

Review

Truth in Fiction: Rethinking its Logic

By John Woods

Cham: Springer International Publishing, 2018. ISBN 978-3-319-72658-8.

Reviewed by Gilbert Plumer

Associate Director for Assessment Projects and Research (retired)
Law School Admission Council
1812 Laurel Oak Dr. N.
Rockledge, FL 32955 USA
plumerge@gmail.com

Abstract: This article reviews John Wood's *Truth in Fiction: Rethinking its Logic* (Springer 2018).

Résumé: Cet article est une critique de *La vérité dans la fiction: repenser sa logique* (Springer 2018) de John Woods.

Keywords: abstract objects, fictional objects, fictionality, formal methods, inference, Law of Noncontradiction, paradox of fiction

John Woods is an international treasure. This book constitutes one reason, among many, to think so. The book is not about argument in fiction or the cognitive value of fiction, or what is commonly called 'truth *through* fiction.' In approximately the past fifteen years, argumentation studies have become increasingly concerned with how works of fiction, taken as wholes, may be argumentative (e.g., Olmos 2017). Since argument can provide the justification needed for knowledge, this question has implications for the cognitive value of fiction. There is nothing, or almost nothing, of this in Woods' book (an exception, reviewed below, might be his discussion of 'the paradox of fiction' in ch. 6: "Salty Tears and Racing Hearts"). Rather, as the book's cover says, it concerns 'truth *in* fiction,' that is, the (often formal) logic and

semantics of claims in fiction, which Woods strikingly regards as “unambiguously true and false together” (e.g., p. xi).

The book is in league with informal logic in that it presents a sustained critique of formal methods in philosophy, especially in logic and philosophy of language. Woods tells us at the beginning that he thinks that “most of what matters for reference, truth and inference in human cognitive practice is not adequately catered for by the standard formal methods, and is better dealt with in a more naturalistic way” (p. 6). A challenge is determining what that naturalistic way is, but certainly the following is a primary case: Woods argues (p. 14 and *passim*) that mathematical logic has “lost sight” of the distinctions between consequence-having, consequence-spotting, and consequence-drawing. Consequence-having is entailment, and it occurs in logical space. Consequence-spotting is an epistemic achievement that occurs in psychological space, although the consequence must be there (be ‘had’) in order to be spotted. Consequence-drawing “occurs in the inferential subspace of psychological space” and cannot be (validly) accomplished unless the consequence is spotted. Spotting and drawing require an agent, unlike having; thus, it is “a costly and revealing mistake” that “in most of the standard approaches to deductive logic, the obtainment conditions on entailment do double duty as rules of valid inference” (p. 14). As an example of such a cost, Woods reminds us that in most standard deductive systems, every proposition has an infinity of consequences. Yet since consequence-spotting and consequence-drawing (inference) require an agent’s belief or belief revision, the infinity of logical consequences cannot be inferred by any agent like us. Any theory that requires that it could or that rational “belief is closed under (deductive) consequence” would make a human being “a doxastic disaster” (p. 15).

In this vein, Woods also discusses, for instance, the shortcomings of canonical notation (esp. p. 210ff., but see Index). Other topics of particular interest to informal logicians include abduction (esp. pp. 44-47, 155-156, 188-189), the Law of Noncontradiction and its Aristotelian roots (esp. sec. 6.4, but again see Index), and the related subjects of dialethic and paraconsistent logics. In addition, Woods argues, partly inspired by Hume, that the vast majority of knowledge and inference is tacit or unreflective—is cognition “down below” (esp. ch. 3). What distinguishes this inference is that “it is *not* advanced by

reasoning,” whereas according to “command and control” epistemology, “inference at its best is delivered by *reasoning*” (p. 52). Woods’ epistemology is supposed to be naturalized both in that it is a variation of causal-response reliabilism and that it “respects empirically discernable facts” (p. 23). This seems to get him into some trouble in considering the paradox of fiction. The paradox of fiction, as it is usually construed, is that although all three of the following propositions seem plausible, they cannot all be true:

- (a) We have genuine emotional responses to certain fictional works.
- (b) We believe that those works are fictional.
- (c) Each of (a) and (b) implies the denial of the other (they are incompatible).

