Epistemology of Logic First: Against Warren’s Defense of Charity in Logic

Jared Warren has argued that charitable interpretation demands that, roughly, epistemically rational people express only truths in using sentences of logic. He defends this thesis as the best thesis in the metasemantics of logic. I argue that Warren’s defense is mistaken. The view that epistemically rational people express only truths with logical sentences requires not a thesis in metasemantics but a substantive and likely implausible thesis in epistemology. My argument has two upshots for Warren’s larger project of defending implicit conventionalism in logic. The first is that, contrary to what he claims, he has not addressed Quine’s old objections to implicit conventionalism. The second is that, without an epistemology of logic, the naturalistic credibility his project is open to question.

Section 1. Charity

Putting things broadly to cover different views of charitable interpretation, our interpretation of another speaker is charitable if we interpret her utterances so that any beliefs we attribute to her on the basis of our interpretation come out rational. (By ‘rational’ I mean epistemically rational). Of course, sometimes people say or believe irrational things, either because they do not have their wits about them or because they have made a mistake they could easily correct. Aside from such cases, I will grant in this paper that our interpretations of speakers should usually be charitable.

Here is an example. Suppose Anna accepts as true ‘It’s not raining’ as she and I stand in a torrential downpour in New York City. The average rational person would not believe that it is not raining in this circumstance. Hence, to interpret her charitably, I should accept (R):

(R) I should reject any interpretation of Anna’s sentence on which she says that it is not raining.

Otherwise, assuming Anna believes what she says, she has an irrational belief.

Suppose by querying Anna I determine that in her idiolect ‘rain’ means snow. So, my interpretation of her sentence is: it’s not snowing. Assuming she says what she believes, she believes that it is not snowing. In the circumstances, her belief is true. But note that only (R) is needed to interpret Anna charitably given how I described the circumstance: to only attribute rational beliefs to Anna in the circumstance, I simply need to avoid interpreting her as saying that it is not raining. I was not done interpreting Anna in the circumstances by making this supposition, but I was done meeting the constraint of charitable interpretation. Without further detail, inferring additionally that ‘rain’ means snow need not appeal to charity.

In fact, we can modify the circumstances so that our charitable interpretation of her sentence comes out false. Suppose the downpour mentioned above was of acid rain. It is easy to imagine Anna having the rational yet false belief that the rain was not acidic due to lack of knowledge of the relevant meteorological facts. We can also imagine more details added which imply that, in her idiolect, ‘rain’ is short for ‘acid rain’. Hence, she says and believes the falsehood that this is not acid rain. A similar point can be made about rejection of sentences. If Anna had rejected ‘It is raining’, we need not interpret ‘It is raining’ as expressing a falsehood if ‘rain’ means ‘acid rain’.

This establishes My Moral:
My Moral: One can charitably interpret a speaker without further interpreting the sentences she accepts as expressing something true or the sentences she rejects as expressing something false.

Charity *often* demands attribution of truths to accepted sentences and falsehoods to rejected sentences, since rational speakers often get things right. That is why charitable interpretation in the literature tends to be associated with attribution of truth to most utterances, likely due to Davidson (1973). But total attribution of truth does not follow – I am charitable to Anna in attributing to her a false belief that this is not acid rain.

**Section 2. Warren’s Defense of Charity in Logic**

Now let us turn to Jared Warren’s work. In “Change of Logic, Change of Meaning” (Warren 2018), Warren defends **CLCM**:

**CLCM:** for any logical sentence \( \phi \), if A and B disagree in their attitudes toward \( \phi \), then A and B mean something different by \( \phi \) provided that (i) neither A nor B are disposed to defer to a common source in adjudicating logical disputes and (ii) neither A nor B is making any relevant computational error in the disagreement (p. 424)

Loosely put: a change of logic entails a change of meaning. He says that “logical sentences” of a given language \( L \) are “[u]sually…the putative logical truths/falsehoods in a language.” (p. 423) So, in a language with classical logic, ‘Either it is raining or it is not raining’ is a logical sentence likely interpreted as true; ‘It is both raining and not raining’ is also a logical sentence likely interpreted as false.

