

A Granularity Problem for Impossible World Semantics

Synopsis:

The thesis of the paper is that there is a subclass of necessarily true [and necessarily false] propositions whose members are distinct yet impossible world semantics won't be able to distinguish between these members. Before giving the argument, some background is covered explaining how impossible world semantics was developed from possible world semantics to be better able to individuate necessarily true [and necessarily false] propositions, without a recourse to structured intensions. For possible world semantics (without structured intensions), two propositions are identical *iff* they are true in all the same possible worlds. Infamously, this would make every necessarily true proposition identical and every necessarily false proposition identical. For impossible world semantics, two propositions are identical *iff* they are true in all the same worlds, possible and impossible. Since there are impossible worlds where necessary truths are false, impossible world semantics can distinguish most necessarily true propositions. And since there are impossible worlds where necessary falsehoods are true, impossible world semantics can distinguish most necessarily false propositions.

However, I argue that impossible world semantics won't be able to distinguish between sentences, roughly [more details in the paper], of the type "if P were true and only true, then P". The fact that the consequent is repeated in the antecedent guarantees the truths of these propositions at all worlds, possible and impossible; so long as the truth value of the counterfactual is related to the truth value of its antecedent and consequent. This isn't solely sufficient for the problem to arise but the "only true" caveat in the sentences expressing these propositions guarantees that these propositions aren't false in any world; the details of why are spelt out in the paper. These two factors are sufficient for every proposition of this type to be considered identical by impossible world semantics.

This is problematic since sentences such as, "If it was true and only true that Obama was president, then Obama is president" and, "If it was true and only true that there are four mugs in the world, then there are four mugs in the world" seem to be expressing different propositions but are of the "problematic" type referenced above.

In the final section, some objections and replies are considered. First, the objection that the problematic propositions are true and only true while also false at some impossible world is considered. The motivation behind this objection is that descriptions of impossible worlds are often inconsistent which means that the paradoxical situation of something being "true and only true" and also being "false" doesn't really prevent the situation from accurately describing an impossible world. In reply, I argue that if impossible world semantics were to utilize this avenue the theory would end up entailing that propositions were distinct from themselves, which is implausible. Next, the objection considered is that the truth-values of counterfactuals are independent of the truth-value of their antecedents and consequents. This would allow the problematic propositions to not be true at some impossible worlds. I claim that this is a very high price to pay and that it doesn't work well with another type of proposition which I discuss.

-----start of paper-----

1. Background

Possible world semantics is a popular theory in which semantic contents are identified with a function from circumstances of evaluation (construed as ‘possible worlds’) to reference. For declarative sentences, the semantic content, sometimes called a “proposition”, is a function which, given any possible world as an input, will output either “true” or “false”. The function outputs ‘true’ if given a possible world in which the sentence, according to our use, would be true in that world and false otherwise. For example, the semantic content of “Jane is tall” is the function that maps all and only the possible worlds in which Jane is tall to “true”.

But possible world semantics has an infamous problem. Since the theory says that the semantic content just is a function from possible worlds to truth values, two sentences express the same semantic content *iff* the content expressed by both are true in all and only the same possible worlds. As such the theory, as-is, entails that all necessarily true sentences have the same semantic content, due to their having the same truth value in all possible worlds (since necessarily true sentences are true in all possible worlds). This unappealing entailment is known as “the granularity problem” (Barwise 1977, Berto 2013). The problem is severe and infamous enough that several lines of response have been attempted.

One-way philosophers have tried to solve this problem is to bring in the notion of “structured intensions” in which the sub-sentential units, which also express functions, can be used to individuate the semantic contents of sentences. I mention this view only to set it aside; it will **not** be evaluated in this paper. Instead I will show how a different ‘solution’ philosophers have proposed to the granularity problem isn’t in and of itself sufficient.

The other best-known solution to the granularity problem is to extend possible world semantics to include impossible worlds (see Yagisawa 1988, Barwise 1997, Berto 2013, Nolan

2013, and Jago 2015). This extended theory, sometimes called “Impossible World Semantics” [hence “IWS”], supposedly solves the granularity problem because the mathematical relations¹ associated with any two non-synonymous necessarily true statements give a different output at least once when given the same impossible world as input. This is because there are impossible worlds where some (though usually not all) necessary truths are false. For example, there is an impossible world, call it *W*, in which triangles don’t have three sides but two plus two still equals four. As such, the proposition expressed by the sentence “two plus two equals four” outputs “true” when given *W* as an input whereas the proposition expressed by the sentence “triangles have three sides” outputs “false” when given *W* as an input. Since there are some inputs for which these mathematical relations give different outputs they are not identical relations, and, as such, IWS is able to distinguish the semantic contents expressed by “two plus two equals four” and “triangles have three sides” even though this are standardly considered necessarily true statements.

