THE EXTRAORDINARY IMPOSSIBILITY OF SHERLOCK HOLMES*

Ben Caplan

Abstract: In an addendum to Naming and Necessity, Saul Kripke argues against his earlier view that Sherlock Holmes is a possible person. In this paper, I suggest a non-standard interpretation of the addendum. A key feature of this non-standard interpretation is that it attempts to make sense of why Kripke would be rejecting the view that Sherlock Holmes is a possible person without asserting that it is not the case that Sherlock Holmes is a possible person.

“Nor is an impossibility of the ordinary kind involved.”

Kripke, “Vacuous Names and Fictional Entities” (2011)

1 Introduction

According to A Study in Scarlet (1887) and other stories by Sir Arthur Conan Doyle, Sherlock Holmes is a person. But Sherlock Holmes is not really a person. That is, it is not the case that, outside of the fiction, Sherlock Holmes is a person. Still, on some views, it is the case that, outside of the fiction, Sherlock Holmes is a possible person. Let’s call the claim that Sherlock Holmes is a possible person possibilism (about Sherlock Holmes).

In “Semantical Considerations on Modal Logic,” Saul Kripke (1963) endorses possibilism. He says, “Holmes does not exist, but in other states of affairs, he would have existed” (85). But, in an addendum to Naming

* It is unlikely that what I say here is entirely novel. While I was a graduate student at UCLA, from 1996 to 2002, the Philosophy of Language Workshop, led by David Kaplan and Joseph Almog, discussed Kripke’s John Locke Lectures, which were available in manuscript form. (We didn’t have access to Kripke’s [2011] “Vacuous Names and Fictional Entities.” But, although Kaplan didn’t participate in the discussion of the paper [see Kripke et al. 1974b], he was at a conference at which Kripke presented the paper in 1973 [see Kripke et al. 1974a].) I don’t think I really understood what was going on in the Language Workshop at the time. Still, if anything I say here is right, then it is possible that I am recollecting it from the Language Workshop.
and Necessity, Kripke (1972, 156–158) rejects his earlier possibilism. He says,

I hold the metaphysical view that, granted that there is no Sherlock Holmes, one cannot say of any possible person that he would have been Sherlock Holmes, had he existed. Several distinct people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes, but there is none of whom we can say that he would have been Holmes, had he performed these exploits. For if so, which one? (158; italics in original)

In this paper, I suggest a non-standard interpretation of the addendum. In Section 2, I present a standard interpretation. In Section 3, I offer some reasons for thinking that it is incorrect. In Section 4, I discuss an argument that I take Kripke to endorse in “Vacuous Names and Fictional Entities.” (In Section 5, I offer some reasons for thinking that an alternative interpretation of “Vacuous Names and Fictional Entities” is incorrect.) And, in Section 6, I suggest that Kripke might be endorsing the same argument in the addendum.

I should say three things at the outset. First, in interpreting the addendum, I am setting aside Kripke’s (2011; 2013) fictional realism, from “Vacuous Names and Fictional Entities” and Reference and Existence, according to which Sherlock Holmes is an actual non-person.1 I am interpreting the addendum in light of those works.2 And Kripke does argue for fictional realism in those works, but he does not do so until after he argues against possibilism.3 So I do not think that his arguments against possibilism in those works presuppose his fictional realism.

Second, some of what I say here is tentative or incomplete. In particular, I take the case against the standard interpretation (in Section 3) to be stronger than the case for the non-standard interpretation (in Section 6); and I take the discussion of “Vacuous Names and Fictional Entities” (in Section 4 and Section 5) to raise a host of further issues that I do not begin to address here.

And, third, most of what I say here is exegetical. The focus of the paper is largely on the merits of different interpretations of the addendum rather than on the merits of the arguments that different interpretations of the addendum attribute to Kripke there. (But the questions are not entirely separate, since an interpretation that attributes a bad argument to Kripke is probably a bad interpretation.)

---

1 For a contrary exegetical strategy, see Zalta 2006, 599.
2 See Section 6, especially footnote 36.
3 Compare Kripke 2011, 59 and 62–65; or compare Kripke 2013, 40–42 and 69–78.
2 The “Too Many Candidates” Interpretation

Consider the following argument for the claim that numbers are not sets.\(^4\)

**Multiplicity:** There is more than one set that one could identify the number 2 with: for example, \{\{\emptyset\}\} and \{\emptyset, \{\emptyset\}\}.

**Arbitrariness:** There is no reason for one of these sets rather than another to be the number 2.

**No Set:** The number 2 is not a set. (From Multiplicity and Arbitrariness)

One can construct a parallel argument for the negation of possibilism. Let’s call this argument the “too many candidates” argument.\(^5\)

**The “too many candidates” argument:**

**Multiplicity:** There is more than one possible person who could be identified with Sherlock Holmes.

---

\(^4\) The argument is inspired by Benacerraf 1965, 5, 62, 67. The argument is, of course, enthymematic. Its missing premises are (i) that, if there is more than one set that one could identify the number 2 with, and there is no reason for one of those sets rather than another to be the number 2, then the number 2 is not one of those sets rather than another; and (ii) that, if the number 2 is not one of those sets rather than another, then the number 2 is not a set.

\(^5\) The “too many candidates” argument is also enthymematic. Its missing premises are (i) that, if there is more than one possible person who could be identified with Sherlock Holmes, and there is no reason for one of those possible people rather than another to be Sherlock Holmes, then Sherlock Holmes is not one of those possible people rather than another; and (ii) that, if Sherlock Holmes is not one of those possible people rather than another, then Sherlock Holmes is not a possible person. The first explicit presentation of the “too many candidates” argument in the literature that I know of is in Thomasson 1999, 17–18. (See also Thomasson 1999, 44–45, although she does not explicitly state Arbitrariness there.) For equally explicit presentations of the argument, see van Inwagen 2008, 43; Wright 2014, 292. See also Sainsbury 2009, x, 85 and Stone 2010, 639, although neither Sainsbury nor Stone explicitly states Arbitrariness. (Thomasson, van Inwagen, Wright, and Stone endorse the argument. Stone does not. See Stone 2010, 641–642.)

