Artifactualism and Authorial Creation

Zsófia Zvolenszky

Department of Logic, Institute of Philosophy, Eötvös University (ELTE)

ABSTRACT. Artifactualism about fictional characters, positing Harry Potter as an abstract artifact created by J. K. Rowling, has been criticized on the grounds that the idea of creating such objects is mysterious and problematic. In the light of such qualms, it is worth homing in on an argument in favor of artifactualism, showing that it is the best way to include the likes of Harry Potter in our ontology precisely because it incorporates authorial creation. To that end, I will be exploring Kripke's fleeting remarks in his "Naming and Necessity" lectures (1972, 156–7) about expressions like 'unicorn' and 'Harry Potter'. Elsewhere, Kripke motivates artifactualism by suggesting that incorporating authorial creation (as artifactualism does) is a move that is intuitive and natural; but beyond this, he doesn't provide any arguments in favor of such a move. My purpose in this paper is to construct such an argument based on considerations about Kripke's general view about proper names, in particular, his seminal causal-historical chain account of reference determination.

I.

Why insist that authors create fictional characters? It does seem natural to say (1):

(1) Harry Potter was created by J. K. Rowling.

Artifactualism about fictional characters, positing Harry Potter as an abstract artifact created by J. K. Rowling, takes (1) at face value. Like other forms of realism about fictional characters, artifactualism posits an ontology that

---

1 For brevity's sake, I'll suppress the qualification 'about fictional characters' and will simply talk of realism, irrealism, Meinongianism, nonactualism, artifactualism, Platonism. Whenever these labels appear unqualified, they are shorthand for theories about fictional characters.
includes the likes of Harry Potter. But realism is not our only option; we could also accept an irrealist analysis of (i) that doesn’t take it at face value: “J. K. Rowling wrote a body of fiction in which Harry Potter is a specific character”. Quite independently of irrealism, several philosophers have had serious qualms about taking (i) at face value: Brock (2010, 338) sets out to “explain why creationism about fictional characters [the view that fictional characters exist by being created by their author(s)] is an abject failure. It suffers from the same problem as theological creationism: the purported explanation is more mysterious than the data it seeks to explain” because it cannot offer a satisfactory account of the spatial and temporal dimensions of fictional characters, for example, their moment of creation. Yagisawa (2001, 154) argues that the most influential creationist views (by Searle and van Inwagen) “are ultimately unsuccessful in establishing creationism”; more generally, he thinks no view on which fictional characters exist can do justice to our intuition that a claim like “Harry Potter doesn’t exist” is true and is entailed by the true “Harry Potter is a fictional character”. In the light of such doubts about creationism, it is worth homing in on an argument for artifactualism (a form of creationism), showing that it is the best form of realism one could adopt precisely because it incorporates authorial creation. The goal of this paper is to expound such an argument.

First, let’s take stock of the various realist positions. We may, along with Mark Sainsbury (2010, 44–114), distinguish three realist alternatives about fictional characters: there really are such things just as there are ordinary concrete objects occupying space and time; but unlike those ordinary objects like cups, saucers and the Big Ben, ...

- fictional characters don’t exist, according to Meinongianism about fictional characters;  

2 An argument for fictional characters as objects created by people is noteworthy in the light of Brock’s (2010, 340–342) criticism. He calls this the Fundamental Thesis: “Fictional characters, to the extent that there are any, are genuinely created by the authors of the works in which their names (or designating descriptions) first appear.” Brock then remarks that “arguments in support of the fundamental thesis are almost completely lacking”. In this paper, I set out to produce precisely this sort of argument.

3 Parsons (1980) is a contemporary proponent of Alexius Meinong’s (1904) eponymous theory.
fictional characters are not actual but merely possible, according to nonactu-
nalism; and

fictional characters are not concrete but abstract, created by the activities of
authors according to artifactualism.

How might the various forms of realism handle (1)? On this point, artifac-
tualism shows a clear edge relative to its two rivals. According to Sains-
bury (2010, 61–63, 82–85), the real advantage of artifactualism concerns
its response to the so-called selection problem: upon introducing the name
‘Harry Potter’ in her novel, how does J.K. Rowling manage to select one
rather than another among the countless candidate objects? According to
Meinongianism, there are countless nonexistent candidates; according to
nonactualism, there are countless merely possible, nonactual candidates.

