ABSTRACT. The model-theoretic argument, which Putnam employs to argue against metaphysical realism, has faced serious objections of many realist opponents. Igor Douven in his recent paper offers a new interpretation of the model-theoretic argument, which avoids the previous objections. The purpose of this paper is to show that Douven’s reconstruction of Putnam’s argument is not successful, and hence that the realist objections still stand.

Let me first introduce Douven’s reconstructed argument, which depends on three premises.

The first premise is derived from the so-called Löwenheim-Skolem argument. Assume the world consists of infinitely many things. Let $T$ be an ideal empirical theory. Then $T$ must be consistent. Then by the completeness theorem it is satisfiable. Thus by the Löwenheim-Skolem theorem $T$ has a model of every infinite cardinality. Pick a model $M$ of the same cardinality as the world. Then we can establish a satisfaction relation $\text{SAT}$ by mapping $M$ one-to-one onto the world. If $T$’s language is
interpreted according to SAT, $T$ comes out true of the world; that is, $T$ is true-on-SAT. This argument seems to show the following:

(P₁) Any consistent empirical theory is true-on-SAT in any world on some interpretation SAT.

According to Douven, (P₁) expresses a necessary truth. Thus, it can be formalized as follows:

(1) $\Box \forall x \exists \text{SAT}[\text{True-on-SAT}(x)]$, where $x$ ranges over all and only (actual and potential) consistent empirical theories. This premise is very plausible. And it is consistent with metaphysical realism. Given a consistent empirical theory, there are some interpretations that make the theory come out true. Then there can be some feature – call it $F$ – that distinguishes these interpretations from other interpretations that make the theory false. According to the metaphysical realist, there is some feature $F^*$, other than $F$, distinguishing among the existing correspondence relations between our vocabulary and reality the putative unique reference-determining correspondence relation. For, as Douven points out, “[t]he realist acknowledges that there is an abundance of correspondence relations between our vocabulary and reality, but he insists that only one of those is intended, only one of those is reference”.

The second premise of the reconstructed argument is the following (P₂):

(P₂) If there is not some feature $F^*$, other than $F$, distinguishing among all the existing correspondence relations between the words of a theory’s language and things in the world the putative unique reference-determining correspondence relation, then a theory is true exactly if it is true-on-SAT on some correspondence relation SAT.

On Douven’s view, Putnam regards (P₂) to be a conceptual truth, at least from the perspective of the metaphysical realist. For if there is no way for the metaphysical realist to discriminate between truth-on-some-interpretation and truth simpliciter, then from the metaphysical realist’s point of view truth-on-some interpretation cannot be logically distinguished from truth simpliciter. Let us abbreviate the antecedent of (P₂) as (N). Then (P₂) can be formalized as (2):

(2) $\Box[(N) \rightarrow \forall x (\exists \text{SAT}[\text{True-on-SAT}(x)] \leftrightarrow \text{True}(x))]$.

I shall critically discuss this premise later.