problem with the encoding/exemplifying distinction. This objection was raised first by Stacie Friend.⁸ The AAT proponent’s distinction between statements that are literally true, such as “Holmes is a fictional character” and statements that are not, such as “Holmes is a violinist” leaves this theory vulnerable to counterexamples. Friend asks, what do we do with statements like, “Holmes is more brilliant than Poirot?” We cannot compare Holmes and Poirot through internal discourse because they do not exist within each other’s fictional world. But if we compare them

⁸ See Friend (2007) 151-2.

through external discourse, then neither of them can be brilliant since they only encode the property of being brilliant.

An AAT proponent might, in this instance, argue that we can compare them through internal discourse by merging their fictional worlds into a new pretense where both exist. This method, however, will not work on a sentence like, “Holmes is cleverer than any real detective.” Here, we are comparing Holmes to real detectives. If we attempt to handle this sentence through internal discourse, we’ll have to say that we’re not talking about real detectives when we say “real detectives;” rather, we’re talking about pretend correlates of some kind. If, on the other hand, we attempt to handle this sentence through external discourse, we can’t compare Holmes to real detectives at all, because Holmes exemplifies the property of *having the encoded property of being a clever detective*, whereas real detectives can exemplify the property of *being a clever detective*. The encoding/exemplifying distinction, then, must be accompanied by an error theory that either explains how we are not talking about real detectives when we say, “real detectives,” or explains how we think we can compare Holmes and real detectives when, in fact, we cannot.

We might want to circumvent this problem by creating a property *kind* distinction, such as that people have normal properties and fictional characters have fictional properties.⁹ Unfortunately, this distinction fares no better. On this account, “Sherlock Holmes is cleverer than any real detective” compares the fictional cleverness of Holmes to the non-fictional cleverness of a real detective. If we can smoothly compare the two properties just as we would compare the properties of two real detectives, then this property distinction is ad hoc: it only exists to keep our fictional characters fictional. If, on the other hand, we allow fictional properties to be functionally distinct from regular properties, we would have no reason to think they can be evaluated on the same scale, even though they share the name “cleverness”.

⁹ This move is common in Platonistic accounts of fictional entities, which distinguish “nuclear” properties (those ascribed in the text) from “non-nuclear” properties (those true of the character “outside the fiction”), see Zalta (1988).

2.2 The reference ambiguity objection

The second objection is a problem with the internal/external reference ambiguity.¹⁰ R. M. Sainsbury mounts this objection by introducing a new problematic sentence,

Kilgore Trout is a writer, and one of Vonnegut's most engaging creations.

Here we see one instance of a fictional character, Kilgore Trout, apparently used in both internal and external discourses simultaneously. This presents a serious problem for internal/external reference ambiguity. If the name does not refer when we use it in internal discourse, then in this instance it does not refer. But if the name does refer when we use it in external discourse, then in this instance it does refer. The best account that AAT can give for a sentence like this is the following paraphrase:

Kilgore Trout exemplifies having the encoded property of being a writer and Kilgore Trout exemplifies the property of being one of Vonnegut's most engaging creations.

While this paraphrase avoids the contradiction by placing the whole sentence in the external discourse, it comes at great intuitive cost. After all, given the structure of the sentence, we'd like to say that in both clauses we are attributing properties to Trout in the *same way*.

The intuitively unappealing nature of this paraphrase reveals a deeper problem with the reference ambiguity. If pressed about what a "fictional character" is, the ordinary person would want to say something like "a person who does not really exist," rather than "an abstract artifact that we pretend is a person." This common sense approach is in conflict with internal/external reference ambiguity because the ordinary person prefers to use the name of a fictional character the same way regardless of the mode of discourse. In disallowing us to refer to the same entity in both contexts, the AAT proponent must develop yet another error theory accounting for our belief that the name has the same referential mechanism regardless of the mode of discourse.

¹⁰ See Sainsbury (2010) 98-9.

Thus AAT, as presented above, is intuitively unappealing enough to require two separate error theories. Recall, though, that the primary motivation for AAT was that the idea of authors creating fictional characters is intuitively appealing. As it stands, then, this theory looks implausible.