There are historical antecedents, among them Kaplan 1973, 505–508; Donnellan 1974, 24–25; Plantinga 1974, 154–155, 159–163; Lycan 1994, 119. (See also Lycan 2015, 26.) But Kaplan, Donnellan, and Plantinga do not explicitly assert Arbitrariness. Plantinga might not explicitly assert Multiplicity either. What Donnellan argues is that Sherlock Holmes (or, in his example, Homer) does not exist in a possible world in which multiple people satisfy the relevant descriptions. At least in principle, this leaves open the possibility that Sherlock Holmes (or Homer) exists in another possible world, in which someone uniquely satisfies the relevant descriptions. Lycan explicitly asserts an analogue of Arbitrariness; but, like Donnellan, he is considering a case in which multiple candidates exist in the same possible world. And the argument that he presents is against a form of haecceitism about merely possible objects rather than against possibilism. (Lycan [1994, 132 n. 22] attributes the argument to Greg Currie. Lycan does not endorse the argument. See Lycan 1994, 122–123; 2015, 28–31.)

On the connection between the “too many candidates” argument and the problem of the many, see Stone 2010, Wright 2014. (On the problem of the many, see Unger 1980, Lewis 1993.)
Arbitrariness: There is no reason for one of these possible people rather than another to be Sherlock Holmes.

The negation of possibilism: Sherlock Holmes is not a possible person. (From Multiplicity and Arbitrariness)

The idea behind the “too many candidates” argument is that there are multiple candidates for being Sherlock Holmes, and the candidates are equally good, so Sherlock Holmes isn’t any of them.

One reply to “the too many candidates” argument is to deny Arbitrariness. On this view, there is a unique closest possible world that satisfies some condition (e.g., being a possible world in which there is a detective who lives at 221B Baker Street); and the possible person in that possible world who satisfies a related condition (e.g., being a detective who lives at 221B Baker Street) is more eligible than the other candidates (see Krasner 2001). Let’s call this reply the eligibility reply.

Another reply to the “too many candidates” argument is to say that, if there is a possible person $h$ such that Sherlock Holmes $= h$, then it is a brute fact that Sherlock Holmes $= h$. (A brute fact is a fact that obtains but not in virtue of any other fact.) If Sherlock Holmes $= h$, then the fact that Sherlock Holmes $= h$ just is the fact that Sherlock Holmes is Sherlock Holmes, and presumably that fact is brute. On this view, the inference from Multiplicity and Arbitrariness to the negation of possibilism is invalid. Just because all the candidates are equally eligible, that does not mean that none of them is Sherlock Holmes. For, if one of them were Sherlock Holmes, then it would not need to be more eligible than the other candidates to be Sherlock Holmes; it would just be a brute fact that it is Sherlock Holmes. Let’s call this reply the brute identity fact reply.

I’m sympathetic to the brute identity fact reply, but many would not be. For example, Benacerraf (1965, 57) considers the possibility that there is “a particular set of sets . . . , which is really the numbers, but such that there exists no argument one can give to establish that it, and not, say, . . . [another set of sets] is really the numbers” (italics in original). He says, “It seems altogether too obvious that this latter possibility borders on the absurd. If the numbers constitute one particular set of sets, and not another, then there must be arguments to indicate which” (57–58).

---

6 For some possibly congenial remarks, see Howell 1979, 172–174. Wright (2014, 294–295) discusses a version of this view on which, for some candidate, it’s a brute fact (i.e. a fact that obtains but not in virtue of any other fact) that it is the most eligible.

7 For similar reasoning, see Salmon 1986, 110–114; 1987. See also Caplan and Muller 2015, 179–180. The reasoning behind the brutality of identity facts parallels Kripke’s (1971, 136; 1980, 3) argument for the necessity of identity. For a view on which identity facts are not brute, see Fine Forthcoming (Damiano Costa has defended a similar view in Costa Unpublished).

8 Or, if the argument is enthymematic, then missing premise (i) is false. See footnote 5.

9 Modulo the assumption that there is more than one set of sets that one could identify the numbers with, Benacerraf is here endorsing a generalization of the contrapositive of missing premise (i) from footnote 4.
On one interpretation, Kripke (1972, 157–158) endorses the “too many candidates” argument in the addendum. Let’s call this interpretation the “too many candidates” interpretation. Kripke seems to endorse Multiplicity: he says, “Several distinct people . . . might have performed the exploits of Holmes” (1972, 158). He seems to suggest Arbitrariness when, after saying “there is none of whom we can say that he would have been Holmes,” he asks “For if so, which one?” (158; italics in original). And he seems to endorse the negation of possibilism: he says, “one cannot say of any possible person that he would have been Sherlock Holmes” (158; italics in original).

Kripke makes analogous remarks about unicorns in the addendum. He seems to endorse an analogue of Multiplicity: he says that there are “several distinct hypothetical species . . . which would have the external appearances postulated to hold of unicorns in the myth of the unicorn” (156–157). And he seems to endorse an analogue of the negation of possibilism: he says, “there is no actual or possible species of which we can say that it would have been the species of unicorns” (157).

3 Against the “Too Many Candidates” Interpretation

There are two reasons for thinking that the “too many candidates” interpretation is incorrect.

The first reason is that Kripke does not explicitly state one of the premises of the “too many candidates” argument—namely, Arbitrariness—in the addendum. In presenting the “too many candidates” argument, others explicitly state Arbitrariness.

