Quine on Conventionalism

John Peloquin

Conventionalism attempts to explain the apriori and necessary character of logical and mathematical truths by arguing that it is a consequence of linguistic conventions governing logical and mathematical expressions. In "Truth by convention",[1] Quine examines several different approaches that conventionalism might take, including (i) claiming that all mathematical terms can be given simple logical definitions; (ii) claiming that all mathematical terms can be given contextual (schematic) definitions in terms of conditional logical sentences; (iii) claiming that mathematical sentences are just abbreviations for conditional sentences containing primitive (undefined) mathematical terms; and finally (iv) claiming that conventions can be used to specify the truth, and hence the meaning, of logical and mathematical sentences, the latter of which may include some primitive mathematical terms. Quine thinks (i) is hopeful, but (ii), (iii), and (iv) fail. In this paper, I examine aspects of (i) and (iv).

The first approach claims that all mathematical terms can be given simple logical definitions. A *simple* definition is a stipulation that one expression, the *defined* expression, serve as an abbreviation for another expression, the *defining* expression. Because such definitions are eliminable—that is, a defined expression may always be replaced by its defining expression—mathematics does not gain any expressive power over logic under this approach. Indeed, mathematical sentences are just abbreviations for logical sentences, so the propositions expressed by mathematical sentences are just the propositions expressed by their corresponding unabbreviated logical sentences, and mathematics is a subset of logic.[2] But if this is so, then how should we understand the conventionalist thesis that mathematical truths are consequences of

[1] [2]

[2] [2], p. 331. While introducing defined terms allows one to express new *sentences*, it does not allow one to express new *propositions*.

linguistic conventions? For this approach to work, the truths of logic must themselves be true by virtue of linguistic conventions governing logical expressions.[3] This can be cashed out in a number of ways, one of which I examine below. The important point is that the definitions cannot be understood as truths relating each mathematical truth to a distinct logical truth. Indeed, a mathematical truth *just is* a logical truth under this approach. The definitions just relate mathematical terms to logical terms, and hence mathematical sentences to logical sentences, by way of abbreviation.

The final approach attempts to use conventions to specify the truth, and hence the meaning, of logical and mathematical sentences. In his treatment of this approach, Quine outlines a method the conventionalist might use for pure logic. The rough idea is to (i) identify a small set of primitive logical symbols, like "$\neg$", "$\rightarrow$", and "$\exists$", (ii) fix a set of sentences in which these symbols appear and adopt the convention that the symbols mean whatever they have to mean in order for all and only the sentences in the set to be true, and finally (iii) define all other logical symbols in terms of these. Doing this specifies the truth or falsity, by convention, of all logical sentences (subject to essential incompleteness limitations).[4]

It might be wondered whether, when using this method, the conventionalist should require the set of sentences designated in (ii) to be consistent. This requirement could be formulated as a condition that the set not specify, for any sentence p, the joint truth of the sentences p and $\ulcorner \neg p \urcorner$. Because the primitive logical symbols are initially treated as meaningless, and only endowed with meaning indirectly by the truth conventions in (ii), technically speaking there is no requirement for consistency. Indeed, the conventionalist is free to specify whatever meanings she wants for the primitive symbols. However, consistency is relevant because the conventionalist is ultimately aiming to capture our intuitive, pretheoretic meanings of "not", "if, then", and "for all". The consistency requirement can be justified by the fact that, intuitively and pretheoretically, we never consider a sentence and its negation to be jointly true, and the

[3] [2], p. 338.
[4] [2], p. 345.

conventionalist wants the meaning of "$\neg$" to match that of negation.[5]

Does the conventionalist approach to pure logic outlined here explain the apriori character of our knowledge of logical truths? Consider the following two questions:

(1) How do we know that the sentence "if snow is white, then snow is white" is true (in the English language)?

(2) How do we know that if snow is white, then snow is white?