3. Augmented Abstract Artifact Theory

3.1 Three necessary constraints

There are a few options open to the AAT proponent. One option is to develop error theories to accommodate the plausibility objections. Another option is to attempt to rescue reference ambiguity by drawing new distinctions which might somehow avoid these plausibility objections. Both of these options strike me as ad hoc. Therefore, the response I will develop in this paper is to *accept* the objections and to rework AAT *without* relying on either the encoding/exemplifying distinction or the internal/external reference ambiguity.

If reference to fictional entities shares the same basic mechanics as reference to human beings, then we can preserve the intuition that fictional characters can be meaningfully compared to human beings. On this view, when we engage in internal discourse, we ascribe properties to characters in the same way as when we engage in external discourse, and this mode of ascription is also the same as when we discuss living human beings. The absence of an encoding/exemplifying distinction entails that the internal/external distinction, whatever its other merits, cannot result in any difference in ontological commitment: pretend assertions must succeed at referring in just the same way as serious assertions.

Moreover, since we are allowing fictional characters and human beings to take the same kinds of properties, there must be a relevant sense in which they are the same kind of thing. And if fictional characters and human beings are somehow the same kind of thing, then there is only one reference mechanism: the mechanism used to refer both to human beings and to fictional characters.

We can now identify the necessary constraints for an augmented abstract artifact theory

(AAAT):

1. Reference through pretend assertion is ontologically committal.
2. Human beings and fictional characters are somehow the same kind of thing.
3. Reference to human beings and to fictional characters employs the same basic mechanism.¹¹

3.2 The central tenet of AAAT

The idea that human beings and abstract artifacts are the same kind of thing is initially preposterous. It seems manifestly false that human beings are abstract entities (*pace* Berkeley). Additionally, it seems impossible that fictional characters are concrete particulars.¹²

The position I propose for AAAT is that the human beings to which we refer when we use a name are a composite of *both* concrete particulars *and* abstract artifacts, while the fictional characters to which we refer are *only* abstract artifacts. In taking this position, I do not hold that a human being really is the composite of a concrete particular and an abstract artifact (as an Aristotelian might); rather, I only hold that it is this composite entity to which we refer and that we are thereby ontologically committed to the two members of this composite.¹³

Additionally, I will follow Amie Thomasson in supposing that fictional characters are social constructs, and that social constructs are unproblematic abstract artifacts. For our purposes, there are only a few conditions that need to be met for a character to be a social construct : (a) it is an abstract entity, (b) it is created by human beings, and (c) it is somehow constituted by the kinds of properties we

11 Nathan Salmon (1998) has proposed a similar approach in which sentences in both internal and external discourses refer to the same entity, but his account still depends on a clear distinction between the internal and external contexts, and he does not address problems in comparing fictional characters to human beings.

12 David Lewis (1978) defends this claim, but his theory of fictional characters as concrete but non-actual entities cannot account for impossible characters, such as a square circle, about whom a make-believe story can be told; neither can it account for contradictory stories. For extended criticism of Lewis' account, see Thomasson (1999) and Sainsbury (2010).

13 I mean this composite entity to be understood in the same way we might understand a dog-walking man to be a composite of a man and the dog he is walking. I make no claim about whether or not we are committed to the existence of a composite as a thing over and above just the man and the dog he is walking.

associate with human beings. That is, a social construct can model a person by being composed of the properties we typically take that person to have.

On this model, reference to a particular fictional character, whether through internal discourse or external, is reference to one and the same social construct. Moreover, the act of referring to a particular human being is dual rather than singular: We refer both to the individual person and to a social construct (a “character”), which we believe accurately models the individual. A notable corollary of this position is that pretend discourse is not unique in being ontologically creative: both serious and pretend discourses generate social constructs.

3.3 How AAAT explains the three discourses

The first task of this theory is to explain the apparent contradictions latent in the three forms of fictional discourse (internal, external, and nonexistence claims).

