Counterfactual Donkeys Don't Get High (but Not Because They're Weak)

Mike Deigan
michael.deigan@yale.edu
(Questions/comments appreciated!)

May 1, 2017

1 Background

(1) If Balaam owns a donkey, he beats it.

On a salient reading, (1) iff $\forall x(D(x) \wedge BO(x) \rightarrow BB(x))$. We can infer from (1) that if Herbert is a donkey Balaam owns, Balaam beats Herbert.

(2) If Balaam owned a donkey, he would beat it.

It seems that we can infer that if Herbert were a donkey Balaam owned, Balaam would beat him.

How best to predict this?

van Rooij (2006), Walker and Romero (2015): give a semantics which allows it to be valid through use of an assignment-sensitive similarity ordering.

Alternative (perhaps implicit in Wang (2009)): valid only relative to contexts with special kind of similarity relation, leave similarity ordering assignment-insensitive.

1.1 High/Low, Weak/Strong

Some terminology:

High: $\forall x((D(x) \wedge BO(x) \;\Box\!\!\rightarrow \ldots)$
Low: $(\exists x D(x) \wedge BO(x)) \;\Box\!\!\rightarrow \ldots)$

Strong: $\ldots \;\Box\!\!\rightarrow \forall x(D(x) \wedge BO(x) \rightarrow BB(x))$
Weak: $\ldots \;\Box\!\!\rightarrow \exists x(D(x) \wedge BO(x) \wedge BB(x))$

(3) a. If I had a dime, I'd put it in the meter.
b. If John owned a donkey, he'd own Platero (since Platero is the cheapest and John's favorite).

vR and WR think the salient readings of (3-a) and (3-b) are low+weak, whereas the salient reading of (2) is high+strong.

Wang in effect thinks there are no high readings.

1.2 WR against Wang

Main claim: there are contexts in which CDSs license high-reading-like inferences even though the special similarity relation Wang would require for this is implausible.

(4) Scenario: Balaam took part in a game show which had the following format. If you win the easy first round, you win Herbert, an obnoxious and disobedient donkey. The reward for the much more difficult second and third rounds are the well-mannered and obedient donkeys Eeyore and Platero, respectively. Losing a round of the game eliminates the player, keeping them from advancing to any later rounds. Balaam was eliminated in the first round, and so remains donkeyless. John, only aware of the game's first round, asserts (2), since he knows about Balaam's short temper. Sarah corrects him with (5).

(5) No, Balaam could have won Platero or Eeyore too, and he wouldn't beat either of them.

It is implausible, WR would contend, to claim that in this context a world where Balaam advances to and wins the third round is just as similar to the actual world as the one where he wins just the first round. But this is what Wang would have to say to give John's utterance a false reading which (5) can be used to disagree with.

1.3 Where we're going

Primary aims: (1) show that WR's argument in favor of a high reading fails and (2) offer independent reason to doubt the existence of high readings.

Secondary aim: clarify what a high reading would involve.

2 Why we don't *need* the high reading

Basic idea: even if we grant there are high readings of the kind vR and WR's accounts predict, we'll still need the kind of unintuitive ordering Wang needs.

"Note that in van Rooij's system, low and weak readings always go together, a feature we will inherit from him. However, it is an open question whether this is empirically the case: It might well be the case that low+strong readings exist. *This issue does not arise for the high reading, as here a weak and a strong reading are indistinguishable anyway*." (Walker and Romero, 2015, p. 295, n. 3, my emphasis).

I think this is wrong. Weak and strong can be distinguished even in high readings. Moreover, the high reading vR and WR get is weak. To fit our judgments about the following kind of case, vR and WR will have to appeal to the very same unintuitive similarity ordering that Wang must appeal to.

(6) SCENARIO: Cory, who is donkeyless, is a bit crazy. He's disposed to take out his anger on his most prized possession when he's angry. He also took part in the game show described in (4), but also lost in the first round. Had he won any rounds, the prize from the most advanced round he won would have become his prized possession, and he would have beaten it, but he wouldn't beat anything else.

Now consider the following.

(7) If Cory owned a donkey, he would beat it.

The salient reading of this is a strong one, and seems false in this scenario. But using the kind of intuitive similarity ordering advocated by WR for cases like (4), it's true on the high reading that vR and WR would give it (it's also true on their low reading). As far as I can tell, the only way for the kind of theory advocated by vR and WR to predict (7) to be false is by appeal to the very same kind of similarity ordering Wang has to appeal to for (4).

But if we are going to have this kind of ordering available anyways, there's no need to posit the high reading to begin with.

3 So what's with this similarity ordering?

Very tentative proposal: for underspecified antecedents, salient alternative ways of making the antecedent true are all equally good, as far as similarity is concerned. The normal contributors to dissimilarity (like those from Lewis (1979)) only concern aspects of the world not directly involved in making the antecedent true.[1]

But this won't work in general.

(8) a. If Balaam had won any rounds, he would have won just the first one. So if he had owned a donkey, he would have only owned Herbert. And if he had owned Herbert, he would have beat him. So if he owed a donkey, he would beat it.
b. If Cory had owned a donkey, he would have owned only Herbert and he would beat him. So if he owned a donkey, he would beat it.