2. The Purposed Granularity Problem For IWS

However, in this paper I will argue, even IWS has the undesirable consequence of entailing identity between distinct semantic contents, at least if it tries to keep the truth value of counterfactuals related to the truth value of its antecedent and consequent (something any semantic theory should want to do).

The problematic propositions that I will be talking about are very much like propositions expressed by sentences of the form “if *P* [or superset of *P*], then *P*” but with a slight twist. As such it is informative to first look at how an IWS theorist could deal with sentences of the type “if *P* [or

¹ I use the more general term “mathematical relation” as opposed to “function” due to possible multiple outputs per input.

superset of P], then P” and then see how their responses won’t work for the problematic propositions I will bring up. Take the following two counterfactuals:

1. If it is raining, then it is raining
2. If ice cream is cold, then ice cream is cold

The theorist endorsing IWS has two options. First, they could deny that the truth-values of counterfactuals are related to the truth values of its antecedent and consequent, but this is immensely unappealing and comes at a huge cost. These drawbacks are enough for us to leave this line of response alone for now but it will be revisited in “Objection 2”. Second, and more likely than not, the IWS theorist distinguishes the propositions expressed by (1) and (2) by claiming that there is some world, call it W^* , such that:

- i. Some world, W^{**} , is the relevant distance away from W^* such that (1)’s truth value at W^* depends upon the truth value of ‘it is raining’ at W^{**}
- ii. Some world, W^{***} , is the relevant distance away from W^* such that (2)’s truth value at W^* depends upon the truth value of ‘ice cream is cold’ at W^{***}
- iii. W^{**} is an impossible world where the proposition ‘it is raining’ is both true and false at W^{**}
- iv. W^{***} assigns true or false, but not both true and false, to the proposition ‘ice cream is cold’.

Given these conditions, (1) will map W^* to both 'true' and 'false'² whereas (2) will map W^* only to 'true'. Since, for some input, the propositions expressed by (1) and (2) give different outputs, IWS is able to individuate them.

Having discussed this, I can now discuss the propositions that will be problematic for IWS. They will be the propositions that are expressed by sentences constructed through following formula:

If every proposition in Γ [where ' Γ ' refers to a consistent set of propositions] and all the propositions classically entailed by Γ were true and only true and the only propositions that were true then, ψ [where ' ψ ' refers to a proposition]

Sentences constructed in this way will express propositions that are either true and only true in every world, possible and impossible (this will happen if ψ is a member of Γ), or false and only false in every world, possible and impossible (this will happen if ψ is not a member of Γ) [*henceforth I will focus on those sentences that are always true*]. Yet, as will be shown by examples, there do seem to be sentences created through the above formula that express non-identical semantic contents. For example, the following sentences are constructed using the above formula:

- A. If every proposition in {all dogs are blue, only two dogs exist, two plus two equals four} and all the propositions classically entailed by this set were true and only true and the only propositions that were true, then all dogs are blue.

² This will depend upon the theory of counterfactuals being used, but this a typical evaluation and the most favorable evaluation for IWS since it best allows IWS to distinguish between tokens of this type of a necessarily true counterfactual.

- B. If every proposition in {water is H₂O, Barack Obama exists, there are five grains of sand} and all the propositions classically entailed by this set were true and only true and the only propositions that were true, then Barack Obama exists.

These sentences express different propositions. The proposition expressed by our use of A [hence ‘PA’] is about dogs and math whereas the proposition expressed by our use of B [hence ‘PB’] is about grains of sand, the molecular composition of water, and the existence of Barack Obama. Yet IWS will, unappealing, entail that these sentences express the same proposition/have the same semantic content because the propositions expressed by these sentences are true in every world.

The reason these sentences express propositions that are true and only true in all possible worlds is due to two factors. First, the fact that the consequent is also featured in the antecedent of the counterfactual guarantees that this counterfactual will be true at all worlds.³ Second, these sentences also say that the propositions referred to in the antecedent and consequent are “only true”. This guarantees that these counterfactuals are only true at all worlds. It does this by way of limiting the worlds that satisfy the antecedent of the counterfactual; if a proposition featured in the antecedent and the consequent of these counterfactuals were true and false at any world, such a world would not satisfy the antecedent. As such, the truth value of the counterfactual will never depend upon worlds where the proposition in the counterfactual is both true and false. This prevents the IWS theorist from completing a similar move to the one made earlier when we were discussing “if P [or superset of P] then P”.