[T]here seem to be . . . no means to choose among them [the candidates]. . . . Selecting any one as identical with a particular character seems hopelessly arbitrary. (Thomasson 1999, 17–18)

10 See Thomasson 1999, 17–18, 44–45, 156 n. 28; Stone 2010, 639; Wright 2014, 292. The “too many candidates” interpretation is at least suggested by van Inwagen (2008, 43). It might or might not be suggested by Cook (1985, 303) and Salmon (1998, 292, 314 n. 31). For other interpretations, see Zalta 2006, 599–604; Ben-Yami 2010; Liebesman 2014. Some of these other interpretations face some of the same objections as the “too many candidates” interpretation. See footnote 12 and footnote 13.

11 On the analogy between Sherlock Holmes and unicorns, see Kripke 1972, 157; 2011, 59 n. 13; 2013, 52–53. See also Section 4.

12 On Liebesman’s (2014, 311–312) interpretation, Kripke might endorse Arbitrariness. Zalta (2006, 599) might also attribute Arbitrariness to Kripke, although Arbitrariness is not among the claims that he uses to fill in “the metaphysical blanks underlying Kripke’s argument.” See Zalta 2006, 604. I place less weight on this first reason for thinking that the “too many candidates” interpretation is incorrect than I do on the second. (On the interpretation I suggest in Section 6, Kripke is also less than maximally explicit.)
All those other-worldly objects are equally good candidates for one or the other of the offices ‘being Holmes’ and ‘being Vulcan.’ (van Inwagen 2008, 43)

There is the claim of arbitrariness . . . there is no principled reason to privilege one [candidate] over another. (Wright 2014, 292; italics in original)

By contrast, the closest Kripke comes to stating Arbitrariness is asking a rhetorical question in arguing against possibilism: “[I]f so, which one?” (1972, 158). (In arguing against the claim that the species unicorn is a possible species, Kripke does not ask the analogue of the rhetorical question.)

The second reason for thinking that the “too many candidates” interpretation is incorrect is that Kripke does not assert the conclusion of the “too many candidates” argument—namely, the negation of possibilism—in the addendum. In presenting the “too many candidates” argument, others assert the negation of possibilism (or of a more general claim).

Not only is there no actual Sherlock Holmes, there is also no possible person, such that if he were actual he would be Sherlock Holmes. (Thomasson 1999, 44–45)

Of all those other-worldly objects . . . Holmes and Vulcan cannot be any of them, and therefore neither Holmes nor Vulcan exists in any possible world. It follows that it is false of each of them that it might have existed. (van Inwagen 2008, 43; italics in original)

Fictional characters . . . could not be possibilia. (Wright 2014, 292)

By contrast, what Kripke says in the addendum is that we cannot assert possibilism. He does not say that there is no possible person who would have been Sherlock Holmes. Rather, he says, “one cannot say of any possible person that he would have been Sherlock Holmes” (1972, 158). And he does not say that there is none who would have been Holmes, had he performed certain exploits. Rather, he says, “there is none of whom we can say that he would have been Holmes, had he performed these exploits” (158).

Similarly, in the case of unicorns, Kripke does not say that none of the distinct mythical species would have been the unicorns. Rather, he says, “one cannot say which of these distinct mythical species would have been

---

13 On Zalta’s (2006, 599) and Liebesman’s (2014, 306 n. 2) interpretations, Kripke asserts the negation of possibilism.
14 Here I agree with Ben-Yami (2010, 175), who reports that Kripke “emphasized in discussion” that he was being “cautious” in this way.
15 Here and subsequently, boldface is mine. In this paragraph and the next, I have removed Kripke’s italics.
the unicorns” (157). He does not say that no actual or possible species would have been the species of unicorns. Rather, he says, “there is no actual or possible species of which we can say that it would have been the species of unicorns” (157). And he does not say that no merely possible world (or counterfactual situation) is one in which there would have been unicorns. Rather, he says, “no counterfactual situation is properly describable as one in which there would have been unicorns” (156).

There is a difference between rejecting a claim and asserting its negation. In particular, one can reject a claim without asserting its negation. I think that this is what Kripke is doing in the addendum: I think that he is rejecting possibilism and other claims that he is arguing against without asserting their negations.

To see the difference between rejecting a claim and asserting its negation, suppose that there is something defective about a claim (e.g., it might lack a truth-value). There might be something similarly defective about its negation, in which case one would not want to assert its negation. But one might still want to indicate that there is something defective about the original claim. And one can do that by rejecting it without asserting its negation. For example, if the claim that Sam stopped believing is neither true nor false because it has a false presupposition (Sam never believed in the first place), then perhaps its negation—the claim that Sam didn’t stop believing—is also neither true nor false, in which case one wouldn’t want to assert its negation. But perhaps one can indicate that there is something defective about the claim that Sam stopped believing by rejecting it without asserting its negation. To do this, one wouldn’t say, “Sam didn’t stop believing.” Instead, one might say, “One cannot say that Sam stopped believing” or perhaps even “One cannot say, of Sam, that she stopped believing.”

Closer to the case at hand, suppose that Nancy asserts a claim about possibility: she asserts that something—for example, that Bess and George are sisters—is possible. And suppose that this is equivalent to asserting that some proposition—namely, the proposition that Bess and George are sisters—is possibly true. Suppose that Carson and Ned are both arguing against Nancy’s claim. On the one hand, Carson agrees with Nancy that there is such a proposition as the proposition that Bess and George are sisters, but he disagrees with Nancy about whether that proposition is possibly true. So Carson asserts that the proposition that Bess and George are sisters is not possibly true. On the other hand, Ned (who is perhaps confused about whether Bess and George exist) disagrees with Nancy about whether there is such a proposition as the proposition that Bess and George are sisters in the first place. So Ned asserts that there is no such proposition as the proposition that Bess and George are sisters. In this case, Carson

---

16 See Parsons 1984. Thanks to Sam Cowling, Andreas Stokke, and Chris Tillman for help with the next three paragraphs.
does something like asserting the negation of Nancy’s claim, whereas Ned does something like rejecting Nancy’s claim without asserting its negation.