In brief, Warren defends **CLCM** as follows. He first defends **Save the Logical Attitudes**:

**Save the Logical Attitudes:** When translating language \( L \) of language community \( C \) into English, we should reject any translation \( t \) that doesn’t preserve (actual or potential) clearheaded attitudes of acceptance, rejection, and neither acceptance nor rejection for logical sentences. (Warren 2018, p. 429)

Warren then defends **Save the Logical Attitudes** by using and defending **Logical Charity**:

**Logical Charity:** For any language \( L \), for any logical sentence \( \phi \): \( \phi \) is true [false] in \( L \) if and only if \( \phi \) is (potentially) clearheadedly accepted [rejected] by \( L \) speakers.

My focus in this paper is not on **CLCM** but Warren’s defense of **Logical Charity**. Before getting to that defense, first note that My Moral implies **Logical Charity** is not a consequence of charitable translation. Fix a speaker \( S \) of a fixed language \( L \). Suppose \( S \) is rational in uttering a logical sentence \( \phi \) of \( L – S \), as Warren puts it, “clearheadedly accepts” or “rejects” \( \phi \). My Moral shows that charitable interpretation of \( S \) does not imply that \( \phi \) is true if accepted or false if rejected. So **Logical Charity** does not follow from the nature of charitable translation. Additional support for **Logical Charity** is necessary. Warren should recognize this, since elsewhere he characterizes charitable interpretation similarly to how I have: “if our proposed translation makes those we are interpreting unaccountably wrong and shockingly irrational then our proposed translation should be rejected” (2015b, p. 8). He there correctly suggests that characterizations of charity stronger than mine and his which imply general attribution of truth are too strong.
So how does Warren defend Logical Charity? He first argues that a “central objection” to Logical Charity fails: the charity trap. According to the charity trap, if we accept Logical Charity, we must accept false views in the metasemantics (a discipline offering an account of the nature of truth, reference, and meaning of sentences) of non-logical areas of discourse. For example, we must accept a view he attributes to Thomas Kuhn called Scientific Charity: “clearheadedly accepting some scientific sentence (e.g., “The earth does not move”) suffices for the truth of that sentence (and similarly for clearheaded rejection and falsity).” (p. 433) Warren thinks that externalist metasemantic theories on which causal connections between scientific terms fix their reference show that Scientific Charity is false: since Copernicus and his opponents both causally interacted with the earth in using their languages’ equivalents of ‘the earth’, the equivalent of “The earth does not move” as uttered by Copernicus’ opponents is false, not true. Warren says that “nothing analogous is plausible” in the case of logic. Hence, our acceptance of Logical Charity does not commit us to endorsing Scientific Charity, since we can offer a different metasemantics for scientific discourse.

Warren also considers a few other metasemantical accounts and argues against extending them to logic. He concludes that “[t]hose who reject Logical Charity face the burden of providing a metasemantically plausible explanation of putative logical disagreement” and that “my discussion makes it clear that Logical Charity neatly coheres with the most plausible contemporary approaches to metasemantics, and does so in a way that avoids the charity trap.” (p. 436)

Warren here is conceding that a bit more needs to be done to defend Logical Charity. But the incompleteness of his argument is not what bothers me. It is the region of philosophy in which Warren thinks a complete argument for Logical Charity takes place. What is clear is that Warren thinks that an adequate argument for or against Logical Charity is offered within metasemantics - we must seek the best explanation in the metasemantics of logic in accepting or rejecting it. To reject it, the “burden” is on us to come up with a different account of how the truth, meaning, and reference of our logical words gets explained, since he thinks we are in the business of offering a metasemantical account for logic.

But this conception of how to defend or reject Logical Charity is a mistake. First, let us note that a bold claim follows from Logical Charity if it is to be a consequence of charitable interpretation. As I said above, an interpretation is charitable if it interprets a speaker’s utterances so that any beliefs we attribute to her on the basis of our interpretation come out rational. In short, rationality places demands on what we can interpret speakers as saying. If Logical Charity is really a consequence of charitable interpretation, then rationality not only places demands on what we can interpret speakers as saying by logical sentences, but rationality also places the demand that what speakers say in accepting any logical sentence is true and that what they say in rejecting any logical sentence is false. More exactly, Total Demand follows:

**Total Demand**: for any logical sentence \( \varphi \), for any \( L \) and speaker \( S \) of \( L \): if \( \varphi \) is (potentially) clearheadedly accepted [rejected] by \( S \), and \( S \) says what she believes, then \( S \) does not express a rational belief by means of her acceptance [rejection] of \( \varphi \) unless \( \varphi \) is true [false].