Like most proposed resolutions of the paradox, including my own (2015, sec. 2), Woods rejects proposition (c), but he does so in an eccentric fashion. He postulates two operant causal chains. First, from our knowledge of the real world leading to (b), and the second, from the fact that the fictional story’s “telling is causal.” He takes empirical research he cites to indicate that “our being told something *gets* us to believe it,” though the belief is defeasible (p. 55, cf. 145). The first chain suffers a “causal override,” but not extinction, by the second chain. Hence, (a) and (b) are compatible in our experience: “we can cry for Juliet knowing that nothing really is she” (p. 149).

Something has gone awry. Woods says that through the “compartmentalization” of belief, “we have the two happily harmonized beliefs that Sherlock lived in 1800s London and so too, by golly, did Prime Minister Gladstone.” But it is just false that we believe that Sherlock lived in 1800s London. What is true is that we believe that *in the story*, Sherlock lived in 1800s London. Often, Woods relativizes knowledge and belief in this way; on the same page he says “reading a story would suffice for knowing what’s true in it” (p. 146). Yet relativizing belief in this way dissolves his proposed resolution of the paradox—why would we cry for Juliet if all we believed is that *in the story*, she is in trouble? I think Woods confuses being “believable” with being believed (pp. 143-144). One might make the case (I do) that a story’s

being believable can account for (a) even though (b) is the case, and so, (a) and (b) are actually compatible. Believability with respect to fiction is quite a different thing than it is with respect to nonfiction. If a work of nonfiction is believable, it is *worthy of belief*, but the term cannot mean this with respect to fiction.

As part of his critique of formal methods, Woods asks us to consider the meaning of the sentence “Holmes abided at 221 Baker Street in the 1880s.” The dominant semantics in analytic philosophy says that a sentence’s meaning is given by its truth conditions, that is, the way the world would have to be if it were true. If you know this worldly way, then you know the sentence’s meaning, and vice-versa. Although this theory may work well enough for simple sentences such as the Wittgensteinian favorite ‘the cat is on the mat,’ what about the Holmes sentence? Woods argues that the sentence is true (“in Doyle’s stories”) and Holmes is referred to in virtue of a causal chain in the real world that notably includes Doyle’s “doings,” such as his writing efforts. We all more or less know how the (real) world has to be for there to be this causal chain. Yet, as Woods points out, “this causal path is not what its sentential meaning is, and is not what it says” (p. 211).

This brings us to the heart of Woods’ book, which is his rejection of all but the first of what he calls the “Basic Laws” of received formal semantics (p. 2):

I *The something law*: Everything whatever is something or other.

II *The existence law*: Reference and quantification are existentially loaded.

III *The truth law*: No statement disobliging the something and existence laws can be true.

And there is the “corollary of the prior three” (p. 3):

IV *The fiction law*: Sentences of fiction fail to refer and cannot be true or false. For, if true, they refer to something existent, and, if false, they are false of something, hence of something that exists.

In rejecting II-IV, Woods says that “we part company with Frege...Russell, Tarski, Davidson, Searle, [Kendall] Walton and Kripke” (p. 9).

As indicated, the critical part of Woods’ view that allows sentences such as “Holmes abided at 221 Baker Street in the 1880s” to be true is that there are fictional entities created by writers, and we can refer to those entities. Woods characterizes entities like Holmes (or fictional places like 221 Baker Street in the 1880s) as “impalpable objects,” which he says “are as common as dirt—numbers, universals, idealizations, models, propositions and, yes, even *stories*,” etc. (p. 37). We stand in “real but impalpable relations” (p. 29) to such objects. In knowing that Holmes abided at 221 Baker Street in the 1880s, the knowing is real, but the objects are impalpable. So, given Woods’ causal reliabilism epistemology, causality is a real but impalpable relation to such objects. In writing the Sherlock Holmes stories, the real Doyle brought it about that the unreal Holmes is who he is and did what he did; Doyle is the “truth maker” for the fictional sentences in his books. This sort of causality requires no causal *contact* between Doyle and Holmes or between us and Juliet (or between us and numbers, propositions, etc.). Woods regards causal-contact epistemologies as “discredited” in view of independent counterexamples such as the fact that we can know about “events happening in galaxies outside our light cone” (p. 53)—events that are not contact-causally connectable to us or physically accessible to us.