In *Reference and Existence*, Kripke suggests that his response to possibilism is closer to Ned’s response to Nancy’s claim than it is to Carson’s response to Nancy’s claim. Speaking of alleged propositions like the proposition that Sherlock Holmes lived on Baker Street (and, presumably, also the propositions that Sherlock Holmes exists and that Sherlock Holmes is a person), Kripke (2013, 42) says, “if I am not mistaken as to the status of these alleged propositions—and I don’t think there are any such propositions—then I am also not mistaken in saying that one cannot say that they would have been true of a certain hypothetical world: for there are no propositions to be true of this hypothetical world.”

Kripke makes it clear in a number of other places that he is not asserting a negation of a claim that he is arguing against. For example, at the beginning of Lecture I of *Naming and Necessity*, he says, “Perhaps according to me the truth should not be put in terms of saying that it is necessary that there should be no unicorns, but just that we can’t say under what circumstances there would have been unicorns” (1972, 24). Here, Kripke does not say that there are no possible worlds in which (or circumstances under which) there are unicorns. Rather, he says that we can’t say in which possible worlds (or under which circumstances) there are unicorns.

Similarly, in “Vacuous Names and Fictional Entities,” Kripke (2011, 69) says, speaking of Lewis Carroll’s (1871) poem “Jabberwocky,”

> Although we can say ‘there are no bandersnatches’ or ‘bandersnatches don’t exist,’ this plainly does not imply that we would know what it would be like for bandersnatches to have existed. Nor is an impossibility of the ordinary kind involved, such as the necessary nonexistence of round squares.

Here, Kripke does not say that it is impossible that there are bandersnatches, just as it is impossible that there are round squares. Rather, he suggests that the impossibility of bandersnatches, unlike the impossibility of round squares, would not be “an impossibility of the ordinary kind.”

And, in *Reference and Existence*, Kripke (2013, 47 n. 15) says,

> I should not be taken to be saying that it is impossible that there should have been dragons or that there should have been unicorns (like composite primes, or, according to me, water that is not H$_2$O). Rather, I should be taken to

---

17 Similarly, Kripke (2013, 40) says, “on my view, if statements containing ‘Sherlock Holmes’ express pretended propositions—or rather, pretend to express propositions—one can’t speak of a pretended proposition as possible.”

18 On the analogy between bandersnatches and unicorns, see Kripke 2011, 69; 2013, 52. Kripke (2011, 69, 72; 2013, 161) uses 1872 rather than 1871 as the year of publication for *Through the Looking-Glass, and What Alice Found There* and omits the hyphen in the title.
be saying] that the counterfactual possibility is ill-defined, given that there are no dragons or unicorns.19

Here, Kripke does not say that it is impossible that there are unicorns, just as it is impossible that there are composite primes or water that is not H$_2$O. Rather, he says that the possibility that there are unicorns is, in some sense, “ill defined.”

I take the passages from Naming and Necessity (both in Lecture I and in the addendum), “Vacuous Names and Fictional Entities,” and Reference and Existence to support the claim that Kripke is not asserting the negation of possibilism in the addendum. But, on the “too many candidates” interpretation, Kripke is asserting the negation of possibilism in the addendum. (The negation of possibilism is the conclusion of the “too many candidates” argument; and, on the “too many candidates” interpretation, Kripke endorses that argument there.) I take this to be a substantial reason for thinking that the “too many candidates” interpretation is incorrect.

4 The “No Modal Criterion for Reference” Argument

4.1 The Argument

Consider the following argument against possibilism. Let’s call this argument the “no modal criterion for reference” argument.20

The “no modal criterion for reference” argument:

No Modal Criterion for Reference: There is no modal criterion for reference for ‘Sherlock Holmes.’
Reject Reference: We cannot say, of any possible person $o$, that ‘Sherlock Holmes’ refers to $o$. (From No Modal Criterion for Reference)
Reject Possibilism: We cannot say, of any possible person $o$, that Sherlock Holmes is $o$. (From Reject Reference)

Leaving aside for now the question of what a modal criterion for reference for ‘Sherlock Holmes’ is, the idea behind the “no modal criterion for reference” argument is that there is something modally lacking about ‘Sherlock Holmes’; so, for any possible person, there is something defective

---

19 I have corrected the text by removing a comma after ‘water.’

20 The “no modal criterion for reference” argument is enthymematic. Its missing premises are (i) that, if there is no modal criterion for reference for ‘Sherlock Holmes,’ then we cannot say, of any possible person $o$, that ‘Sherlock Holmes’ refers to $o$; and (ii) that, if we cannot say, of any possible person $o$, that ‘Sherlock Holmes’ refers to $o$, then we cannot say, of any possible person $o$, that Sherlock Holmes is $o$. Exercise for the author: figure out whether the eligibility and brute identity fact replies to the “too many candidates” argument apply to the “no modal criterion for reference” argument. Conjecture from the author: the eligibility reply is irrelevant, but the brute identity fact reply blocks the inference from Reject Reference to Reject Possibilism.
about the claim that ‘Sherlock Holmes’ refers to that person and hence there is something defective about the claim that Sherlock Holmes is that person.

