Doing the impossible with possible worlds

Louis Rouillé
Institut Jean Nicod
louis.rouille@ens.fr

September 6, 2019

1 The problem of fictional "truth"

1.1 The problem

Explain the following contrast:

(1) Hamlet is a human being. ("true" in *Hamlet*)

(2) Hamlet is a crocodile. ("false" in *Hamlet*)

1.2 The counterfactual analysis of fictional statements

Ref: [Lewis 1978]

Thesis: Fictional statements are disguised counterfactuals.

Csq: The contrast between (1) and (2) is explained by the following constrast:

(3) If Shakespeare's *Hamlet* were told as known fact rather than fiction, Hamlet would be a human being.

(4) If Shakespeare's *Hamlet* were told as known fact rather than fiction, Hamlet would be a crocodile.

Reminder: Lewis's general truth-conditions for counterfactual statements:

- $\phi > \psi$ is true at w iff all the ϕ-worlds most similar to w are also ψ-worlds.

1.3 Problems for possible-world accounts of fictional "truth"

Ref: [Currie 1990], p. 54:

> Possible worlds are *determinate* with respect to truth; [...] they are *consistent* [...] but fictional worlds are always indeterminate and sometimes inconsistent.

Ref: [Walton 1990], p. 64:

> [...] fictional worlds are not possible worlds. Two differences, especially, have been discussed elsewhere: Fictional worlds are sometimes impossible and usually incomplete, whereas possible worlds (as normally construed) are necessarily both possible and complete.

1.3.1 Incompleteness

(5) Hamlet has blood is his veins. ("true" in *Hamlet*)

(6) Hamlet is of blood group O. (*neither* "true" *nor* "false" in *Hamlet*)

(7) Hamlet is not of blood group O. (*neither* "true" *nor* "false" in *Hamlet*)

Completeness in PWS: In a Kripke model, $K =< W, R, v >$, for all proposition p:

(Completeness) $\forall w \in W, v_w(p) = 1 \vee v_w(\neg p) = 1$

1.3.2 Inconsistency

(8) Watson has a war wound. ("true" in *The Adventures of Sherlock Holmes*)

(9) Watson has a wound on his shoulder. (*both* "true" *and* "false" in *The Adventures of Sherlock Holmes*)

(10) Watson has no wound on his shoulder. (*both* "true" *and* "false" in *The Adventures of Sherlock Holmes*)

Consistency in PWS: In a Kripke model, $K =< W, R, v >$, for all proposition p:

(Consistency) $\forall w \in W, v_w(p) = 1 \leftrightarrow v_w(\neg p) = 0$

1.3.3 Dual problems

One can see that (Completeness) and (Consistency) are formally equivalent, given:

(Classical neg) $v(p) = 1 - v(\neg p)$

Proof:

(Consistency)

$\equiv \forall w \in W,\ (v_w(p) = 1 \wedge v_w(\neg p) = 0) \vee (v_w(p) \neq 1 \wedge v_w(\neg p) \neq 0)$

Given (Classical neg):

$\equiv \forall w \in W,\ v_w(p) = 1 \vee v_w(p) = 0$

$\equiv \forall w \in W,\ v_w(p) = 1 \vee v_w(\neg p) = 1$

Which is (Completeness).

2 In search for duality

2.1 Inelegance of non-dual solutions

2.1.1 Lewis's solution

Incompleteness: Technically, there is no problem for Lewis, for the solution is hardwired in his semantics.

Inconsistency: It requires special fixes, presented in [Lewis 1983], called the "method of union and intersection":

> Perhaps we should take the maximal consistent fragments, obtained by deleting the bare minimum that will give us consistency. [...] what do we do with our several consistent fragments (or corrections) when we have them? [...]
>
> I suggested this *method of intersection*: ϕ is true in the original fiction iff ϕ is true in every fragment. Now I would favor instead this *method of union*: ϕ is true in the original fiction iff ϕ is true in some fragment. (Not that we need to choose once and for all – we can have both methods, distinguishing two senses of truth in inconsistent fiction.)[1]

But as [Phillips 1999] puts it:

> A better account would be a unified theory which supplied truth-conditions for every fiction, rather than singling out inconsistent fictions for special treatment.[2]

2.1.2 Non-normal world solution

Ref: [Berto 2017]

Incompleteness: Introduce incomplete worlds, e.g. where bivalance fails.

Inconsistency: Introduce impossible worlds, e.g. a paraconsistent world. But as [Hanley 2004] puts it:

> [...] the shift to the method of union actually does shed some light on the nature of fctional truth, by showing the relationship between indeterminacy and inconsistency. Truth in fiction is not truth-at-a-world, but rather truth at a *set* of worlds.[3]

[1] p. 276.
[2] p. 281.
[3] p. 117.

