Merricks: Truth and freedom

Remark 1. Merricks' notation regarding propositions is: *that X*. I render it as: $[\![X]\!]$.

TRUTH AND BEING. Let's agree, says Merricks, that truth depends on the world. Thus:

(12-1)   The proposition $[\![$THERE ARE NO WHITE RAVENS$]\!]$ is true *because* there are no white ravens.

Or if you don't believe in propositions, then you should still agree that:

(12-2)   The sentence 'There are no white ravens ' (or its assertion) is true *because* there are no white ravens.

This is a platitude, but it creates a challenge for fatalism, so far as the fatalist wishes to say explicitly that truth *determines* being. Perhaps he reasons this way:

   (i) $[\![$I RAISE MY HAND TOMORROW$]\!]$ is true (or: true already today).
  (ii) So I will raise my hand tomorrow.
 (iii) But if it's the case that I will raise my hand tomorrow, then I can't choose freely whether to do it or not.
  (iv) So I can't choose freely—that or any other act.

The characteristic feature of this argument is the inference from (i) to (ii) that rests exactly on the dubious claim that truth determines being. Curiously, Merricks never states the correct principle of dependence. Perhaps, then, we are compelled to accept one of the following:

'*Truism*'. The truth of a proposition *P* depends on (is determined by): how things are with regard to *P* (?? the case that *P*?? the world described by *P*??)

Let's comment on this a bit...

   In any event, not every fatalist argument is threatened by this direction of dependence. The fatalist may construe an argument that doesn't invoke the dependence claim.

Remark 2. Note the plainly invalid fatalist argument mentioned by Merricks. Some of them are also in Vihvelin[10].

THE MAIN ARGUMENT. The fatalist argument that avoids the bad truth-being dependence instead appeals to realism about the future (see Vihvelin[10], 308). But, as Merricks notes, we have to decide whether propositions are true *simpliciter* or true-at-*t*. Suppose they are true *simpliciter*, timelessly. Then we could have the following argument:

   (i) I have free will only if I have control over a certain part $\mathfrak{F}$ of the future.
  (ii) I have control over $\mathfrak{F}$ only if $\mathfrak{F}$ is not 'fixed '.
 (iii) $\mathfrak{F}$ is not 'fixed ' iff the relevant proposition has no truth value.
  (iv) $[\![$I RAISE MY HAND TOMORROW$]\!]$ is true. (RF)
   (v) Therefore, the proposition $[\![$I RAISE MY HAND TOMORROW$]\!]$ has a truth value.
  (vi) Therefore, I have no control over raising my hand tomorrow. [from (ii), (iii), (v)]
 (vii) Since $[\![$I RAISE MY HAND TOMORROW$]\!]$ was selected arbitrarily (i.e. since RF holds generally), the same holds for every other proposition.
(viii) So I have no free will.

This argument is weak in the transition from (v) to (vi). 'Why,' you say, 'surely by raising my hand I can change the truth value? What's so mysterious about changing truth values if I can change the world?' Your grounds seem good generally, but especially so if we accept the idea that truth depends on the world, not *vice versa*. Of course, it's no good to reply that you can't *change* the world—that would beg the question.

   In order to make the transition more convincing I should say something along these lines:

(v-1) The proposition $[\![$I RAISE MY HAND TOMORROW$]\!]$ has an *unchangeable* truth value.
(v-2) I have control over future ϕ-ing only if $[\![$I AM ϕ-ING IN THE FUTURE$]\!]$ has a changeable truth value.
  (vi) Therefore, I have no control over (=can't change) raising my hand tomorrow.

But this involves the fatalist in a possibly fatal dispute over 'unchangeably '.

Remark 3. Compare van Inwagen's discussion of unchangeable truth values.

Instead, Merricks grants that propositions are *not* timeless, and that they can change their truth values in the course of time. Thus the fixity of the past is expressed in his premiss (1) by a claim about a truth a thousand years ago. We can revise our argument above as follows:

(v-M1) The proposition ⟦I ʀᴀɪsᴇ ᴍʏ ʜᴀɴᴅ ᴛᴏᴍᴏʀʀᴏᴡ⟧ was true a thousand years ago. [implicit in Merricks' premiss (1)]

(v-M2) I have control over future ϕ-ing only if I have control over the truth value of ⟦I ᴀᴍ ϕ-ɪɴɢ ɪɴ ᴛʜᴇ ꜰᴜᴛᴜʀᴇ⟧.

(v-M3) But the truth value of ⟦I ʀᴀɪsᴇ ᴍʏ ʜᴀɴᴅ ᴛᴏᴍᴏʀʀᴏᴡ⟧ was fixed in the past.

(v-M4) Therefore, I have no control over the truth value of ⟦I ʀᴀɪsᴇ ᴍʏ ʜᴀɴᴅ ᴛᴏᴍᴏʀʀᴏᴡ⟧.

(vi) Therefore, I have no control over (=can't change) raising my hand tomorrow.

Conjoined with the premisses above, this is a more compelling reconstruction of the ideas driving Merricks' Main Argument. But how does its original form fare? As it stands, it's not logically valid. To be plausibly valid, you must accept these inference rules (where '$\boxdot$' stands for 'Having no choice over'):

(12-3)   a. $\Box A \Rightarrow \boxdot A$.

b. $[\boxdot A, \boxdot(A \rightarrow B)] \Rightarrow \boxdot B$.

Then we have:

(i) $\boxdot P$ [Merricks' premiss (1)]

(ii) $\Box(P \rightarrow Q)$ [Merricks' premiss (2)]

(iii) Therefore, $\boxdot(P \rightarrow Q)$ [by (12-3a)]

(iv) Therefore, $\boxdot Q$ [by (12-3b): Merricks' conclusion (3)].

The problem here is the rule (12-3b). I really don't know what to say about it—unless you treat $\ulcorner \boxdot A \urcorner$ as synonymous with $\ulcorner \Box A \urcorner$. That's what we implicitly assumed earlier in the discussion of Taylor's fatalism. Then it would be a familiar modal rule.

Alternatively, Merricks' premiss (2) is not about necessity at all. It compresses two different claims. First, there is a principle:

(2-1) If I have no choice over A, then I have no choice over B that is logically entailed by A.

The claim (2-1) *looks* plausible to me. Perhaps we may try to substantiate it by a *reductio ad absurdum*. Secondly, there is this:

(2-2) [The proposition A is true at t] $\Rightarrow A$.

With (2-1) and (2-2) replacing (2), the Main Argument becomes valid. However, I worry that Merricks shouldn't accept (2-2), or indeed (2-1) despite its plausibility. As we saw, he holds that propositions are not timeless, and that they change their truth values. So how to understand *any* entailment $\ulcorner A \Rightarrow B \urcorner$? Should we, in testing the entailment, look at the values of A and B at the same time t, or at all times?

Question 4. Examine the logical structure of the Main Argument in the light of the above remarks.

YSB