In “Vacuous Names and Fictional Entities,” Kripke offers the “no modal criterion for reference” argument. He says,

Under what circumstances . . . would Sherlock Holmes have existed . . . ? Well, not simply if someone or other would have done these things in the story, . . . because many possible people might have done the things in the story. In fact, some actual people might have done the things in the story, if the circumstances had been different, in another possible world. Charles Darwin, if he had decided to go into another line of work, might have made an excellent detective around London at the time and fought with some analogue of Moriarty. This is not to say of him, or of anyone else, that he would have been Sherlock Holmes or might have been Sherlock Holmes. . . . [O]ne cannot say which person would have been designated. There is no criterion to pick out one as opposed to another.\(^{21}\) (2011, 59)

Kripke seems to endorse No Modal Criterion for Reference: speaking of Darwin and the other “possible people who might have done the things in the story,” he says, “There is no criterion to pick out one as opposed to another.” He endorses Reject Reference: he says, “one cannot say which person would have been designated.” And he says something that suggests that he endorses Reject Possibilism: speaking of Darwin, he says, “This is not to say of him, or of anyone else, that he would have been Sherlock Holmes or might have been Sherlock Holmes.”\(^{22}\)

Kripke (2011, 59) does not say what a “criterion to pick out one as opposed to another” is. Perhaps such a criterion for a name \(n\) is a non-circular property \(F\) such that it is necessary that, for any object \(o\), if \(o\) has \(F\), then \(n\) refers to \(o\); or perhaps such a criterion for a name \(n\) is a property \(F\) such that it is necessary that, for any object \(o\), if \(o\) has \(F\), then \(n\) refers to \(o\) in virtue of the fact that \(o\) has \(F\). (Either way, the property being an \(x\) such that \(n\) refers to \(x\) is not such a criterion for any name \(n\), either because

\(^{21}\) I have corrected the text by transposing ‘had’ and ‘he’ in the sentence that begins with ‘Charles Darwin.’

\(^{22}\) Kripke (2011, 68) offers a parallel argument, from the claim that there is no modal criterion for truth for the sentence ‘Sherlock Holmes exists’ to the claim that we cannot say, of any possible world \(w\), that Sherlock Holmes exists at \(w\): speaking of the sentence ‘Unicorns exist,’ he says, “it is not sufficient just to be able to say that it is false; one has to be able to say under what circumstances it would have been true, if any. And there seems to be no clear criterion here. . . . I cannot say of a counterfactual situation that it is correctly describable as one in which ‘Sherlock Holmes was fond of cricket,’ ‘Sherlock Holmes was a detective,’ or ‘Sherlock Holmes exists’.”
that property is circular or because it is not the case that \( n \) refers to an object in virtue of the fact that the object has that property.\(^{23}\) Let’s call such a criterion a modal criterion for reference. The rough idea is that, if a property \( F \) is a modal criterion for reference for a name \( n \), then \( n \) refers to any possible object that has \( F \).

4.2 The Brute Reference Fact Reply

One reply to the “no modal criterion for reference” argument is to say that a modal criterion for reference for ‘Sherlock Holmes’ is not required. If there is a possible person \( h \) such that ‘Sherlock Holmes’ refers to \( h \), perhaps it is a brute fact that ‘Sherlock Holmes’ refers to \( h \).\(^{24}\) On this view, the inference from No Modal Criterion for Reference to Reject Reference is invalid.\(^{25}\) Just because there is no modal criterion for reference for ‘Sherlock Holmes,’ that does not mean that, for any possible person, there is something defective about the claim that ‘Sherlock Holmes’ refers to that person. For, if there were a possible person who ‘Sherlock Holmes’ refers to, that person would not need to satisfy any modal criterion to be the referent of ‘Sherlock Holmes’; it would just be a brute fact that ‘Sherlock Holmes’ refers to that person. Let’s call this reply the brute reference fact reply.

Although I am sympathetic to the brute identity fact reply to the “too many candidates” argument, I’m not especially sympathetic to the brute reference fact reply to the “no modal criterion for reference” argument. In the case of identity facts, if some possible person \( h \) is such that Sherlock Holmes = \( h \), then the fact that Sherlock Holmes = \( h \) just is the fact that Sherlock Holmes = Sherlock Holmes. And, I think, it is plausible that that fact is brute. By contrast, in the case of reference facts, if some possible person \( h \) is such that ‘Sherlock Holmes’ refers to \( h \), then the fact that ‘Sherlock Holmes’ refers to \( h \) might be the fact that ‘Sherlock Holmes’ refers to Sherlock Holmes. But, I think, it is less plausible that that fact is brute.\(^{26}\)

4.3 The “Modal Criterion from the Fiction” Strategy

Another reply to the “no modal criterion for reference” argument is to deny No Modal Criterion for Reference. On this view, although a modal

\(^{23}\) On the need for non-circularity in the theory of reference, see Kripke 1972, 68–73.

\(^{24}\) On brute semantic facts, see Breckenridge and Magidor 2012; Kearns and Magidor 2008; 2012. (Kearns and Magidor 2008 is about the extension of predicates; Breckenridge and Magidor 2012 is about the reference of names; Kearns and Magidor 2012 is about both.)

\(^{25}\) Or, if the argument is enthymematic, then missing premise \((i)\) is false. See footnote 20.

\(^{26}\) If we’re comparing the reasoning behind the brutality of identity facts to Kripke’s argument for the necessity of identity (see footnote 7), notice that Kripke does not similarly argue for the necessity of reference. On his view, reference facts are contingent. See, for example, Kripke 1972, 77–78.
criterion for reference for ‘Sherlock Holmes’ might be required, there is such a criterion. There are various ways of denying No Modal Criterion for Reference. For example, perhaps the property being someone who did the things in the story is a modal criterion for reference for ‘Sherlock Holmes.’ Let’s call this way of denying No Modal Criterion for Reference the “modal criterion from the fiction” strategy.

There are two ways to argue against the “modal criterion from the fiction” strategy. One way is to point out that different possible people have being someone who did the things in the story in different possible worlds. So, if ‘Sherlock Holmes’ refers to any possible object that has the property, then it refers to different possible people in different possible worlds, in which case it would not be a rigid designator (roughly, an expression that refers to the same object in every possible world in which that object exists). But ‘Sherlock Holmes’ is a rigid designator, so being someone who did the things in the story is not a modal criterion for reference for ‘Sherlock Holmes.’ This first way of arguing against the “modal criterion from the fiction” strategy assumes that ‘Sherlock Holmes’ is a rigid designator.