Consider the following three sentences:

1. Sherlock Holmes is a detective.
2. Sherlock Holmes is a fictional character.
3. Sherlock Holmes does not exist.

If all three of these sentences refer to the same social construct, we have at least one ambiguity on our hands: sentence 3 must be paraphrased to allow this social construct to exist at all. According to the model above, this paraphrase is simple:

4. Sherlock Holmes does not exist as a concrete particular, but Holmes does exist as a social construct.

Because, on the standard account of ontological commitment, “Sherlock Holmes” needs to refer in order for the nonexistence claim to be true, every account of fictional discourse must find some suitable paraphrase of nonexistence claims. What matters philosophically for handling a nonexistence

claim, then, is that the paraphrase is plausible. Assuming that reference to human beings is ontologically dual as I propose it is, this paraphrase looks very plausible.

On this account, the apparent contradiction between sentences 1 and 2 is also fairly easy to manage. If reference is dual, then every name refers to a social construct, or a “character.” That such a character is a detective then becomes common sense. Whenever we say that a human being is a detective, we are not ascribing a property to the concrete particular; rather, we are ascribing a property to the social construct with which we associate the concrete particular. On this theory, the standard objection that only a physical human being can be a detective falls flat: on the contrary, only a social construct can be a detective, because *being a detective* entails playing a particular role in a narrative, which is what a character does. Because a character is just the sort of thing that can be a detective, the modifier “fictional” only distinguishes a character that *is not* associated with a concrete particular human being from one that *is* associated with such a human being.

4. Objections to Augmented Abstract Artifact Theory

In the remainder of this paper, I will respond to objections that might be raised against AAAT.

4.1 That AAAT falsifies an analytic truth of reference

One might object that it is a definitional truth of reference that the name “Obama” refers to Obama and not to a social construct, and since this is an implausible change to the theory of reference, AAAT itself is implausible.

The prevailing theory of reference is Saul Kripke’s direct reference model.¹⁴ On this account, reference does not involve definite descriptions or concepts or any other intervening abstract element; rather, it is purely ostensive: it is analogous to pointing. When I say “Obama,” I mean to point to the

¹⁴ See Kripke (1972).

concrete particular, and if I cannot do this, I will point to someone or something that can. In this way, reference to Obama is fixed by a social chain in which I point to someone or something that points to something else and the chain always ends with the person you intended when you spoke the name. The benefit of this model is that it allows me to successfully refer to e.g. Henry Kissinger, even if the only description I can give of him is *a former Secretary of State*, which does not uniquely identify Kissinger.

Given this model of reference, it appears that social constructs are left completely out of the loop. After all, one cannot point to a social construct.

To answer this objection, we will need to clearly distinguish reference to a social construct from reference to a concrete particular, though they confusingly share a name. When I speak the name “Obama,” I ostensibly single out the concrete particular, but I also invoke the character who plays the POTUS role in the prevailing narrative of the real world.

An Obama social construct is often what people are describing when they speak the name “Obama”. In the narrative world we think of as the real world, Obama the construct has, among other things, the properties definitive of his role as POTUS. By contrast, Obama the concrete particular is only closely approximated by this description. We might assert, for example, the sentence, “Obama wants US citizens to have access to affordable healthcare without threatening the health insurance industry.” This sentence is true if and only if Obama really does want such a thing. Naturally, we do not know for sure whether the ACA is something he wanted or something he settled for, and we can imagine disputes arising concerning Obama’s actual intentions. That is, the sentence is not obviously true of the concrete particular. However, it appears to be true of the social construct *by definition*. On AAAT, this sentence expresses a particular property that truly belongs to at least one Obama social construct, and insofar as it meets this truth condition, it is true regardless of what the concrete particular wants. Thus, we might say, “Your version of Obama wants that, but I disagree: I think Obama wants to bankrupt the insurance market.” In this case, it is true within the fictional world where Obama is the

kind of person you imagine him to be; that is, it is part of your Obama construct that the construct corresponds to a man, but this is only true in a fictional world. However, if we want to judge whether the sentence is true in the actual world, we must compare this particular social construct to the concrete particular. Thus, we often speak falsely of the concrete particular but truly of the social construct.¹⁵

The reference mechanism has two parallel tracks. One reference track is the chain of direct ostension that deals exclusively with concrete particulars pointing to each other. The other reference track deals with the make-believe game we play involving the various roles we attribute to ourselves and other people. On this second track, referring to a social construct entails either stating the properties of the construct or ostensibly singling out some expression of these properties (like a text). When we speculate about Obama's true sentiments, we are engaging in a form of make-believe. Since, on this theory, make-believe is an ontologically creative act, our speculation generates social constructs to which we refer. It is a *feature* of AAAT that make-believe about real people is not privileged over other forms of make-believe. The theory handles both forms of pretense exactly the same way. Thus, all the properties we say are true of Obama (including being a person) actually belong to Obama social constructs. The name, "Obama," then refers to both the concrete particular and to the socially constructed role the concrete particular apparently plays, and only rarely do we think about these two entities as distinct from each other.