These are very different questions. The first concerns our knowledge of a metalinguistic proposition about the truth of a conditional sentence. This knowledge will depend, in part, on our knowledge of the meanings of the words in the sentence, knowledge about the truth predicate, and so on. The second question concerns our knowledge of a conditional proposition, which does not involve any sentence. This knowledge need not be based on any metalinguistic knowledge. Indeed, we might just know apriori that propositions imply themselves, in the same way we know apriori that objects are identical to themselves. How can the conventionalist account for this? The conventionalist can answer (1) by showing that the truth of "if snow is white, then snow is white" follows from the truth conventions specified for pure logic.[6] The conventionalist might try to argue that, knowing this, we can conclude apriori that if snow is white, then snow is white, and so this approach also provides an answer for (2). But this will not work, because *linguistic conventions are not knowable apriori.* One can know apriori that if snow is white, then snow is white without knowledge of any linguistic conventions, and the conventionalist approach cannot account for this. In general, the conventionalist cannot explain apriori knowledge by appeal to linguistic conventions which are not knowable apriori.

Quine thinks this approach is also circular, and leads to infinite regress. It is circular because, in specifying truth conventions, it makes use of our intuitive, pretheoretic, notions of "not", "if, then", and "for all".[7] We are led into infinite regress when we attempt to infer the truth of

[5] [2], p. 345–6.

[6] For reasons of space, I omit the explicit reconstruction of this argument. In [2], p. 344, simply substitute "snow is white" for "time is money" in the sentences in Quine's argument.

[7] [2], p. 352.

particular logical sentences from the general conventions. For example, in order to infer the truth of "if snow is white, then snow is white" (or equivalently "snow is white or snow is not white", given our definition of "or") from the conventions, we must make use of a convention which, essentially, encodes modus ponens (the convention states that if p and q are sentences, and p is true, and $\ulcorner$If p, then $q\urcorner$ is true, then q is also true). But the only way we can infer the truth of a sentence from this convention is by use of modus ponens, which the convention itself is supposed to justify. Therefore we are engaged in an infinite regress.[8] This is a problem for the conventionalist because she should like to be able to infer truth of particular logical sentences from the conventions. But the inferences do not follow from the linguistic conventions alone.

It seems like this problem may not be specific to the conventionalist, as the problem seems similar to one described by Carroll in "What the Tortoise Said to Achilles".[9] If one inquires as to the justification of the inference from premises (a) if p, then q and (b) p to the conclusion (z) q, one sees that this is justified by modus ponens. But one might be tempted to encode modus ponens as a premise in the argument, namely (c) if (if p, then q) and p, then q. But if one then inquires as to the justification for the inference in the new argument, from premises (a), (b), and (c) to the conclusion (z), one sees that this is again justified by modus ponens. Again one might be tempted to encode this as a premise, and so on to an infinite regress. This puzzle illustrates that modus ponens cannot be treated merely as a premise (or schema of premises) and be expected to serve us, but must be treated as a fundamental inference rule telling us how we may legitimately act upon premises. Without being able to act upon premises in the manner prescribed by modus ponens, we would never be able to get off the ground in our logical reasoning, as illustrated by the puzzle.

Quine does not seem to think this is a problem for logic ordinarily, but thinks it is a problem for the conventionalist because the conventionalist essentially tries to capture modus ponens in a linguistic truth convention, something analogous to a premise. Given the use of this con-

[8] For reasons of space, I omit a detailed reconstruction. In [2], p. 351–2, simply apply Quine's argument to our example sentence.

[9] [1]

vention, it seems the conventionalist will experience the infinite regress described by Carroll, because in attempting to infer something from the convention, she must use an inference rule ultimately justified by the convention itself.

The conventionalist might reply that the convention is correctly understood as a *proposition* encoding something analogous to an inference rule for *sentences* (namely, telling us how to move from true sentences to true sentences), and since the the convention itself need not be formulated as a sentence, and logic is concerned only with the truth of logical sentences, there is no infinite regress. The problem with this reply is that, logic is not just concerned with the truth of sentences, and we must use inference rules on propositions when working with the conventions. This is really just a variant of the problem seen earlier, that knowing the truth of sentences is not the same as knowing the truth of propositions. This confusion seems to be a major failing of conventionalism.

References

[1] Carroll, L. "What the Tortoise Said to Achilles." *Mind* 4 (1895), p. 278–80.

[2] Quine, W. V. "Truth by convention." *Philosophy of Mathematics*, eds. Paul Benacerraf and Hilary Putnam, p. 329–54.