Another way to argue against the “modal criterion from the fiction” strategy is to point out that different possible people have being someone who did the things in the story in the same possible world. (Admittedly, this possible world would be a bizarre one, in which a violin-playing detective fights with a mathematics professor in one part of the universe while a duplicate detective fights with a duplicate professor in another part of the universe.) So, if ‘Sherlock Holmes’ refers to any possible object that has the property, then it refers to different possible people in the same possible world, in which case it would not be a singular term (roughly, an expression that refers to at most one object in each possible world). But ‘Sherlock Holmes’ is a singular term, so being someone who did the things in the story is not a modal criterion for reference for ‘Sherlock Holmes.’ This second way of arguing against the “modal criterion from the fiction” strategy does not assume that ‘Sherlock Holmes’ is a rigid designator.

This second way of arguing against the “modal criterion from the fiction” strategy is part of Kripke’s argument for the claim that there is no modal criterion for reference for ‘unicorn.’ Speaking of a story in which there are unicorns and mere lookalikes (which he calls “fool’s unicorns”), Kripke (2011, 66) says,

---

27 On rigid designation, see Kripke 1972.
28 Lycan (1994, 119) considers such a possible world. Unless London could be multi-located, the story would need to be purely qualitative; otherwise, only one of the duplicate detectives would live in London—the city that is identical with the city in the actual world—in that possible world. The story also could not contain superlatives, since only one of the duplicate detectives could be the person who, in that possible world, is the foremost authority on cigar ash.
29 For a possibly congenial remark, see Thomasson 1999, 45. For criticism, see Liebesman 2014, 306.
We are given a partial identification of them [unicorns]; [in the story] there are other criteria that would pick them out from fool’s unicorns, but [outside of the story] we are not told what these criteria are. . . . Thus of no possible animal can we say that it would have been a unicorn. One can merely say that it would look the way unicorns are supposed to. If a possible world contained two very different species, both fully conforming to the aforementioned story, one could not say which of them would have been unicorns.

Here, being a species whose members have such-and-such surface characteristics (e.g., looking like a white horse with a single horn) is not a modal criterion for reference for ‘unicorn,’ not because different possible species have that property in different possible worlds, but rather because different possible species have that property in the same possible world.30

4.4 Another Way of Denying No Modal Criterion for Reference

So far, we’ve been considering the “no modal criterion for reference” argument against possibilism; a reply to the “no modal criterion for reference” argument that denies No Modal Criterion for Reference; the “modal criterion from the fiction” strategy for denying No Modal Criterion for Reference; and two ways of arguing against the “modal criterion from the fiction” strategy.

But the “modal criterion from the fiction” strategy is not the only way to deny No Modal Criterion for Reference. Just because being someone who did the things in the story is not a modal criterion for reference for ‘Sherlock Holmes,’ that does not mean that no property is. And, surprisingly, it seems that, on Kripke’s view, there is another property that is a modal criterion for reference for ‘Sherlock Holmes.’

In Naming and Necessity, Kripke provides what he later describes as “criteria for naming” (2013, 23). He says, “In my own case, . . . the criterion is given by a historical chain in which the reference is transmitted from link to link” (23 n. 23).31 So, for example, something like the property being the object that was baptized at the origin of the historical chain that leads up to our actual use of ‘Aristotle’ is a modal criterion for reference for ‘Aristotle’: ‘Aristotle’ refers to any possible object that has that property.

30 Exercise for the author: figure out whether either of the two ways of arguing against the “modal criterion from the fiction” strategy can be turned into an argument for Multiplicity and Arbitrariness. Conjecture from the author: the first way can; the second way can’t on its own; but the second way can together with the additional premise that uniqueness is not relevant to eligibility. Thanks to David Sanson for help with this assignment.

31 Kripke (1972, 97) suggests that perhaps he has not provided sufficient conditions. But, if so, no name would have a modal criterion for reference, so there wouldn’t be anything especially bad about ‘Sherlock Holmes.’
So why isn’t the property *being the object that was baptized at the origin of the historical chain that leads up to our actual use of ‘Sherlock Holmes’* a modal criterion for reference for ‘Sherlock Holmes’? If it were, No Modal Criterion for Reference would be false. There would be a modal criterion for reference; it’s just that nothing would satisfy it.

There is a difference between the claim that there is a modal criterion for reference for ‘Sherlock Holmes’ *that nothing satisfies* and the claim that *there is no modal criterion for reference* for ‘Sherlock Holmes’. The latter claim is No Modal Criterion for Reference; it is what the “no modal criterion for reference” argument relies on. In particular, the inference to Reject Reference relies on No Modal Criterion for Reference rather than on the claim that there is a modal criterion for reference for ‘Sherlock Holmes’ that no possible person satisfies. If there were a modal criterion for reference for ‘Sherlock Holmes’ that no possible person satisfies, then for any possible person we could assert the negation of the claim that ‘Sherlock Holmes’ refers to that person. That is, for any possible person, we could assert that *it is not the case* that ‘Sherlock Holmes’ refers to that person. By contrast, if there is no modal criterion for reference for ‘Sherlock Holmes,’ then for any possible person there is something defective about the claim that ‘Sherlock Holmes’ refers to that person and we should reject that claim without asserting its negation. What we can assert is the claim that, for any possible person, *we cannot say* that ‘Sherlock Holmes’ refers to that person. This is Reject Reference.

So, if Kripke is in a position to say that there is a modal criterion for reference for ‘Sherlock Holmes’ that nothing satisfies, then he is in a position to deny No Modal Criterion for Reference in a way that undermines the “no modal criterion for reference” argument. This seems puzzling.

### 5 The “No Epistemic Criterion for Reference” Argument

In the previous section, I suggested that Kripke endorses the “no modal criterion for reference” argument in “Vacuous Names and Fictional Entities.” This led to a puzzle, since it seems that Kripke is in a position to deny the premise of that argument: namely, No Modal Criterion for Reference.