³ This alone, as seen by the preceding paragraphs, does *not* guarantee that the counterfactual is *only true* at all worlds. The clever maneuver to distinguish propositions expressed by sentences of the type “if P [or superset of P] then P” concedes that these sentences are true at all worlds but claims that these propositions are true and false at different worlds, which allows IWS to individuate them.

To make the issue more concrete, take PA, for example. For a world to satisfy the antecedent of PA, no proposition outside of what is classically entailed by the set of propositions {all dogs are blue, only two dogs exist, two plus two equals four} can be true and all propositions in that set or classically entailed by the set are true and **only** true. Because of the ‘only true’ requirement in the antecedent no world where it is false (be this only false or true and false) that all dogs are blue, will satisfy the antecedent. But for it not to be the case that PA is true and only true at every world there would need to be a world that satisfies the antecedent of PA yet where it is false (be this only false or true and false) that all dogs are blue. As such, according to this analysis of counterfactuals, PA is true and only true at all worlds, possible and impossible. And parallel reasoning entails that all propositions expressed by sentences created using the formula above, including PB, are true and only true at all worlds possible and impossible. Since IWS entails that propositions are identical whenever they share the same truth value in all worlds, possible and impossible, IWS entails that PA and PB, and all the propositions expressed by sentences following the above formula, are identical. But this seems blatantly false. After all, some of these propositions will be about kangaroos whereas others will not, some will be about the Game of Thrones TV show whereas others will not, some will be about Obama whereas others will not, and so on.⁴

A summarized version of the argument presented above is distilled below in premise-conclusion format for easier viewing of the reasoning. While two propositions were picked from the problematic subclass of necessarily true propositions, it is important to remember there are many such propositions (this is not represented in the below formulation of my argument).

⁴ For those wondering what the “and the only propositions that are true” clause is adding; it is more relevant when we are discussing propositions that are false in all worlds, possible and impossible, but it does also add a little buffer against some odd possible counters that could be made.

- 1) If IWS was correct then propositions are identical *iff* they share the same truth value(s) in all worlds, possible and impossible.

Let 'W' designate an arbitrary world, possible or impossible

- 2) Counterfactuals are true at W *iff* either:
 - i. the consequent obtains at the nearest (or some other specified distance⁵) world, possible or impossible, from W where the antecedent obtains, or
 - ii. There is no world where the antecedent obtains.
- 3) In any world where the antecedent of PA obtains it is true and only true that all dogs are blue.
- 4) In any world where the antecedent of PB obtains it is true and only true that Obama exists.
- 5) By [2] and [3] PA is true and only true at W
- 6) By [2] and [4] PB is true and only true at W
- 7) Since 'W' designated an arbitrary world, possible or impossible, from [5] it follows that PA is true and only true at all worlds, possible or impossible.
- 8) Since 'W' designated an arbitrary world, possible or impossible, from [6] it follows that PB is true and only true at all worlds, possible or impossible.
- 9) By [1], [7], and [8]; if IWS was correct then PA and PB are identical
- 10) But PA and PB are about different things and so PA and PB cannot be identical.
- 11) By [9] and [10], IWS is not correct, at least as-is.

3. Objections and Replies

⁵ The details of this don't matter for my argument because what I am arguing holds for all worlds.

Objection 1: There is an impossible world W where all of the proposition in the set {all dogs are blue, only two dogs exist, two plus two equals four} are true and only true but the proposition expressed by “all dogs are blue” is also false. This is, after all, an impossible world and so descriptions of it may be paradoxical and inconsistent yet still be accurate. The proposition expressed by “all dogs are blue” may very well be only true at this world and also be false at this world, since, again, W is an impossible world and so descriptions of W may be inconsistent. In this world, PA would be both true and false whereas PB may be only true and thus IWS does not entail the identity of PA and PB.

This avenue of response faces a dilemma, the first horn of which is shown by the line of reasoning below.

