

Truth in Fiction and Relating Logic

Daniela Glavaničová¹ and Matteo Pascucci²

¹ Comenius University in Bratislava, Bratislava, Slovakia
daniela.glavanicova@gmail.com

² Slovak Academy of Sciences, Bratislava, Slovakia
matteopascucci.academia@gmail.com

Truth in fiction is one of the ongoing debates in analytic philosophy that presents intriguing challenges and puzzles for logic. To begin with, the step from characterizing the explicit content of a work of fiction to characterizing its implicit content is not straightforward. Authors of fictions can also unintentionally contradict themselves, forgetting about what they wrote earlier. Perhaps the best known example of this is Watson’s inconsistently located war wound in the stories about Sherlock Holmes. Moreover, the story evolves: what is true at some point of the storyline need not be true at a different point of the storyline, such as Elizabeth Bennet’s earlier opinions on Mr. Darcy. Also, there is the problem of unreliable narrators, such as Humbert Humbert in Nabokov’s *Lolita*. In addition, some genres, such as time-travel stories, are specifically challenging when it comes to their logical assumptions. For instance, in the series *Dark*, a single character is a mother and a daughter of another character. And of course, authors can intentionally contradict themselves in their works of fiction, as in *Sylvan’s Box* by Priest [4] (in the story, a box is both empty and has something in it).

The seminal work in this area was written by David Lewis [3]. Lewis proposed three analyses of which propositions are true in a work of fiction. This account has been influential till the present day; see, for instance, Woodward [6], Badura and Berto [1], and García-Carpintero [2].

Arguably, truth in fiction requires a fine-grained analysis. As Badura and Berto [1] argue, there are some fictions that are best interpreted as diverging from reality; in fact, this divergence can venture as far as the impossible. They mention the logical impossibility of *Sylvan’s Box* as one example, and a metaphysical impossibility of a fictional character becoming a real human as another. Moreover, in our opinion, if one wants to track the distinction between the explicit and the implicit content of a work of fiction in a rigorous way, then one can acknowledge that works of fiction typically do not contain equivalences of classical logic (or of other logics). Therefore, the fact that a proposition ϕ is part of the explicit content of a fiction does not entail that any proposition ψ which is classically equivalent to ϕ is also part of the content of that fiction. Relating logic allows for representing pertinence to a work of fiction as well as pertinence to a proposition in a work of fiction, and allows for a fine-grained analysis. For this reason, we will employ a fine-grained framework of relating semantics.

In our formal framework relating semantics will be combined with a bimodal language of tense logic [5]. The idea behind this is that we want to treat a work of fiction f as a story that is presented in a certain temporal order, namely the order of narration (which may not correspond to the order of the events it concerns, since it may include flashbacks, time travel, different timelines, etc.). For instance, a written novel can be identified with a sequence of pages (or sections) of a book. Each of these can be identified, in turn, with a set of sentences (or propositions) and can serve as a parameter for evaluating the truth of an arbitrary sentence within the novel. Moreover, what is mentioned at some point in the novel remains as a trace in later parts. Our formalism captures the perspective of a reader of a novel f , who, after having gone through the whole book, compares what is stated at some point (via the tense-logical operator N_f) with what is stated in earlier parts (via the tense-logical operator H_f) or later parts (via the tense-logical operator G_f).

We keep track of the distinction between the explicit and the implicit content of a fictional work by making use of two modal operators. First, we define an operator Exp_f which reads “it is part of the explicit content of fiction f that” in a purely tense-logical way, as follows:

$$\text{Exp}_f\phi =_{def} H_f\phi \vee N_f\phi \vee G_f\phi$$

Second, we take our formal language to include a relating operator \sim_f such that a formula like $\phi \sim_f \psi$ reads “according to fiction f , ϕ is related to ψ ”. This choice allows us to define an operator Imp_f which reads “it is part of the implicit content of fiction f that” in terms of \sim_f and Exp_f , as follows:

$$\text{Imp}_f\phi =_{def} \text{Exp}_f\psi \wedge (\phi \sim_f \psi)$$

Thus, the notion of relatedness plays a central role in our approach. In the final part of our presentation, we discuss various ways in which relatedness can be understood in this context and show the formal consequence of each of the proposed interpretations.

References

- [1] C. Badura and F. Berto. Truth in fiction, impossible worlds, and belief revision. *Australasian Journal of Philosophy*, 97:178–193, 2019.
- [2] M. García-Carpintero. ‘Truth in Fiction’ reprised. *British Journal of Aesthetics*, 62:307–324, 2022.
- [3] D. Lewis. Truth in fiction. *American Philosophical Quarterly*, 15:37–46, 1978.
- [4] G. Priest. Sylvan’s box: A short story and ten morals. *Notre Dame Journal of Formal Logic*, 38:573–582, 1997.
- [5] A. Prior. *Past, Present and Future*. Clarendon Press, 1967.
- [6] R. Woodward. Truth in fiction. *Philosophy Compass*, 6:158–167, 2011.