Carnap's Problem, Definability and Compositionality

April 7, 2024

According to inferentialists the meaning of logical vocabulary is determined by its use in inferences. Hard-line versions of the position –e.g. (Došen 1989), Peregrin 2014)– characterise meaning directly in terms of inference rules. To be a certain connective, on these accounts, is just to be governed by certain rules. A moderate form of inferentialism is favoured by Hacking (1979), Garson (2013) or Murzi and Topey (2021). Moderate inferentialists don't identify meanings with inference rules, but still claim that the meaning of logical vocabulary can be, in some sense, read off its role in inference.

A well-known result by Carnap poses a problem for moderate inferentialism. In his *Formalization of Logic* (1943), Carnap pointed out that there are non-normal interpretations of classical logic: non-standard interpretations of the connectives and quantifiers that are nevertheless consistent with the classical consequence relation of the appropriate language. So, if we take inferential roles to be given by consequence relations, the meaning of classical logical vocabulary cannot be read off inferential roles. Let us call this Carnap's Problem.

In a recent paper Bonnay and Westerståhl (2016) put forward a solution to Carnap's Problem. Their approach is to limit the space of possible interpretations by 'universal semantic constraints'. According to Bonnay and Westerståhl, if we restrict attention to interpretations that are (a) compositional, (b) non-trivial and (c) in the case of the quantifiers, invariant under permutations of the domain, Carnap's Problem is avoided.

In this talk I will point out two problems with Bonnay and Westerståhl's approach, and suggest a way to fix them.

The first problem concerns the main result of Bonnay and Westerståhl's paper, a characterisation of the interpretations of $\forall$ that are consistent with the classical consequence relation of a language:

(BW) Let $\mathcal{L}$ be a language with $\forall$ primitive, let $\mathcal{M} = (D, I)$ be an $\mathcal{L}$-structure and let $Q \subseteq \mathcal{P}(\mathcal{D})$ be the denotation of $\forall$ (seen as a generalized quantifier). Then a weak model $\mathcal{M}, Q$ is consistent with the classical consequence relation for $\mathcal{L}$ iff Q is a principal filter closed under the interpretation of terms in $\mathcal{M}$.

Crucially, Bonnay and Westerståhl only prove (BW) for first-order languages supplemented with predicate *variables* (in other words, for *second*-order languages without second-order quantifiers). I will show —adapting the methods of (Antonelli 2013)— that (BW) fails for first-order languages, and as a result the normal interpretation of quantifiers is not fixed. The underlying problem here regards definability. It's well-known that given a first-order language and a structure for it, there usually are subsets of the domain that cannot be defined by a formula. This general fact can be easily exploited to define non-normal interpretations, and makes Carnap's Problem for first-order languages particularly challenging.

The second problem with Bonnay and Westerståhl's approach concerns the way they define interpretations. Although they hold that we should restrict attention to compositional interpretations, their normal interpretations of first-order languages aren't compositional after all. I will also show that if we redefine interpretations to avoid this problem, compositionality, non-triviality and invariance under permutations don't pin down the standard meaning of logical vocabulary. In this case the underlying problem is Bonnay and Westerståhl's demand for compositionality itself. The usual, two-valued semantics for first-order logic is *not* compositional. While different compositional semantics for first-order languages are available, they all involve a large range of semantic values. This, in its turn, makes Carnap's Problem more difficult to solve: more semantic values means more possible interpretations, and therefore more non-normal ones that may be consistent with a given consequence relation. In the first-order case, then, demanding compositionality is counter-productive.

After expanding on these problems I'll propose a way to modify Bonnay and Westerståhl's solution that avoids them. Roughly put, I'll argue that there are plausible reasons to strengthen the notion of an interpretation being consistent with a consequence relation, and that this is enough to clinch the usual interpretation of classical logical vocabulary.

References

[Antonelli (2013)] Antonelli, G. A. (2013). On the general interpretation of first-order quantifiers. *Review of Symbolic Logic* 6 (4):637-658.

[Boghossian (2000)] Boghossian, P. (2000). Knowledge of Logic. In Paul Boghossian and Christopher Peacocke (eds.), *New Essays on the A Priori.*

[Bonnay and Westerståhl (2016)] Bonnay, D. and Westerståhl, D. (2016). Compositionality Solves Carnap's Problem. *Erkenntnis* 81 (4):721-739.

[Carnap (1943)] Carnap, R. (1943). *Formalization of Logic.* Cambridge: Mass., Harvard University Press.

[Došen (1989)] Dosen, K. (1989). Logical Constants as Punctuation Marks. *Notre Dame J. Formal Log.*, 30, 362-381.

[Hacking (1979)] Hacking, I. (1979). What is logic? *Journal of Philosophy*, 76 (6):285-319.

[McGee (2000)] McGee, V. (2000). 'Everything'. In G. Sher and R. Tieszen (Eds.), *Between Logic and Intuition: Essays in Honor of Charles Parsons.* Cambridge: Cambridge University Press.

[Murzi and Topey (2021)] Murzi, J. and Topey, B. (2021). Categoricity by convention, *Philosophical Studies* 178 (10):3391-3420.

[Peregrin (2014)] Peregrin, J. (2014). *Inferentialism: Why Rules Matter.* Palgrave Macmillan, London