- 1) If IWS was correct, then propositions are identical *iff* they have the same truth value in all worlds possible and impossible.
- 2) PA maps W only to the true. [stipulation in response *need so that W satisfies the antecedent of PA*]
- 3) By [1] and [2], if IWS was correct, then PA is different than any propositions which maps W to false.
- 4) PA also maps W to the false. [stipulation in response]
- 5) By [3] and [4] if IWS was correct, then PA it is distinct from itself.
- 6) No proposition is distinct from itself
- 7) By [5] and [6], IWS is not correct as-is

The second horn of the dilemma, which would need to be taken to avoid the first, is that the imaginary objector didn't really mean 'only' when they said that the proposition expressed by “all dogs are blue” maps W to true and *only* true. Taking the second horn of this dilemma abandons

this avenue of response because then *W* doesn't satisfy the antecedent of *PA*, but taking the first horn leaves the theory with the absurdity of individuating propositions so finely that they have become distinct from themselves. I take the second horn to be vastly preferable and thus I take this avenue of response to be sufficiently closed.

Objection 2: Reject the standard way to evaluate counterfactuals. The truth values of compound propositions (or at least counterfactual propositions) aren't related to the truth values of their constituents. The truth value of a counterfactual isn't related to the truth of its antecedent or its consequent, at least in impossible worlds.

This line of response may fall under what Nolan (forthcoming) calls the "American Plan". Indeed, this route has been discussed and developed to an extent that what I am about to say may not be completely decisive, though I believe my concerns to be at least compelling. First, Bjerring and Schwarz (2017) have a wonderful article discussing the immense costs of giving up the fruitful and overwhelmingly plausible notion that the truth values of compound sentences (such as counterfactuals) are what they are standardly considered. Giving up the predominate, at least for world based semantics, account of counterfactuals is a significant loss, especially *for* a world based semantics. An analysis of counterfactuals in terms of possible worlds has historically been one of the strongest virtues of world based semantics and so to abandon this is to give up a lot of ground. This standard account about the truth conditions of counterfactuals has been fruitful, strongly advocated, and enjoys a greater degree of plausibility than the alternative.

Further, I also worry about sentences such as the following:

A* In a world where, all of the propositions in {all dogs are blue, only two dogs exist, two plus two equals four} and all the propositions classically entailed by this set were true and only true and the only propositions that were true, then all dogs are blue.

B* In a world where, all of the propositions in {water is H₂O, Barack Obama exists, there are five grains of sand} and all the propositions classically entailed by this set were true and only true and the only propositions that were true, then Barack Obama exists.

After all, it seems, according to any world based semantics, A* is true *iff* all dogs are blue in a world where every proposition in {all dogs are blue, only two dogs exist, two plus two equals four} and every proposition classically entailed by this set were true and only true and the only propositions that were true. But if that is the case, then the problem is still alive. This problematic subclass of sentences is like the one previously discussed except that instead of being of the form “if... were true and only true then...” they are of the form “In a world where ... were true and only true, ...”.

4. Conclusion

It has been argued that the theory I have called “Impossible World Semantics” still individuates propositions too coarsely despite the fact that this theory was born from Possible World Semantics in order to avoid this problem. This runs counter to what various philosophers have suggested (e.g., Yagisawa 1988, Barwise 1997, Berto 2013, Nolan 2013, and Jago 2015). After explaining the class of propositions that Impossible World Semantics would have trouble with and arguing that Impossible World Semantics does indeed have trouble with them, various objections were considered and replies to these objections were provided. I take this paper to show that a minimal constraint for Impossible World Semantics to individuate propositions plausible is that it must either drop the notion that the truth value of a counterfactual is related to the truth values of their constituents, which comes at a huge cost, or it must incorporate structured intensions, in which case further argumentation would need to be done in order to show that

impossible worlds are actual doing theoretic work in the hybrid theory. This is just a minimal constraint; there may still be other bars that would need to be passed for IWS to individuate propositions plausibly in all cases.

References

- Barwise, J., 1997, "Information and Impossibilities", *Notre Dame Journal of Formal Logic*, 38: 488–515.
- Berto, F., 2013, "Impossible Worlds", *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2013/entries/impossible-worlds/>
- Bjerring, J.C., and Schwarz, W., 2017 "Granularity Problems", *Philosophy Quarterly*, 67 (266): 22-37
- Jago, M., 2015, "Impossible Worlds", *Noûs*, 49 (4): 713-728
- Nolan, D., 2013, "Impossible Worlds", *Philosophy Compass*, 8: 360–72.
- , forthcoming, "Impossibility and Impossible Worlds" in Otavio Bueno & Scott Shalkowski (eds.), *The Routledge Handbook of Modality*. New York, USA: Routledge Press.
- Yagisawa, T., 1988, "Beyond Possible Worlds", *Philosophical Studies*, 53: 175–204.