Sainsbury (2010, 63) doesn’t see “how a Meinongian can offer any sensible
account of how an author’s or reader’s thoughts are supposed to engage
with one rather than another nonexistent entity”. We are about to see
that a more decisive objection emerges against the Meinongian once we
consider the difficulties that the nonactualist encounters when it comes to
the selection problem and other problems.

In the “Addenda” to his “Naming and Necessity” lectures, Kripke (1972,
156–7) motivates two theses for expressions like ‘unicorn’ and ‘Harry Pot-

---

4 Lewis (1978) put forth such a view. This position is sometimes called possibilism
about fictional characters. See also Kripke’s earlier (1963) view about Sherlock Holmes.

Salmon (1998), Thomasson (1999) are prominent proponents who hold that authors’ cre-
ative process of writing novels, stories, etc. creates fictional characters. This position is
sometimes called creationism about fictional characters.

There is a position in logical space for holding that fictional characters are abstract but
exist timelessly, and authors don’t create but discover them—we might call such a view
Platonism about fictional characters. Zalta’s (1983) unorthodox neo-Meinongian proposal
can be considered an instance of such an account. The only kind of abstract-object theory
I will consider in this paper is artifactualism, given the overwhelming popularity and
attention that this position has been enjoying (compared to Platonism), as well as the
advantages that I think it has over rival theories (Platonism included) precisely because
it treats fictional characters as human-created objects. The arguments expounded here
carry over to Platonism also, but I will relegate discussion of that to footnotes.

6 Ultimately, Sainsbury (2010) rejects artifactualism in favor of irrealism. For a re-
response strategy that the artifactualist can adopt to fend off Sainsbury’s criticism see
ter’:

- The metaphysical thesis: There is no basis for counting any merely possible object as Harry Potter, Sherlock Holmes, a unicorn, etc.
- The epistemological thesis: There is no basis for counting any actual object as Harry Potter, Sherlock Holmes, a unicorn, etc.

In the metaphysical thesis, Kripke’s target seems to be the nonactualist (given that he is talking about merely possible entities, the nonactualist’s candidates for fictional characters). At the end of the paper, we will see, however, that both theses bear on Meinongianism also. Along the way, we will also see that the two arguments are at root intimately connected.

Elsewhere (see Kripke 1973/2011, 1973/2013), Kripke motivates artifactualism by suggesting that incorporating authorial creation (as artifactualism does) is a move that is intuitive and natural; but beyond this, he doesn’t provide any arguments in favor of such a move. My purpose in this paper is to construct such an argument based on considerations about Kripke’s general view about proper names, in particular, his seminal causal-historical chain account of reference determination (Kripke 1972).

2.

Behind Kripke’s metaphysical thesis is what we might call the insufficient-specificity problem.7 The Harry Potter novels specify many details about Harry; but they also leave a lot of other details unspecified, for example, which of various parental cells Harry came from. Due to such lack of specificity in the novels, we have no basis for deciding between two distinct merely possible candidates (they originate from distinct zygotes, say) that are just like Harry is described in the novels, which of them is Harry Potter. Notice that it is in part due to insufficient specificity in the novels that Sainsbury’s selection problem arises—for the Meinongian as well as the nonactualist.

The epistemological thesis turns out to generate an even deeper problem for the nonactualist, one that we shall see (at the end of the paper)

7 Kaplan also emphasizes insufficient specificity as an obstacle to naming nonexistents (1973, 506; 1989, 609).
affects the Meinongian also. Behind the epistemological thesis is what we might call the *coincidental-resemblance problem*, which Kripke discusses in connection with the mythical species of unicorn:

...the mere discovery of animals with the properties attributed to unicorns in the myth would be no means to show that these were the animals the myth was about: perhaps the myth was spun out of whole cloth and the fact that animals with the same appearance actually existed was mere coincidence. In that case, we cannot say that the unicorns of the myth really existed; we must also establish a historical connection that shows that the myth is *about* these animals. (Kripke 1972, 157, emphasis in the original)

Kripke is making two points here: even if we find animals qualitatively like the unicorns of the myth, that wouldn’t justify counting them as unicorns given (i) the lack of historical connection between the newly found species and the use of the expression ‘unicorn’; and given that (ii) the unicorn myth was “spun out of whole cloth”, not created in the right way, to make the term apply to the newly found species. The upshot of (i) and (ii): we would have no more than mere qualitative coincidence between unicorns as described in the myth and the actual species discovered. And for a proper name, reference takes more than coincidental resemblance, so we don’t have any candidate actual objects to count as unicorns.  

