# Logic 450 Writing Assignment

One of the goals in this class is to get good at communicating technical ideas. For this assignment, you'll write a short essay about an application of formal logic to the philosophy of language, mind, mathematics, or metaphysics. (I'm also open to other topics.)

**Due: April 28** (by email to jeff.russell@usc.edu)

**Length:** 1000–2000 words

You may turn this in any time during the semester. Please talk to me if you think an extension would be helpful.

I'm also happy to give you comments on a draft. This is optional, but recommended. The deadline for turning in a draft is **April 14**.

The deadline for discussing an alternative topic with me is **April 14**.

## Topics

Choose one of the following topics, or discuss with me if you'd like to write on a different topic (no later than April 14).

Each of these questions is very big, and could lead in a lot of different directions. I'm not expecting anything like a comprehensive answer. I'm looking for two things: (1) a succinct, clear explanation of a philosophical problem, and (2) a brief discussion of one interesting idea about how to approach it.

I am happy to recommend further reading.

1. **(Formalism)** Hilbert (among others) argued for a *formalist* philosophy of mathematics: for a mathematical statement to be correct merely consists in its being derivable by certain basic rules from certain basic axioms. ("The subject matter of mathematics is, in accordance with this theory, the concrete symbols themselves whose structure is immediately clear and recognizable." But I should note that my brief characterization of formalism simplifies Hilbert's actual view a bit.) Do Gödel's Incompleteness Theorems show that this view is false?

   *Reading:* Hilbert, "On the Infinite"

2. **(Rationalism)** Many historic philosophers have held that all truths can, in principle, be deduced *a priori* from first principles. (These philosophers might include Leibniz, Spinoza, Carnap, and more recently Chalmers, though they differ about what the "first principles" might be.) Gödel's Theorems make serious trouble for this idea: not even all truths of *arithmetic* follow in any straightforward way from a simple set of axioms. Does this refute the rationalist's ambitious proposal?

   *Reading:* Chalmers, *Constructing the World*, sec. 6.2.

3. **(The Computational Theory of Mind)** Gödel's Theorem shows that for any axiomatizable true theory of arithmetic, there is a sentence—the *Gödel Sentence* for that theory—which is unprovable, and also its negation is unprovable. But *we* can recognize that the Gödel sentence is true—indeed, we can prove it! Lucas and Penrose have each argued that this shows that our capabilities to recognize mathematical truths—our intuitive "axioms"—cannot be recursively enumerated, and so humans can do something that no finite machine can do. Thus the human mind cannot be adequately described as a finite computational machine. Is this argument correct?

   *Reading:* J.R. Lucas, "Minds, Machines, and Gödel"

   (Apology: this article, published in 1961, includes some pretty bad offhand sexism.)

4. **(Skolem's Paradox)** We can write down the principles of set theory as a first-order theory, which is called ZFC. One of the things we proved using set theory was that there is an uncountable set. So the first-order formulation of this sentence is a consequence of ZFC. But the Löwenheim-Skolem Theorem tells us that if ZFC is consistent, it has a *countable* model. Every first-order consequence of ZFC is true in this structure. So in particular, the first-order formulation of "There is an uncountable set" is true in this structure. Is this a contradiction? What does this show?

   *Reading:* Skolem, "Some Remarks on Axiomatized Set Theory"

5. **(Realism)** The Lowenheim-Skolem Theorem shows that any first-order theory that has an infinite model has many different models of vastly different sizes—and so, in particular, these models are not isomorphic to one another. But this would apply in particular to the theory eventually delivered by "ideal science"; even this idealized theory cannot pin down what the structure of the world is really like, since any of these vastly different structures would still satisfy the theory. As Putnam puts it, 'the *total use of the language* ... does not "fix" a unique "intended interpretation."' But if this is so, then how can our language correspond to the unique structure of the "ready-made world"?

   *Reading:* Putnam, "Models and Reality"

   *Further reading:* Lewis, "Putnam's Paradox"

6. **(Truth)** Consider a theory consisting of all true English sentences. It seems like this is a theory that is logically consistent, can represent its own syntax, and can represent itself (using the predicate "is a true English sentence"). This apparently conflicts with Tarski's Theorem. What should we conclude?

   *Reading:* Tarski, "The Semantic Conception of Truth"

7. **(Necessity)** Quine argues that there is an important philosophical difference between expressing necessity using a *predicate of sentences*, as in " 'Seven is less than nine' is a necessary truth", and alternatively using a *sentence operator*, as in "It's necessary that seven is less than nine". The second kind of use has "special dangers": it tempts us to "quantify into" this sentence operator, as in "Something is necessarily less than nine", which "leads us back into the jungle of Aristotelian essentialism" (p. 174). According to Quine, the only safe way of using a necessity operator is to understand it in a way that can be reduced to the linguistic predicate.

   Montague's Theorem is a variant of Tarski's Theorem: it shows that no predicate of sentences "necessarily true" can obey the certain basic principles of modal logic. (Roughly, these principles say: each logical truth is necessarily true; and each necessary truth is true.) What does this theorem tell us about the prospects for Quine's "linguistic" theory of necessity?

   *Reading:* Quine, "Three Grades of Modal Involvement"; Montague, "Syntactical Treatments of Modality"