4.2 That AAAT substitutes one unacceptable reference ambiguity for another

At this point, a skeptical person might object that I have introduced a reference ambiguity for *real people* instead of fictional characters. But there is no reason to think that our reference to real

¹⁵ Tarski's truth condition can still function in this reference framework given that names in the metalanguage only refer to concrete particulars and not social constructs as well. Second-order knowledge, however, is still a problem as it is for any correspondence theory of truth.

people is ambiguous; when we say “Obama,” we never mean to talk about anything besides just the concrete particular.

It is true that I am introducing a new reference ambiguity, but whereas the internal/external reference ambiguity is not supported by the data, the concrete particular/social construct reference ambiguity *is* supported by the data. I might say, for example, “No matter how much I learn about Obama, I do not really know him unless I spend time with him.” Suppose I do meet the man and I discover some shocking truth about him, such as that he intends to become emperor of the world. I might tell everyone I know “We are wrong about Obama,” but I might alternatively say, “Obama is not who we think he is.” While the direct ostension model can accommodate the first claim, it must paraphrase the second claim. However, if there is a distinction between the concrete particular and the social construct, then it is uncontroversially true to say “Obama is not who we think he is,” without the need for a paraphrase: “Obama” in this case signifies only the concrete particular, and “who” signifies a specific social construct which, in this case, does not correspond to the concrete particular.

Every name for a real human being has at least two distinct uses, and this is intuitively obvious when we consider the significance of reputation. In fact, each name has a multitude of uses, each specific to the particular version of the person we are discussing. Large chunks of the American population speak about and respond to an Obama social construct who is a foreigner bent on destroying the country (among other unpalatable traits). It is not a stretch to say that we take ourselves to be speaking only approximately about the concrete particular when we make judgments about his character or when a news source reports on Obama’s approval rating. In these instances, we are speaking about a caricature of the concrete particular that fits into a specific shared narrative. If we are wrong in our beliefs about Obama, it’s not that we falsely attributed properties to the concrete particular; it’s that the social construct we name when we say “Obama” does not model the concrete particular as accurately as we think.

4.3 That AAAT cannot fix reference for a social construct

Another objection: there is no way to determine which social construct we have in mind when we say “Obama,” so we cannot fix our reference on the same social construct.

This is actually a feature of the theory and not a problem. We can easily fix the reference for the social construct for Sherlock Holmes because everything we know about Holmes is contained in the canon, which is the books Conan Doyle wrote about him. Thus there exists what we might call a *canonical construct* that can uniquely identify the character Holmes. The sources of knowledge for a real human being, however, are far more diverse and the relationships between these sources far more complex than that of a character like Holmes. The canonical construct for Obama can only be the construct that most accurately models the concrete particular. But who has access to the properties of this construct? His family is the best available candidate, but even they surely lack access to some of the concrete particular’s traits. We should by no means expect agreement about the canonical construct associated with the name “Obama,” but we *can* all agree that the criterion for judging whether the cluster of properties constitutive of the social construct corresponds to the concrete particular is whether the picture of Obama it creates is similar to the concrete particular we perceive.

The canonical construct is the particular version of the social construct that experts (whoever they might be) agree is correct, whether the construct models a concrete particular or is merely a character. The privilege that a canonical construct enjoys is not that it gets to exist and the other social constructs do not. On the contrary, there exist as many different social constructs even of Sherlock Holmes as there exist versions of him in specific make-believe games. Similarly, there are as many social constructs of President Obama as there are conceptions (understood as clusters of properties) of him discussed among human beings. Rather, the privilege the canonical Obama social construct enjoys is that, in most cases, this is the construct we intend to refer to and link to the concrete particular,

though we often fail in that act of reference and succeed instead in referring to some other Obama social construct. This failure to refer to the construct we intended does not, however, impact our success in referring to the concrete particular through direct ostention. In other words, very few of us are experts on Obama, but we can still successfully refer to the concrete particular without being experts.