In the case of the expression ‘unicorn’, the coincidental-resemblance problem thus arises as a result of two distinct problems: (i) **historical unconnessedness** and (ii) **unsuited mode of introduction**. Pure myth-making mode and pure fiction-writing mode both give rise to expressions that aren’t introduced in the right way to refer to actual objects. Right after the passage above, Kripke (1972, 157–158) repeats the same point with respect to ‘Sherlock Holmes’ also: “it is theoretically possible though in practice fantastically unlikely, that Doyle was writing pure fiction with only coincidental resemblance to [an] actual man”. A crucial consideration emerges from these fleeting remarks about unicorns and Sherlock Holmes: given (ii) the

---

8 Kaplan quotes Harry Deutsch: “reference is no coincidence” (Kaplan 1989: 608).

9 One may accept this point with respect to fictional objects but not mythical objects. The points in my paper do not require a stance about mythical characters, which raise a host of distinctive issues, see for example Caplan 2004, Braun 2005.
way the myth/fiction was created, and (i) the fact that we encounter historical unconnectedness, the result is that we find no more than coincidental resemblance to actual objects.

Both theses and all the problems considered so far have taken it for granted that the candidate objects to count as Harry Potter are concrete, spatiotemporal objects. It is therefore well to keep this qualification in mind. For example, for (ii) we get: the fiction-writing mode in which the expression ‘Harry Potter’ had been introduced into the language is unsuited for the name to refer to an actual concrete, spatiotemporal object. For (i) we get: actual, concrete, spatiotemporal objects as potential referents for the name are historically unconnected to the introduction and subsequent use of ‘Harry Potter’.

It’s crucial to note that of the two problems (i) and (ii), unsuited mode of introduction is the more fundamental one, explaining historical unconnectedness of the relevant sort. Given that (ii) Rowling’s intention was to create a fictional character rather than refer to a flesh-and-blood person with introducing the name ‘Harry Potter’, (i) ‘Harry Potter’ was never historically linked (in the relevant way) to an actual orphaned boy wearing glasses, with a Z-shaped scar on his forehead, growing up in suburban England learning wizardry in a boarding school, and so on, and the name cannot refer to any actual concrete boy with spatiotemporal dimensions.

The unsuited-mode problem generalizes to concrete, spatiotemporal objects of all sorts, merely possible ones included; this way, we get:

*the unsuited-mode problem generalized*: the fiction-writing mode of introducing proper names into the language is unsuited for them to have as their reference concrete, spatiotemporal objects, whether they be actual or merely possible.

It is well to generalize in the same way the coincidental-resemblance problem also:

*The coincidental-resemblance problem generalized*: there is no more than mere qualitative coincidence between concrete, spatiotemporal objects (whether they be actual or merely possible) and fictional characters as described in works of fiction.
Therefore (in the light of the generalization to merely possible objects), as we dig deeper, the pair of problems behind the epistemological thesis turn out to target nonactualism.

As before, in the case of ‘Harry Potter’, the unsuited-mode problem generalized underlies the generalized coincidental-resemblance problem. And both problems are in the background of the metaphysical thesis also: the generalized unsuited-mode problem provides the following additional reason for holding the metaphysical thesis. If the character of Harry Potter is not fully specified in the novels, then what grounds do we have at all for choosing between two distinct merely possible concrete, spatiotemporal objects which to count as Harry Potter when, given J. K. Rowling’s fiction-writing mode of introducing ‘Harry Potter’, it would be a matter of sheer coincidental resemblance for the name to refer to either of those candidate objects? With respect to names from fiction, the unsuited-mode problem (and in its wake, the coincidental resemblance problem) therefore raises a key issue underlying both the metaphysical and the epistemological theses discussed by Kripke; this is a striking detail to bring to the surface given that Kripke mentions the unsuited-mode problem in passing only (saying no more than the two half-sentences quoted above), devoting far more attention to the metaphysical thesis.

3.