2.2 Dual solutions

2.2.1 Definitions

Starting point: A fiction (is tantamount to a long counterfactual which) selects a *set* of possible worlds.

$F = \{w|@R_f w\}$

- with @ designating the actual world.
- f the fiction and R_f the accessibility relation obtained by processing "were f told as known fact rather than fiction".
- F is the set of all the *complete specifications* of f.

Deletion procedure:

- Extract a contradiction, which is of the form $p \wedge \neg p$.
- Run one specification of the story with p.
- Run another specification of the story with $\neg p$.
- Store p in D, the set of deleted propositions.
- Apply this procedure for each contradiction.

2.2.2 Supervaluations

Let us enrich our Kripke models with a supervaluation function s^+: $K^+ =< W, R, v, s^+ >$. One can define "supertruth" over a set of possible worlds F in the following manner:

- $s^+_F(p) = 1$ iff $\forall w \in F$, $v_w(p) = 1$
- $s^+_F(p) = 0$ iff $\forall w \in F$, $v_w(p) = 0$
- p is *indeterminate* otherwise.

Csq:

$s^+_{F_{Hamlet}}(6)$ is *indeterminate*.

$s^+_{F_{Hamlet}}(7)$ is *indeterminate*.

2.2.3 Subvaluations

Let us enrich our Kripke model with a subvaluation function s^-: $K^- =< W, R, v, s^- >$ One can define "subtruth" over a set of possible worlds F in the following manner:

- $s^-_F(p) = 1$ iff $\exists w \in F$, $v_w(p) = 1$
- $s^-_F(p) = 0$ iff $\exists w \in F$, $v_w(p) = 0$

Csq:

$s^-_{F_{Sherlock}}(9) = 1$

$s^-_{F_{Sherlock}}(10) = 1$

But: $s^-_{F_{Sherlock}}(9 \wedge 10) = 0$

2.2.4 Combining the two

One can define fictional "truth" in general:

$$FT(p) = \begin{cases} s^+_F(p), & \text{if } p \notin D \\ s^-_F(p), & \text{otherwise.} \end{cases}$$

3 The revenge of inconsistent fictions

Distinction: *accidentally* vs *essentially* inconsistent fictions.

3.1 Sylvan's box

3.1.1 The inconsistency

Ref: [Priest 1997]

> Looking in the box was somthing like that: the experience was one of occupied emptiness. [...] The box was really empty and occupied at the same time. The sense of touch confirmed this.[4]

The reading test:

> [...] let us take an old-fashioned comprehension test on the story.
>
> *Question 1*: In which country did the meeting take place? *Answer*: Australia.
>
> *Question 2*: Was Richard at the farmhouse? *Answer*: No.
>
> *Question 3*: Was the box empty? *Answer*: Yes and no.
>
> *Question 4*: How many times did Nick leave the property? *Answer*: Once.
>
> *Question 5*: Was the box shot off to the moon at the end of the story? *Answer*: No.
>
> Other answers are wrong, or in the case of Question 3, at least incomplete.[5]

- Interestingly, *FT* gets all the answers right.

[4] p. 575.
[5] p. 579.

Priest's fourth moral:

> Nor can we simply break the information up into (maybe maximally) consistent chunks and infer from each of these. If we could, we would have to infer that the character were astonished by the fact that the box was empty and/or by the fact that it had a figurine in it, which they most certainly were not. The logic emplyed is not, therefore, a nonadjunctive paraconsistent logic.[6]

- The characters were astonished (supertrue, hence fictionally true) because the box was full *and* empty (subtrue, hence fictionally true).
- The characters were astonished because the box was full (subtrue, but not fictionally true).
- The characters were astonished because the box was empty (subtrue, but not fictionally true).

3.1.2 Impossible fictions as double-bind phenomena

Def: a proposition p is fictionally "true" iff there is a work of fiction which prescribes to imagine that p.

Ref: [Walton 1990], [Walton 2013], [Woodward 2014] (p. 829-830).
As for impossibilities, as remarked in [Walton 1990]:

> Can one imagine impossibilities? Not, presumably, if imaginability is a good test for possibility. But then can contradictory or metaphysically impossible propositions be fictional, on our account? I am inclined to think that even contradictions can be imagined in the relevant sense. But our understanding of fictionality is safe even if they cannot be. There can be prescriptions to imagine a contradiction even if doing so is not possible. (A badly drafted law might require one to do something and also to refrain from doing it.) There may also be separate prescriptions to imagine p and to imagine not-p, without a prescription to imagine their conjunction. The set of propositions fictional in a given world might be inconsistent even if no contradiction is fictional.[7]

3.2 Sylvan's *fempty* box

Ref: [Badura and Berto 2019]
What if Sylvan's box was full-and-empty as a single requirement?

- Let us call this being *fempty*.
- Fempty is a *simple*, *contradictory* predicate.