4.4 That AAT cannot explain correspondence

One might also object that a social construct is incapable of modeling a human being, so there is no meaningful sense in which we can say that there is correspondence between canonical construct Obama and concrete particular Obama.

Consider a different social construct: money. The American Dollar construct corresponds to specific concrete tokens because only these tokens meet the criteria for being money. Whatever else this social construct might be, it surely includes the rules that govern the what counts as money in the currency game of make-believe. So we pretend of paper, coin and digital tokens that they are money and can be used as such in our game of make-believe if and only if they meet the criteria for being money as established in our game of make-believe.

We can implement this same relationship for human beings, but in reverse. The canonical Obama construct corresponds to the person because the rules for being that person match the person's actual function in the social order. Thus, rather than treating the social construct as a cluster of properties, we can alternatively treat it as a cluster of criteria for being Obama (and often these criteria will take the form "having property X"). In the case of both money and Obama, such criteria are designed to ensure that counterfeit does not circulate. To say that the Obama construct corresponds to Obama, then, means that the social construct is an articulation of the criteria for being Obama. The concrete particular should be the only token that fits the social construct's role.

At this point, we can imagine a feedback loop between a token concrete particular and the social constructs we refer to along with that concrete particular. Just as the criteria for being money vary depending on counterfeiting practice and technology, so the criteria for being Obama can change depending on how much information is publicly available about concrete particular Obama. If a social construct does not accurately describe Obama's role, then the construct becomes a kind of counterfeit. Again, here we might say, "He's not who you think he is."

4.5 That AAAT cannot account for trans-world discourse

A final objection: If in speaking about either fictional characters or human beings we always refer to social constructs, then we should not be able to speak about fictional characters as anything other than human beings.

This objection can be met by establishing different contexts for reference. Consider Sainsbury's counterexample sentence:

Kilgore Trout is a writer, and one of Vonnegut's most engaging creations.

Whatever Kilgore Trout is, he is a writer within the relevant fictional world and one of Vonnegut's most engaging creations outside the fictional world. We'd want to establish a similar relationship between Trout and his own creations. That is, within the fictional world Vonnegut created, which we can call *Trout's world*, Trout is a writer and whatever he writes is his own creation. If Trout wrote about a writer named Teddy Sturgeon who also happens to be an engaging character, then we might say that "Teddy Sturgeon is a writer, and one of Trout's most engaging creations." Now we have precisely the same relationships between all the different terms and the only thing that has changed is context. And we'd want to say that within the fictional world where Trout exists (as a concrete

particular) the canonical Trout construct corresponds to concrete particular Trout, but the canonical Teddy Sturgeon construct does not correspond to any concrete particular.¹⁶

The above similarity suggests that there exist nested narrative worlds and that the actual world numbers among those nested worlds. Similarly, we can continue to nest fictional worlds within fictional worlds, beginning with the actual world (where Vonnegut is a writer) and ending however many layers deep we might like to go. So if we want to match our intuitions about what we mean when we use all of these names, we'd say that we use them all in precisely the same ways, including the properties that entail concreteness. The only difference in our usage becomes the indexing world. In the actual world, Vonnegut is a writer and Trout is a character. In Trout's world, Trout is a writer, Vonnegut does not exist, and Sturgeon is a character.¹⁷ In Sturgeon's world, Sturgeon is a writer and Trout and Vonnegut do not exist. This analysis shows the flexibility of nonexistence claims, and how they change depending on narrative context. An appropriate paraphrase would look like this:

Kilgore Trout is a writer in his world, and one of Vonnegut's most engaging creations in ours.