To avoid the puzzle, one could deny that Kripke endorses the “no modal criterion for reference” argument in “Vacuous Names and Fictional Entities.” Perhaps what Kripke (2011, 59) has in mind when he speaks of a “criterion to pick out one as opposed to another” is, not a modal criterion for reference for a name *n*, but rather an *epistemic criterion for reference* for a name *n*: namely, a non-circular property *F* such that it is a priori that, for any object *o*, if *o* has *F*, then *n* refers to *o*. In that case, perhaps the

---

32 A student once complained to a professor about getting a failing grade: “This suggests that the work I did was unsatisfactory, when in fact I did no work at all.”

33 Thanks to a referee for suggesting something along these lines.
The Extraordinary Impossibility of Sherlock Holmes

The “no epistemic criterion for reference” argument:

No Epistemic Criterion for Reference: There is no epistemic criterion for reference for ‘Sherlock Holmes.’
Reject Reference: We cannot say, of any possible person o, that ‘Sherlock Holmes’ refers to o. (From No Epistemic Criterion for Reference)
Reject Possibilism: We cannot say, of any possible person o, that Sherlock Holmes is o. (From Reject Reference)

The idea behind the “no epistemic criterion for reference” argument is that there is something epistemically lacking about ‘Sherlock Holmes’; so, for any possible person, there is something defective about the claim that ‘Sherlock Holmes’ refers to that person, and hence there is something defective about the claim that Sherlock Holmes is that person.

Kripke is not obviously in a position to deny the premise of the “no epistemic criterion for reference” argument: namely, No Epistemic Criterion for Reference. First, the property being someone who did the things in the story is not an epistemic criterion for reference for ‘Sherlock Holmes.’ On Kripke’s (2011, 56) view, it is conceivable that (by some “bizarre accident”) someone in the actual world did the things in the story but, because Conan Doyle wasn’t writing about that person, ‘Sherlock Holmes’ does not refer to that person. And, second, even if being the object that was baptized at the origin of the historical chain that leads up to our actual use of ‘Sherlock Holmes’ is a modal criterion for reference for ‘Sherlock Holmes,’ it might not be an epistemic criterion for reference for ‘Sherlock Holmes,’ since it might not be a priori that ‘Sherlock Holmes’ refers to an object that has that property.

But it would be puzzling if Kripke were to be using an epistemological premise—namely, No Epistemic Criterion for Reference—to support a metaphysical conclusion: namely, Reject Possibilism. I take it that Reject Possibilism is the conclusion of the argument that Kripke (1972, 158) endorses in the addendum: namely, the conclusion that, “granted that there is no Sherlock Holmes, one cannot say of any possible person that he would have been Sherlock Holmes, had he existed” (italics in original). Kripke (1972, 158) explicitly describes that conclusion as a “metaphysical view.” He distinguishes it from the epistemological view that the “mere discovery that there was indeed a detective with exploits like those of Sherlock Holmes would not show that Conan Doyle was writing about this man” (157; italics in original). (Similarly, in the case of unicorns, Kripke distinguishes the “metaphysical thesis that no counterfactual situation is

34 See also Kripke 1972, 157–158; 2013, 26–27.
properly describable as one in which there would have been unicorns” from the “epistemological thesis that an archeological discovery that there were animals with all the features attributed to unicorns in the appropriate myth would not in and of itself constitute proof that there were unicorns” [156; italics in original].)

And, in the addendum, the consideration that might be taken to support the claim that being someone who did the things in the story is not an epistemological criterion for reference for ‘Sherlock Holmes’—namely, that it is conceivable that someone in the actual world did the things in the story but ‘Sherlock Holmes’ does not refer to that person—is part of Kripke’s (1972, 157–158) argument for the epistemological view rather than the metaphysical one.

So, to maintain the distinction between Kripke’s arguments for the metaphysical and epistemological views, and to avoid attributing to Kripke an argument that infers a metaphysical conclusion from an epistemological premise, I have taken the argument that Kripke endorses in “Vacuous Names and Fictional Entities” to be the “no modal criterion for reference” argument rather than the “no epistemic criterion for reference” argument.

As a result, further philosophical and exegetical issues remain, among them whether Kripke is in fact in a position to deny No Modal Criterion for Reference and why, if he were in a position to do so, he would endorse the “no modal criterion for reference” argument in “Vacuous Names and Fictional Entities.”

35 One possibility that I am sympathetic to (but lack the space and wherewithal to develop here) is that, although there is a modal criterion for reference for ‘Sherlock Holmes,’ that property does not count as a “criterion to pick out one as opposed to another” (Kripke 2011, 59). For that property to count, ‘Sherlock Holmes’ must be appropriately connected to it. But, because it was introduced in the course of Conan Doyle’s storytelling activities (rather than, say, by someone intending to refer to something), ‘Sherlock Holmes’ is not appropriately connected to that property. It might be that, according to the fiction, ‘Sherlock Holmes’ is appropriately connected to that property; but, outside of the fiction, ‘Sherlock Holmes’ is not.

Thanks to David Sanson and Joshua Spencer for suggesting what I take to be ideas along these lines.

Another possibility that I am sympathetic to (but also lack the space and wherewithal to develop here) is that Kripke ultimately rejects the “no modal criterion for reference” argument in favor of a different argument, one that infers Reject Possibilism from the claim that ‘Sherlock Holmes’ does not refer (rather than from Reject Reference). Let’s call this argument the “no reference” argument. The argument for the claim that ‘Sherlock Holmes’ does not refer relies on the claims (i) that no possible object satisfies being the object that was baptized at the origin of the historical chain that leads up to our actual use of ‘Sherlock Holmes’ and (ii) that having being someone who did the things in the story is not sufficient for being the referent of ‘Sherlock Holmes’ (either because different possible people might have being someone who did the things in the story or perhaps because it is conceivable that someone has that property in the actual world but ‘Sherlock Holmes’ does not refer to that person). (See the comparison of ‘unicorn’ and ‘tiger’ in Kripke 2011, 65–66.) The inference to Reject Possibilism might go via the claim that ‘Sherlock Holmes is a person’ does not express a proposition (and hence does not express a proposition that is possibly true); see the discussion of Kripke 2013, 40, 42 in Section 3 at footnote 17. As I see it, the main question about the “no reference”
modal criterion for reference” argument to see whether it might be the argument that Kripke endorses in the addendum.