Unlike many, Woods holds that fictional entities are not possible objects in the sense that the actual world could have been such that it included, e.g., Holmes. Apparently, Woods’ main reason is that Doyle does not have that sort of (divine) truth-making power: “no one can speak an entity into real existence,” so Holmes could not “possibly be, have been, or made to be a resident of our world” (p. 164). But this seems confused. In considering whether the world could have been different than the way it actually is in that Holmes and his escapades were a part of it—could they have existed?—we are not considering whether Doyle or anyone else could have brought that about. It is not a question of causality. Rather, we are considering whether a world in which Holmes and his escapades exist is reasonably accessible from this world (is possible given the basic facts of this world), or whether there is something significantly impossible (not merely, e.g., an occasional

inconsistency in Doyle's texts) about Holmes and his escapades, or the like.

Woods also holds that fictional entities are concrete objects, not abstract objects: Holmes "is a man of the human kind and a human of the fictional kind" (p. 153). Being human, Holmes has hair and a spine, for example, as all humans do; he is no "freak." This leads to a kind of specificity or full realization for fictional objects that at least initially seems implausible. For instance, vagueness aside, Woods maintains that for any time t of Holmes' fictional existence, there is a particular n that numbers the hairs on Holmes' head (or the cells in his body, etc.). How can that be, you might object, Doyle never specified any such thing in the novels, and there is no way to determine n at t for Holmes, unlike there may be for a real human? Yet Woods has several good arguments for his view that fictional objects are fully determinate. One is that this objection mistakes "epistemic indeterminacy for ontic indeterminacy" (p. 80). Just because you may not be able to find out the number of cells in your body, does not mean that there is not that number (again, vagueness aside). Another argument derives from the indisputability of something like Woods' "world-inheritance thesis" (p. 81). A simplified version of this idea is what Friend (2017) calls "the *Reality Assumption*: the assumption that everything that is (really) true is also fictionally the case, unless excluded by the work" (p. 29) (in his review of Woods' book, García-Carpintero 2018 also makes this simplification). The argument is that for any time t of a real human's life, there is (e.g.) a particular n that numbers the hairs on the person's head. Thus, given world-inheritance, the same is true of any fictional human.

A third argument revolves around the "*no spines-no readers thesis*" (p. 82). Essentially, this idea is that if people believed that fictional characters and objects generally were somehow incomplete "freaks," they would not become readers in the first place; but since there are (many) readers of fiction...(QED). However, Woods' argument is broader than this. He wants to "emphasize...relations which readers *experience* themselves bearing to the story's subjects and events...If readers believed that literary texts referred to nothing, that they engendered nothing that's true or believable" as per Basic Laws II-IV, the stories would be "*unengageable*" in our experience, and there would no point in reading or writing them (p. 33; cf., e.g., 17). The structure

of this argument is close to the classical structure of transcendental argument. Transcendental arguments on the order, for example, of Davidson's directed against skepticism about other minds (1991, pp. 159-160), reason that since certain aspects of our experience or inner world are undeniable, the external world must have certain features, on the grounds that its having these features is a necessary condition of our experience being the way it is. Woods is saying that the completeness of fictional objects and the falsity of Basic Laws II-IV (etc.) is a condition of the possibility of our experience of reading fictional works being the way that it is, i.e., engaged. This is a little puzzling to me because Woods sometimes (p. 189) appears to think that this sort of argument of his is a Peircian abduction ("The surprising fact *C* is observed. But if *A* were true, *C* would be a matter of course. Hence there is reason to suspect that *A* is true."—e.g., p. 155). In any case, García-Carpintero's (2018) concern that "Woods's naturalism consists merely in an appeal to his own intuitions" is overstated, given such argument patterns in the book and points that I mentioned earlier about Woods' naturalism, not to mention Woods' citation of a significant amount of psychological research (esp. sec. 7.4) and documentation of real world inconsistency management (esp. sec. 9.3), not that I think empirical research is very often dispositive in philosophy or logic.