In order to make this work, we must grant that these social constructs, fictional or not, have sub-clusters of properties (or correspondence criteria) that are indexed to the various worlds in which the entity exists in some way. Kilgore Trout, then, has the property of being a writer in the sub-cluster indexed to the world(s) into which Vonnegut wrote him. Trout also has the property of being one of Vonnegut's most engaging creation in the sub-cluster indexed to *our* world. One can imagine a long list of different world-indexed property subsets for characters such as Spiderman. Whether we are talking

¹⁶ One of the properties found in a social construct is "being a concrete particular," which borders on introducing a paradox, though this is the same intuition Wittgenstein meant to capture when he said, "the world is the totality of facts, not things." On this model, when we describe a concrete particular, we build a social construct and compare it to the particular. Only on the actual world, though, will we find concrete particulars that correspond to social constructs. While Trout is a concrete particular in Trout's world, we cannot compare social construct Trout to concrete particular Trout in Trout's world because we do not live in that world.

¹⁷ In fact, Vonnegut wrote himself into some of the Kilgore Trout stories as a kind of demigod. Since writers typically do not do this, the simple nonexistence claim I make here, though literally false in this case, is more generalizable to other cases.

about a different construct or the same construct in a different world should be sensitive to whether we perceive ourselves to be talking about the same character or a merely a similar character.

5. Conclusion

The contention of this paper is that abstract artifact theory can be augmented in order to avoid the two major objections to the theory's plausibility.

The proposed augmentation of the theory may seem too radical a possibility to entertain. In fact, on my account, the theory is no longer merely a theory about fictional characters, but a theory about how reference works in general, fictional discourse being merely a special case of the more general rule that discourse about anything whatsoever is typically discourse about social constructs along with an ostensive singling out of the objects to which we take those social constructs to correspond. While this theory is surely subject to many more objections I have not yet dreamed up, the potential benefits of the theory are dramatic, since it transforms the empty name problem of ontological commitment into a pseudo-problem.

Moreover, the possibility that all reference to real objects involves simultaneous reference to social constructs allows us to more clearly describe some surprising situations in which apparently true property ascriptions concerning human beings don't seem to match the data, such as in cases of race, gender, and even holding a belief.¹⁸ If our property ascriptions in fact involve social constructs and we only compare these constructs to concrete particular human beings, then we do not need to face the contradictions entailed by truly saying that a person belongs, for example, to a particular race, despite the apparently constructed nature of race. The social construct truly belongs to the race, but the concrete particular is inaccurately modeled by this construct.

¹⁸ See Astuti and Knight (2008). This problem arises when a belief and its opposite are simultaneously held, typically without the agent's awareness of the problematic situation.

References

- Astuti, R. and N. Knight. (2008) "Some Problems with Property Ascription", *Journal of the Royal Anthropological Institute*, 14: S142-S158.
- Friend, S. (2007) "Fictional Characters", *Philosophy Compass* 2/2, Oxford: Blackwell, pp. 141-156.
- Kripke, S. (1972) *Naming and Necessity*, Oxford: Blackwell.
- Lewis, D. (1978) "Truth in Fiction", *American Philosophical Quarterly*, 15: 37-46.
- Quine, W. V. O. (1953) *From a Logical Point of View*, Second Edition, Cambridge, Mass.: Harvard University Press.
- Salmon, N. (1998), "Nonexistence", *Nous*, 32 (3): 277-319.
- Schiffer, S. (1996), "Language-Created Language-Independent Entities", *Philosophical Topics*, 24 (1): 149-167.
- Sainsbury, R. M. (2010) *Fiction and Fictionalism*, London: Routledge.
- Searle, J. (1975), "The Logical Status of Fictional Discourse", *New Literary History*, 6 (2): 319-332.
- Thomasson, A. L. (1999) *Fiction and Metaphysics*, Cambridge: Cambridge University Press.
- Thomasson, A. L. (2003) "Speaking of Fictional Characters", *Dialectica*, 57 (2): 205-223.
- Van Inwagen, P. (1977) "Creatures of Fiction", *American Philosophical Quarterly*, 14: 299-308.
- Voltolini, A. (2006) *How Ficta Follow Fiction*, Dordrecht: Springer.
- Walton, K. (1990) *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, Mass.: Harvard University Press.
- Zalta, E. N. (1988) *Intensional Logic and the Metaphysics of Intentionality*. Cambridge, Mass.: The MIT Press.