6 The “No Modal Criterion for Reference” Interpretation

At the end of the addendum, Kripke (1972, 158) says,

I am aware that the cryptic brevity of these remarks diminishes whatever persuasiveness they may otherwise possess. I expect to elaborate on them elsewhere, in a forthcoming work discussing the problems of existential statements, empty names, and fictional entities.

I take it that the material contained in the forthcoming work that Kripke mentions can be found in “Vacuous Names and Fictional Entities” and Reference and Existence. And, at the end of the paragraph in which the passage quoted near the beginning of Section 4 occurs, Kripke (2011, 59 n. 13) cites the addendum. This suggests that Kripke takes himself to be offering the same argument in both places.
I suggest that, in the addendum, Kripke endorses the “no modal criterion for reference” argument. ⁴³ He endorses Reject Possibilism: he says, “one cannot say of any possible person that he would have been Sherlock Holmes”; and, speaking of several possible people, he says, “there is none of whom we can say that he would have been Holmes” (1972, 158; italics in original). And he presents part of an argument for No Modal Criterion for Reference. In particular, he argues that being someone who performed the exploits of Holmes is not a modal criterion for reference for ‘Sherlock Holmes,’ because “Several distinct people . . . might have performed the exploits of Holmes” (158; italics in original). Let’s call this interpretation the “no modal criterion for reference” interpretation.

On the “no modal criterion for reference” interpretation, the argument in the addendum unfolds as follows.

(P1) “Several distinct people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes” (Kripke 1972, 158).

(C1) So being someone who performed the exploits of Holmes is not a modal criterion for reference for ‘Sherlock Holmes.’ (From [P1]; implicit)

(C2) So No Modal Criterion for Reference is true: there is no modal criterion for reference for ‘Sherlock Holmes.’ (From [C1]; implicit)

(C3) So Reject Reference is true: we cannot say, of any possible person o, that ‘Sherlock Holmes’ refers to o. (From [C2]; implicit)

(C4) So Reject Possibilism is true: “granted that there is no Sherlock Holmes, one cannot say of any possible person that he would have been Sherlock Holmes, had he existed”; that is, “there is none of whom we can say that he would have been Holmes, had he performed these exploits” (Kripke 1972, 158; italics in original). (From [C3])

No Modal Criterion for Reference supports Reject Reference, which in turn supports Reject Possibilism; and (P1) supports NoModal Criterion for Reference. So (P1) ultimately supports Reject Possibilism.

But, in the text, (P1) is separated from Reject Possibilism by ‘but,’ which indicates contrast rather than support. So it would be better if we read ‘but’ as ‘so’ here. And, on the “no modal criterion for reference” interpretation,

---

³⁷ One could also read something like the “no reference” argument (see footnote 35) into the addendum. It is possible that at the time of writing the addendum Kripke did not distinguish the “no modal criterion for reference” argument and the “no reference” argument, that he favored the former in “Vacuous Names and Fictional Entities,” and that he later favored the latter in Reference and Existence. (On the relations among the addendum, “Vacuous Names and Fictional Entities,” and Reference and Existence, see footnote 36.) At the end of the day, the claim I am most committed to is that Kripke does not endorse the “too many candidates” argument in the addendum; this is consistent with his endorsing the “no reference” argument in the addendum, with his endorsing the “no modal criterion for reference” argument there, and with the addendum being indeterminate between those two arguments.
the rhetorical question at the end is misleading at best, since it suggests Arbitrariness, which is not at issue here. So it might be better to pretend the rhetorical question isn’t there. On the “no modal criterion for reference” interpretation, then, the passage from the addendum should be read as follows.

I hold the metaphysical view that, granted that there is no Sherlock Holmes, one cannot say of any possible person that he \textit{would have been} Sherlock Holmes, had he existed [i.e., Reject Possibilism is true]. Several distinct people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes [so \textit{being someone who performed the exploits of Holmes} is not a modal criterion for reference for ‘Sherlock Holmes,’ so No Modal Criterion for Reference is true, so Reject Reference is true], but [so] there is none of whom we can say that he would have \textit{been} Holmes, had he performed these exploits [i.e., Reject Possibilism is true]. For if so, which one? (Kripke 1972, 158; italics in original)

Admittedly, the “no modal criterion for reference” interpretation requires some emendations to the text.\textsuperscript{38} But it explains why Kripke cites the addendum in “Vacuous Names and Fictional Entities” (because he endorses the “no modal criterion for reference” argument in both places), why he doesn’t explicitly state Arbitrariness (because it’s not a premise of the “no modal criterion for reference” argument), and why he rejects possibilism without asserting its negation (because that’s what you’d do if you were asserting the conclusion of the “no modal criterion for reference” argument).

Ben Caplan
E-mail: caplan.16@osu.edu

References:

\textsuperscript{38} There is a precedent for emendation. Kripke himself suggests emending Kripke 1971, 146 so as to replace ‘and’ with ‘or’ and ‘so’ with ‘though.’ (See Kaplan 1989, 570 n. 8.) But in that case Kripke is suggesting that a lecture was “mistranscribed from the tape of the talk,” whereas the addendum was written rather than transcribed. (Again, see Kaplan 1989, 570 n. 8.)

Acknowledgements For comments and discussion, thanks to students at Ohio State and Umeå, to participants at a talk at Umeå, and to David Braun, Scott Brown, Sam Cowling, Eileen Nutting, Romina Padro, David Sanson, Joshua Spencer, Andreas Stokke, Pär Sundström, Amie Thomasson, Chris Tillman, and a referee for this journal.