Although, as Woods says, "ontologists have free rein to raise questions" about the status of impalpable objects and how we can have any interaction with them "until the heat death of the universe" (p. 37), the preceding should give a flavor of his view. In some ways, arguments may be like Woods' fictional objects. For instance, Simard Smith and Moldovan (2011) conclude that, though they are abstract objects, arguments are *author-created*, temporal objects.

There is one last big issue to consider. It is the foundational view of the book noted above that claims in fiction are "unambiguously true and false together." According to Woods, this applies to "every sentence of a full story that owes its truth, even in slender part, to its author's sayso," thereby constituting "fiction's *systemic* inconsistency," as opposed to merely localized inconsistency between parts of a text (p. 128). He adds that almost always this systemic inconsistency "has been taken as lights-out for the sentences of fiction, and correspondingly a vindication of its Basic Laws" (p. 127). Woods' solution to the problem is to insist that although the sentences in question all have the

form of ‘p and not p’ (cf. p. 127) and so are inconsistent, *they are not contradictory*. Get it? I don’t. Woods explains (p. 131): “That they are true and false together makes them inconsistent. Since they aren’t true and false in all the same respects spares them the contradictoriness that flows from the fullest formulation of LNC” (the Law of Noncontradiction). I still don’t get it.

Aristotle has several formulations of LNC. Woods cites (p. 129) what he terms “the ontological formulation.” I prefer this translation: “The same thing cannot at the same time both belong and not belong to the same object and in the same respect” (*Metaphysics* 1005b 19-20, trans. Apostle 1966, pp. 58-59). Aristotle’s idea of a respect here is used by Woods to develop his notion of a “site” (pp. 86-87, 130), which we have already seen hint of. One kind of site is a fictional story; another kind is the real world. So the example “Holmes abided at 221 Baker Street in the 1880s” is true in Doyle’s stories yet false in the real world. The only other cases of sites that Woods gives are the law and “theoretical models.” Something may be true in the real world, for instance, that the defendant is guilty of murder, yet false in the law prior to the jury’s verdict, and afterward for that matter if the jury finds the defendant innocent. Or vice versa. Theoretical models “can be sites for the true idealizations of empirically false reality.” Interestingly, Woods maintains that sites are not possible worlds. His principal reason seems to be that possible worlds are “free for the asking” in that they exist independently of whatever we do, whereas “it takes some real doing to bring about a site” (p. 87). Sites are the creations of writers, societies, theoreticians, and presumably for the real world, the divine (cf. p. 164).

But it remains a mystery why Woods continues to insist that sentences of fiction that owe their truth to the “author’s sayso” are inconsistent, after showing how they are not contradictory. Surely, he is right that the “siteness” of their truth and falsity is not ambiguity or multiple meaning—they do not mean different things in the story and in the real world (pp. 87, 126, 148). No problem. Sites appear to be comparable in some ways to contextual parameters of evaluation, about which there is a large literature (see, e.g., my 1993). For instance, consider time and place of utterance for indexical sentences, e.g., ‘it is raining.’ Just because this sentence is true at some indices and false at others does not mean that it is inconsistent or that it is semantically ambiguous. It seems misleading for Woods to continue to maintain

inconsistency. At one point he characterizes fictional stories as being “inconsistent with the world” (p. 77). Yet if that is all the inconsistency amounts to, it appears trivial; after all, the stories are fictional. It does not mean that the sentences involved have the form of ‘p and not p’ any more than ‘the defendant is guilty’ or ‘it is raining’ have that form.