So we also need an ordering which *is* sensitive to 'similarity' differences having to do with how the antecedent comes true.

I wish I had something systematic to say about why these orderings arise when they do. I'm also somewhat doubtful that a very shifty similarity theory is the way to go here, but no matter how we end up doing things, it seems that in some contexts the counterfactual will care about such differences, in others they won't.

This is a bit like backtracking.

(9) a. If he jumped, he would have died.
b. He wouldn't have jumped unless there was a net. So if he had jumped there'd be a net there to save him and he wouldn't have died.

4 Why we don't want a high reading

There may be reasons independent of WR's argument to think there's a high reading. E.g., alternative-sensitive theories of counterfactuals (Alonso-Ovalle, Ciardelli) + alternativey theory of indefinites.

But the high reading becomes less plausible when we stop limiting ourselves to constant domains.

[1] Regarding truthmaking I have in mind Fine's recent stuff on truthmaker semantics. 'Salient' is here a bit of a black box that should be further theorized.

Allie and Bert think Mary the potter probably didn't make anything yesterday.

(10) *Allie*: If Mary had made a vase, she would have made it from glass.

Presumably (10) does not require for non-trivial truth that there are actually existing things which, if Mary had made them and they were vases, they would have been glass. So we should interpret $\forall$ of the high reading as including the merely possible in its domain. But then the high reading requires there to be no possible thing which is such that if Mary made it and it was a vase, it would not have been made from glass. This seems implausibly strong.

Suppose the conversation contintues.

(11) a. *Bert*: But she could have made it from clay!
b. *Allie*: Yes, I know she could have, but I still think that if she had made a vase, she would have made it from glass.

Now suppose it turns out that Mary in fact *did* make some vases yesterday.

Case 1: Mary made two vases, both of glass.

In this case it seems that on any reading (11-b) is true. But it is not true of all possible vases, if Mary had made them, she would have made them from glass—if she had made one of the clay vases she could have made, she would have made it from clay, not glass.

Note also that there is a reading of (11-b) which is true in Case 1 but not true in

Case 2: Mary made one vase of glass and one of clay.

Thus there is a distinction between strong-low and weak-low readings of CDSs. And we can't make sense of the Case 1 judgment by treating it as weak reading.

5 Where does this leave us?

Given strong centering, Wang's theory handles the vase cases nicely (see below), but as far as I can tell doesn't predict weak readings. I'm also not sure the test semantics is the best way to go, or indeed a minimal change approach to world-selection. There are likely other issues that will arise once we consider the analogues of other phenomena noticed in the indicative donkey sentence literature. But one thing about the account I think we should try to preserve: no high readings.

5.1 Wang gets the vase case right

w - world
g - assignment (partial function from variables to (possible) individuals)
i - possibility, $\langle w, g\rangle$
I - set of possibilities
s - state, set of possibilities with same domain for g
i subsists$_p$ in s' iff $\exists i' = \langle w, g\rangle \in s' : w = w' \wedge g \subseteq g'$
s subsists$_s$ in s' iff $\forall i \in s, i$ subsists$_p$ in s'
$I' = \{i \in I : i$ is a ϕ−possibility with assignment $g\}$
$/\phi/_g = \bigcup_{i \in I'}\{i\}[\phi]$
$<_w$ - similarity relation over worlds relative to w
$f(/\phi/_g, \langle w, g\rangle) = \{\langle w', g'\rangle \in /\phi/_g : \forall \langle w'', g''\rangle \in /\phi/_g, w' <_w w''\}$
$s[\phi]^c = \bigcup_{\langle w,g\rangle \in s} f(/\phi/_g, \langle w, g\rangle)$
$s \vDash \phi$ iff $s[\phi]$ exists and s subsists$_s$ in $s[\phi]$
$s[\phi > \psi] = s, ifs[\phi]^c \vDash \psi$, otherwise $\emptyset$.
$s[\exists x\phi] = \bigcup_{d \in D}(s[x/d][\phi])$

(11-b): $\exists xMMV(x) > G(x)$

Vase case 1:
$@ : MMV(a), MMV(b), G(a), G(b)$
$s[/\exists xMMV(x)/_g] = \{\langle @, g[x/a]\rangle, \langle @, g[x/b]\rangle\}$
$s[/\exists xMMV(x)/_g]G(x) = \{\langle @, g[x/a]\rangle, \langle @, g[x/b]\rangle\}$
so $s[/\exists xMMV(x)/_g] \vDash G(x)$
so $s[\exists xMMV(x) > G(x)] = s$

Vase case 2:
$@ : MMV(a), MMV(b), G(a), \neg G(b)$
$s[/\exists xMMV(x)/_g] = \{\langle @, g[x/a]\rangle, \langle @, g[x/b]\rangle\}$
$s[/\exists xMMV(x)/_g]G(x) = \{\langle @, g[x/a]\rangle\}$
so $s[/\exists xMMV(x)/_g] \nvDash G(x)$, since $\langle @, g[x/b]\rangle$ does not subsist
so $s[\exists xMMV(x) > G(x)] = \emptyset$