Just how bizarre the idea of reference based on coincidental resemblance is—the conception of reference for ‘Harry Potter’ to which the nonactualist is committed—can be brought out based on considerations about nonfictional names that fail to refer. The French astronomer Le Verrier put forth a hypothesis about the existence of an intra-Mercurial planet which he named ‘Vulcan’, to explain perturbations in the orbit of Mercury. There were various independent sightings mistakenly believed to be of Vulcan before enthusiasm dwindled; by 1916, Einstein’s general theory of relativity confirmed that the perturbations were produced by the gravitational field of the Sun; there was no intra-Mercurial planet at all; the Vulcan-hypothesis was refuted; ‘Vulcan’ turned out not to refer to anything.

What about a counterfactual situation in which the Vulcan-hypothesis is a success story? Imagine a counterfactual scenario with the laws of
physics slightly different, and there being an intra-Mercurial planet affecting the orbit of Mercury; Le Verrier puts forth his hypothesis; there are sightings converging on the planet, which comes to be called ‘Vulcan’, the name featured in Le Verrier’s prior hypothesis. But that is not our term ‘Vulcan’ that comes to name the counterfactual planet, but a different one. It is preposterous to think that in coining the name in the actual world, Le Verrier managed to name that counterfactual object even though his naming attempt failed in the actual world. ‘Vulcan’ might have been a success story just as ‘London’ might have been introduced as a name for a river instead of a city; but all that is irrelevant to how and whether these strings, as parts of our language, were introduced and subsequently used. Le Verrier strove to name an actual concrete, spatiotemporal object; due to his failure to do so, he didn’t by coincidence name a nonactual concrete, spatiotemporal object (as the nonactualist would have it); doing so was no part of his intention. So ‘Vulcan’ doesn’t refer to any concrete objects in any counterfactual situations. Kaplan (1973, 506–508) makes this point eloquently with respect to a mythical name like ‘Pegasus’. But what is far more interesting is that the point holds for ‘Vulcan’! We can say the following about this name of our language, as well as other proper names intended for concrete objects or for fictional characters: if it cannot make it here, it won’t make it anywhere. If the name doesn’t manage to refer to a concrete, spatiotemporal object here, in the actual world, it doesn’t refer to such an object in other possible worlds either. Elsewhere (Zvolenszky 2007), I call this the inverse-Sinatra principle for proper names.

The inverse Sinatra principle is quite general, covering names like ‘Vulcan’, ‘Pegasus’, and ‘Harry Potter’. And the reason why these names don’t make it anywhere given that they cannot make it here (in the actual world), is because nonactual concrete objects are, at best, coincidentally similar Vulcan, Pegasus and Harry Potter, as these are described in various bodies

---

10 See Kripke (1971, 145; 1972, 77, 102–3, especially fn. 51).

11 Even an irrealist about fictional characters can, based on the considerations about Vulcan and unicorns above, accept the inverse-Sinatra principle.

12 Frank Sinatra sang about New York City: “If I can make it there, I’ll make it anywhere”. In the inverse-Sinatra principle (to keep it parallel with the song), I use the modal auxiliary ‘can’, by which I mean (as the song’s ‘can’ does) ‘is able to’; I don’t mean metaphysical possibility. Thanks to Nathan Wildman for prompting me to clarify this.
of text. We thus have a nonfictional variant of the coincidental resemblance problem.

Notice that ‘Vulcan’ and ‘Harry Potter’ differ in one crucial detail: for the case of ‘Vulcan’, the unsuited-mode problem doesn’t arise. Le Verrier’s intention had been to introduce ‘Vulcan’ for a concrete, spatiotemporal object; so a historical connection, if there had been one, linking uses of the name to an actual concrete object, could have served to fix the reference of ‘Vulcan’, circumventing coincidental-resemblance-related qualms. A historical connection can be secured in the actual world only—there is absolutely no historical connection between our use of ‘Vulcan’ and a merely possible concrete, spatiotemporal object. And in the absence of an actual historical connection, qualms about coincidental resemblance do arise, leading to the metaphysical thesis about ‘Vulcan’: if the specification of Vulcan isn’t complete, allowing that several distinct merely possible concrete objects fit the specification equally, then we have no basis for counting any one of them as Vulcan. (Notice that here, as before, my argument leading to the metaphysical thesis for Vulcan was crucially linked to considerations about coincidental resemblance and historical unconnectedness, which were originally identified behind the other thesis—the epistemological one. With respect to ‘Vulcan’, too, we see that the two theses are intimately connected.)