[6] p. 580.
[7] pp. 65-6.

  - Fempty(b) cannot be *deleted*.

- Is *Sylvan's box* requiring to imagine that the box is fempty?
  - If yes, then it is even stranger than it first seems, for it actually introduces a new predicate "fempty" which means the same as "both full and empty" without being a complex predicate.
    * Now it looks like Nelson Goodman's "grue" and "bleen", considered as simple predicates...

As [Lewis 1983] puts it:

> [...] but where we have an inconsistent fiction, there also we have several consistent fictions that may be extracted from it. (Perhaps not in the very hardest cases – but I think those cases are meant to defy our efforts to figure out what's true in the story.)[8]

3.3 A diagonal argument against unrestricted impossible worlds

- Give me your logic, and I will produce a fiction which violates this logic:

[Routley 1979]:

> Given that the logic of a fictional world may be any logic, it follows that *there is no general unifom logic of fiction*. For the intersections of all logics is a null logic, no logic, as each purported logical principle is cancelled out by a logic where it does not hold good. Consider, to illustrate, one of the more promising principles for a logic of fiction, formed by introducing a fictional functor 0 (Woods'*olim* operator) read, say, "it holds in fiction that", namely the principle $O(A \wedge B) \rightarrow OA$. Spelled out semantically the principle has it that if $A \wedge B$ holds in the world of an arbitrary work N then so does A. But consider now a novel where the principles of connexive logic govern, and where hence $A \wedge \neg A$ may hold though A does not. The world of such a novel repudiates $O(A \wedge B) \rightarrow OA$. In claiming that there is no uniform logic of fiction, it is not implied that fiction has no logic, far less that it is illogical. In general, *each work will have its own internal logic*: it is simply that the emerging set of *common* logical principle will be zero. The semantical structure will reflect this situation.[9]

[Proudfoot 2006]:

[8] p. 276.
[9] pp. 10-1.

> Both the possible worlds semanticist and the impossible worlds semanticist are caught by the diversity of fiction. Just as there are more fictions than the possible, so the class of "impossible fictions" includes more cases than the impossible worlds semanticist can deal with. An author can make just about anything the case in a fiction, and this is one reason why it is simply wrong-headed to attempt to analyse truth in fiction in terms of such an orderly apparatus as the worlds semantics. The things that Lewis Carroll makes true in *Alice in Wonderland* defeat a possible worlds semantics. By shifting to an impossible worlds semantics one can perhaps cope with *Alice in Wonderland*, but other fictions will defeat this modification. As a final striking and simple example, consider a logician's fiction in which both A and its negation as defined by the familiar truth table are true (and true only). Even impossible worlds semanticists agree that there are no worlds in which A and its truth-table negation are true simpliciter (for by the truth-table if A is true then its negation is false). [Footnote: Proponents of impossible worlds deal in weaker-than-classical negations, e.g. the so-called De Morgan negation.][10]

[10] p. 31.

References

Badura, Christopher and Francesco Berto (2019). "Truth in Fiction, Impossible Worlds, and Belief Revision". In: *Australasian Journal of Philosophy* 97.1, pp. 178–193.

Berto, Francesco (2017). "Impossible worlds and the logic of imagination". In: *Erkenntnis* 82.6, pp. 1277–1297.

Currie, Gregory (1990). *The nature of fiction*. Cambridge University Press.

Hanley, Richard (2004). "As good as it gets: Lewis on truth in fiction". In: *Australasian Journal of Philosophy* 82.1, pp. 112–128.

Lewis, David (1978). "Truth in fiction". In: *American philosophical quarterly* 15.1, pp. 37–46.

— (1983). "Truth in fiction with postscripts". In: *Philosophical papers, vol. 1*. New York: Oxford University Press, pp. 261–280.

Phillips, John (1999). "Truth and inference in fiction". In: *Philosophical Studies* 94.3, pp. 273–293.

Priest, Graham (1997). "Sylvan's Box: A Short Story and Ten Morals". In: *Notre Dame Journal of Formal Logic* 38.4, pp. 573–582.

Proudfoot, Diane (2006). "Possible worlds semantics and fiction". In: *Journal of Philosophical Logic* 35.1, pp. 9–40.

Routley, Richard (1979). "The semantical structure of fictional discourse". In: *Poetics* 8.1-2, pp. 3–30.

Walton, Kendall (1990). *Mimesis as Make-believe: On the Foundations of the Representational Arts*. On the Foundations of the Representational Arts. Harvard University Press.

— (2013). "Fictionality and imagination reconsidered". In: *From Fictionalism to Realism*. edited by C. Barbero, M. Ferraris, and A. Voltolini. Cambridge Scholars Publishing.

Woodward, Richard (2014). "Walton on Fictionality". In: *Philosophy Compass* 9.12, pp. 825–836.