The book covers many other topics, mostly in the philosophy of literature and mostly expertly. In a few places, Woods’ view comes off as a little naïve. For example, he says “readers of fictional texts respond to them as reports knowing full well that they report nothing that’s actually happened” (p. 138; cf., e.g., 77). This is ahistorical in the respect that “recent scholarship has shown” that the “modern *concept* of narrative fiction developed slowly in early-modern Europe” (Gallagher 2006, pp. 337-338). In 1720, Defoe was avowing that Robinson Crusoe was a real person, and in 1742 Fielding had to argue that his characters, such as Tom Jones, were imaginary. Yet by later in the century, “the public had been trained to read novels as stories about thoroughly imaginary” people (p. 344). And one does wonder what Woods would think about the creative ways that the concept of fiction is thrown around these days. My favorite: “I hesitate to say that it is *inconceivable* that a work of fiction could be entirely true” (Friend 2012, p. 190).

As indicated, Woods calls his (first) solution to what he regards as fiction’s inconsistency problem, the “no-contradiction” solution. He has an alternative solution should it turn out that the inconsistency after all is contradiction, which he calls the “no-bother option” (e.g., p. 173). His main worry about the no-contradiction approach appears to be that his concept of sites is theoretically shallow and may turn out not to be viable (e.g., pp. 130, 188). The no-bother solution (ch. 9) says that consumers of fiction recognize its systemic inconsistency but are not bothered by it; it is just one more area among many areas in their beliefs where inconsistency has to be managed through various measures. The most interesting measure consists of “irrelevance filters” that compensate for the fact that even in non-formal English, anything logically follows from a contradiction (pp. 181-182). Woods is not entirely happy with this solution either, because although it is “more empirically assured,” it also could use theoretical development (pp. 188-189). But again, once truth is relativized to a contextual parameter, it is not clear there is an inconsistency problem in fiction to be solved. The real

work is developing a plausible account of the nature of fictional objects and how we can refer to them, thereby allowing sentences of fiction to be true (in the story), and Woods does that job admirably.

References

- Apostle, Hippocrates G., trans. 1966. *Aristotle's Metaphysics*. Bloomington: Indiana University Press.
- Davidson, Donald. 1991. Three varieties of knowledge. In *A. J. Ayer: Memorial Essays*, ed. A. Phillips Griffiths, 153-166. Cambridge: Cambridge University Press.
- Friend, Stacy. 2012. Fiction as a genre. *Proceedings of the Aristotelian Society* 112(2): 179-209.
- Friend, Stacy. 2017. The real foundation of fictional worlds. *Australasian Journal of Philosophy* 95(1): 29-42.
- Gallagher, Catherine. 2006. The rise of fictionality. In *The Novel, Volume 1*, ed. Franco Moretti, 336-363. Princeton, NJ: Princeton University Press.
- García-Carpintero, Manuel. 2018. Review of John Woods, *Truth in Fiction: Rethinking its Logic*. Notre Dame Philosophical Reviews. URL accessed 31 December 2019: <<https://ndpr.nd.edu/news/truth-in-fiction-rethinking-its-logic/>>.
- Olmos, Paula, ed. 2017. *Narration as Argument*. Cham: Springer.
- Plumer, Gilbert. 1993. A here-now theory of indexicality. *Journal of Philosophical Research* 18: 193-211.
- Plumer, Gilbert. 2015. A defense of taking some novels as arguments. In *Proceedings of the eighth Conference of the International Society for the Study of Argumentation*, eds. Bart Garssen et al., 1169-1177. Amsterdam: SicSat. URL accessed 31 December 2019: <<http://rozenbergquarterly.com/issa-proceedings-2014-a-defense-of-taking-some-novels-as-arguments/>>
- Simard Smith, Paul L. and Andrei Moldovan. 2011. Arguments as abstract objects. *Informal Logic* 31(3): 230-261.