The foregoing observation allows us to highlight a more general point of advantage for the artifactualist position over both Meinongianism and nonactualism.

According to artifactualism, Harry Potter is an actual object. Yet the fact that he is an actual artifact makes room for a certain kind of causal-historical dependence on the physical world: in the 1990s, J. K. Rowling’s creative activities bring it about that Potter is an actual abstract object. The sort of dependence in place allows Harry Potter qua abstract artifact to be the kind of referent for Rowling’s name ‘Harry Potter’ with respect to which issues having to do with historical unconnectedness and, in turn, coincidental resemblance, and, in turn, the epistemological thesis, do not arise. (Notice that before, we noted that for names of fictional characters, no historical connection to concrete, spatiotemporal objects is of the relevant, reference-fixing sort. Meanwhile, the point made here is that for the artifactualist, a historical connection to an actual abstract artifact is
precisely what fixes the reference of ‘Harry Potter’.

By contrast, alternative realist accounts that make Harry Potter a concrete object whose existence does not causally depend on us either because the object is nonexistent (according to Meinongianism) or because it is nonactual (according to nonactualism), face a challenge. First, these theorists have to explain why those objects are candidates of the right ontological status to count as the referents of ‘Harry Potter’. As we have already seen, on this point, the nonactualist founders already. The Meinongian can get past this hurdle: he may suggest that his nonexistents are objects of thought and hence have just the right sort of ontological status to be suitable targets of authors’ intended reference. But on the next hurdle the Meinongian stumbles: if his nonexistent objects are of a suitable sort as objects of fiction-writing, what historical connection is there to account for Rowling’s ‘Harry Potter’ referring to one of countless nonexistent candidate objects (each equally faithful to the way Potter is depicted in the novels but varying in details left unspecified—about sock color, etc.)? The Meinongian cannot provide such a historical connection: causal-historical connection between his timelessly nonexistent objects and actual concreta (like authors) is extremely problematic, downright unintelligible even. And because of historical unconnectedness, the Meinongian is confronted with qualms about having to work with no more than coincidental resemblance between Harry Potter as specified in the novels, and various qualitatively similar Meinongian nonexistents. And, on the one hand, coincidental resemblance does not suffice for reference, according to the epistemological thesis; and, on the other hand, with insufficiently specified characters like Harry Potter, coincidental resemblance leaves room for the metaphysical thesis (and also the selection problem) to arise.\(^{13}\)

\(^{13}\) This line of argument brings to the fore why the only abstract-theory contender we considered for fictional characters was artifactualism: it is the only view according to which Harry Potter is created and hence historically linked to goings on in the actual world. Platonism, a theory according to which Harry Potter is a timelessly existing abstract object (akin to numbers, sets), would, like Meinongianism and nonactualism, run into problems with historical unconnectedness and hence coincidental resemblance, and, in their wake, the metaphysical and epistemological theses. For an attempt to combine the advantages of artifactualism and Meinongianism, see Zalta’s (2000, 2006).
Once fleshed out, Kripke’s (1972) fleeting remarks about fictional characters can be summarized as follows: qualitative resemblance is insufficient to determine the reference of a proper name; a causal-historical connection between names and their referents is necessary to determine to whom or to what proper names refer. For names of actual concrete objects like ‘J.K. Rowling’ and ‘London’, this overarching lesson transparently emerges from the second lecture of *Naming and Necessity*. It is considerably less transparent that Kripke reiterates the very same lesson for names of fictional characters. Of the forms of realism considered, artifactualism is the only one that can heed this lesson.¹⁴

References


¹⁴ This paper has benefited from comments by participants at the conference *Realism within Phenomenology and within Analytic Philosophy* held at Kaposvár University (Hungary) in January 2012, as well as the conference *Modal Metaphysics: Issues on the (Im)Possible II* held at the Slovak Academy of Sciences in October 2014. Special thanks are due to Tibor Bárány, Zsolt Bátori, Zsolt Kapelner, András Simonyi and Ádám Tuboly for many thoughtful and incisive suggestions on an earlier draft of this paper. This research has been supported by the Integrative Argumentation Studies Grant No. 19648 received from the Hungarian Scientific Research Fund (OTKA).