Total Demand is a bold claim, since it fails to hold in other areas of discourse. One can accept the truth of sentences of astronomy or sentences about what is in one’s local grocery store and express rational beliefs about those subject matters by those sentences even though one’s beliefs...
are untrue. According to **Total Demand**, the state of affairs in logic is different: rationality demands that the *all* sentences one accepts [rejects] express truths [falsehoods].

The key point is this: since **Total Demand** presupposes a substantive claim about what it is rational to believe, Warren is wrong to think its defense take place wholly within metasemantics. For what is rational to believe depends heavily upon our substantive views on how to figure out the truth of what our sentences express. (I am not committing to an ontology of propositions here, but merely contrasting what is said from the strings used to say them, whatever that fundamentally amounts to). For example, what is rational to believe in chemistry or about my grocery store depends heavily upon how to figure out the truth about chemicals or the grocery store. In short, it depends significantly on epistemology. So, defending **Total Demand** requires taking substantive views in the epistemology of logic that require defense.

Perhaps I do not have to defend a substantive epistemology to defend how I interpret speakers’ utterances about what is in my local grocery store, even if I am committed to such an epistemology. So one might think I am placing too heavy a burden on Warren. But **Total Demand** is a strong claim that needs explanation – it tells us that rationality *demands all* of beliefs as expressed by logical sentences to come out true if accepted or false if rejected. The default position in a given area of discourse is that this is not the case. For example, someone with an incorrect astronomical theory can be rational while expressing untrue beliefs of her theory with astronomical sentences she accepts. So, the default position in logic is this:

**Default**: Someone with an incorrect logical theory can be rational while expressing untrue [true] beliefs in that theory with the sentences she accepts [rejects].

Warren must explain why **Default** is wrong. And in fact, it seems that what likely are the two best ways to show it is wrong will be at best controversial, if not implausible. One way is to argue that one cannot but express a true [false] belief by ones accepted [rejected] sentences when one is being rational because endorsing the correct, or a correct, logical theory is a *requirement of rationality itself*. And assuming rationality is normative, this is a version of the thesis that *logic is normative*, a thesis that has recently received much discussion and criticism in the philosophy of logic literature.\(^1\) It is a rather strong version at that: it does not merely say that logic comes with normative constraints, but that general normative constraints *include commitment to a logic itself*. Another way is to deny the presupposition of **Default** and argue that there actually are no incorrect logical theories – something akin to Carnap’s Principle of Tolerance.\(^2\) Hence, rational speakers cannot fail to express true [false] beliefs by their accepted [rejected] logical sentences because *no logic anyone rationally endorses can fail to be correct*. Those with some familiarity with the philosophy of logic can see that neither of these ways are without their difficulties.

**Section 3. Quine and Naturalism in Logic**

---

1. There is much recent literature on the normativity of logic; for recent critical work, see Russell (2017) and Blake-Turner and Russell (forthcoming).
2. See Carnap (1934, pp. 51-52).
My argument has two upshots. The first is that there is a limitation in Warren’s broader philosophical aims that Quine pointed out long ago. In several recent papers, Warren (2018) included, Warren has aimed to support the thesis of **Logical Conventionalism**:

**Logical Conventionalism:** “The rules governing the use of logical connectives fully explain the truth of logically true sentences in our language.” (Warren 2017, p. 124).

As his discussion in (Warren 2015b) and (Warren 2017) shows, those “rules governing the use of logical connectives” are certain inference rules governing the use of those connectives, including but not limited to the introduction and elimination rules for them. On Warren’s view, the logical conventions of a language are the collection of inference rules for every connective of the language. These logical conventions are rules which are often implicit in speakers’ behavior and thus need not be introduced explicitly all at once (2017, p. 124). This is how Warren in Warren (2017) argues his account avoids Quine’s objection in Quine (1936) to the idea that we can introduce the truths of logic by explicit conventions. Warren’s conventionalism is thus an implicit conventionalism, which he argues in Warren (2017) also avoids Quine’s separate objections to implicit conventionalism.

On Warren’s view, “[t]he meaning determining rules for any expression are automatically valid (necessarily truth preserving)” (Warren 2015b, p. 4). Although Warren says his defense of his implicit conventionalism awaits a forthcoming book, I take it Warren thinks the inference rules for a connective c are meaning determining for c because we can extend CLCM from Warren (2018) to logical rules. Since Warren supports CLCM by means of Logical Charity, I take it he thereby supports his implicit conventionalism by means of Logical Charity.

What is Quine’s criticism of implicit conventionalism? In Quine (1936), he says:

In dropping the attributes of deliberateness and explicitness from the notion of linguistic convention we risk depriving the latter of any explanatory force and reducing it to an idle label. We may wonder what one adds to the bare statement that the truths of logic and mathematics are a priori, or to the still barer behavioristic statement that they are firmly accepted, when he characterizes them as true by convention in such a sense. (p. 106)

In (1954), he says that implicit conventionalism simply is a “metaphorical” version of what he calls “linguistic doctrine of logical truth” (p. 115). The clearest implication of the linguistic doctrine of logical truth he can ascertain is that, roughly, a change of logic implies a change of meaning. But he thinks this simply follows from what he takes to be the obviousness of classical logic. Since rational speakers typically do not fail to accept obvious truths and do not reject obvious falsehoods, any sentence which appears to go against classical logic – such as ‘It is snowing and it is not snowing’ – must be interpreted differently. Given this, difference in meaning of logical sentences is “adequately accounted for by mere obviousness of logical

---

3 See also Warren (2015a), (2015b), and (2017).
4 Warren adopts a “holist inferentialism” on which the inference rules for c are “the total package of rules whose instances can contain occurrences of the expression even if those rules don’t explicitly involve the expression.” (2015b, p. 13)
5 Some confirmation: on (2018, p. 440, fn. 56), he says: “I develop a full philosophical theory of logic in part II of my book, Shadows of Syntax. The theory developed there provides a full background into which the arguments of this paper [i.e., arguments for Save the Logical Attitudes, Logical Charity, and CLCM] nicely fit.” He also says that CLCM can be extended to inference rules on (2018, p. 423).
principles, without help of a linguistic doctrine of logical truth” (pp. 112-113). So, insisting logical truth is linguistic “leaves explanation unbegun.” (p. 113)

Quine’s main point in both papers is that implicit conventionalism boils down to an undefended and unarticulated epistemology. And that is what I have argued about Warren’s view. Implicit conventionalisms like Warren’s are theses in metasemantics which conceal theses in epistemology. Worse, we do not know which epistemology gets concealed. As Quine points out in (1936), for all we know, implicit conventionalists must accept the a priori - rationality demands us to express logical truths since logic comes from capital-R Reason. Or, for all we know, they must accept Quine’s strong (and, for what it is worth, I believe incorrect) view that classical logic is obvious. Or they must accept an extremely permissive, Carnapian epistemology on which logics cannot fail to be correct. Assuming implicit conventionalism merely is a “linguistic doctrine” “leaves the explanation unbegun.” Hence, contrary to what Warren argues, he has not addressed Quine’s objections to implicit conventionalism.

The second upshot is about the metaphilosophical aspirations of Warren’s project. Warren wants a naturalistic metasemantics of logic (see 2015b, p. 1). But Warren’s failure to address Quine’s objection shows his metasemantics is consistent with many epistemologies of logic, and so likely consistent with many anti-naturalistic epistemologies of logic – whatever naturalism is, surely not every version of the thesis that logic is a priori is naturalistic. The naturalistic credibility of his project is open to question. I recommend, then, that a naturalistic philosophy of logic requires doing naturalistic epistemology of logic first.

Works Cited

Carnap, R. (1934). The Logical Syntax of Language. Translated by A. Smeaton. La Salle, Ill.: Open Court Publishing.


