Position paper: Proof-Theoretic Semantics as a viable alternative to Model-Theoretic Semantics for natural language

Nissim Francez

This paper is a response to the last question in the CFP, about alternatives to model-theoretic semantics (MTS). I do not present any specific results, just argue for proof-theoretic semantics (PTS) as such an alternative. PTS is well-established within Logic (e.g., [8]). See [11] for an overview. I have extended PTS to fragments of English.

The paper has two parts: 1. a brief exposition of PTS, and 2. criticism of MTS as a theory of meaning and advantages of PTS as such a theory.

1. In a nutshell, the PTS programme can be described as follows.

I. Sentences: replace the received approach of taking their meanings as truth-conditions (in arbitrary models) by taking their meanings as canonical derivability conditions (from suitable assumptions) within a “dedicated” natural-deduction proof-system in which the derivability conditions are formulated. Such a system is presented in [3]. In a sense, the proof system should reflect the “use” of the sentences in the considered fragments, and should allow recovering pre-theoretic properties of the meanings of sentences such as entailment, assertability conditions and consequence drawing.

II. Subsentential phrases (down to lexical units): replace their denotations (extensions in arbitrary models) as meanings by their contributions to the meanings (canonical derivability conditions) of sentences in which they occur. This adheres to Frege’s context principle [5], made more specific by the incorporation into a type-logical grammar [6] for the fragment considered. This is elaborated upon in [4].

Dummett introduces ([1], p. 48) an important distinction between content and ingredient sense. The content of an (affirmative) sentence $S$ is the meaning of $S$ “in isolation”, on its own. The ingredient sense of $S$ is what $S$ contributes to the meaning of an $S'$ in which $S$ occurs as a sub-expression, a component. This distinction is incorporated in the PTS in [4]. A grasp of the ingredient sense always includes a grasp of the meaning, but not necessarily vice versa. This distinction, incorporated in the PTS in [4], is propagated to sub-sentential phrases too.

The success of PTS is manifested in [2], where conservativity of determiners is proved rather than stipulated as a universal.

2. There is a vast literature containing critical arguments against MTS as a theory of meaning. I present here briefly only some of the main ones, those pertaining directly to NL. Some involve philosophical considerations, and others - not. My
personal position is closely related to the latter sort of criticism.

I. The most famous criticism is Dummett’s *manifestation argument*, e.g., [1], Ch. 13, associating meaning of an affirmative sentence with the understanding of that sentence, an understanding manifesting itself as involving the ability (at least in principle) to verify the sentence as a condition for its assertability. Trans-verificational truth is rejected since it is not reflecting a cognitive process of understanding (this is where the philosophical position of anti-realism emerges); this is followed by a rejection of *bivalence*, according to which every (affirmative) sentence has a truth value (either true or false), independently of any ability to verify what that value is. Since MTS cannot identify uniquely a model corresponding to the actual world, verifiability means *deciding*, given an arbitrary model, whether the truth-conditions obtain in it. In general, this task is impossible even for the simplest sentences, involving only predication, as set membership is not decidable\(^1\) in general. This contrasts the common situation where derivability in proof-systems is decidable, due to the availability of terminating proof-search procedures.

II. Another kind of criticism of MTS questions its *explanatory power*. The received wisdom regards MTS as a formalization of the relationship between language and the world. Quine (in [9]) relates to this view as “the museum myth”: NL expressions are stuck on objects like labels in a great museum. The claim is that no theory can succeed in directly relating language to the world. At most, language is related to some meta-language (e.g., some set-theoretical language), used to specify models and truth-conditions in them. This is true for MTS in general, but is particularly relevant to the case of NL, which is its own ultimate meta-language. Since I find this criticism a very compelling one, independent of philosophical stand on metaphysical issues, I want to elaborate more on it. Consider the usual MTS definition of conjunction ‘and’, using the usual models: \(M\models S_1 \text{ and } S_2 \iff M\models S_1 \text{ and } M\models S_2\). How does such a clause define the meaning of ‘and’? Unless the meaning of ‘and’ (in the meta-language, here English) is *already* known, this does not define meaning at all! Otherwise, a similarly structured definition of a connective ‘blob’ would be equally well-defined by \(M\models S_1 \text{ blob } S_2 \iff M\models S_1 \text{ blob } M\models S_2\).

III. One may feel some dissatisfaction with the *ontological commitment* accompanying MTS, relating to various entities populating models: possible-worlds, events (and their participants), times, locations, degrees, kinds and many more. As emphasized by Paoli [7] when adhering to PTS, the definition of meaning need not appeal to any external apparatus; it can use the (syntactic!) resources provided by the rules of the underlying deductive system, which are artefacts of this system, devoid of any ontological commitment. A related issue, associated with entities in models, is the problematic possibility of quantifying over “absolutely everything”, accompanying MTS, see Rayo and Uzquiano [10].

IV. There is a criticisms of MTS as a theory of meaning, pointing to an advantages of PTS as such a theory, which is independent of cognitive and/or epistemic considerations, as well as from metaphysical ones. A notorious problem of MTS is its coarse granularity of meaning, where logically equivalent propositions, which have the same truth-conditions, are assigned the *same* meaning. As a simple example, in propositional Classical Logic, we have the equivalence \(\varphi \land \psi \equiv \neg (\neg \varphi \lor \neg \psi)\). Both

\(^1\)There is no precise statement by Dummett as to what is taken as “decidable”. It is plausible, at least in a computational linguistics context, to identify this notion with *effectiveness* (i.e., *algorithmic decidability*).
sides of the equivalence are assigned the same meaning (here, same truth-table). However, those two proposition do differ in several aspects involving meaning, most notable in inference. While it is fairly natural to regard a transition from \( \varphi \land \psi \) to \( \varphi \) as some kind of an “elementary” transition, a transition from \( \neg (\neg \varphi \lor \neg \psi) \) to \( \varphi \), while valid, cannot count as elementary, and its validity needs explanation by means of decomposition to more elementary steps. A fortiori, the same applies to more complicated, less transparent logical equivalences. In particular, all logical validities are assigned the same meaning. In natural language this discrepancy is even more salient. Identifying the meanings of every girl is a girl with that of every flower is a flower, and even with that of no bank is a non-bank, is clearly inadequate. In PTS, directly appealing to inferential roles for conferring meaning, finer-grained meanings are obtained, not suffering from this deficiency.

In summary, while the application of PTS as a formal semantics of NL is still in its infancy, I believe it is a serious contender to the traditional, well-established MTS. [In the talk, I intend to explain and exemplify the proof-theoretic concepts involved in PTS, which might not be familiar to many semanticist entrenched in MTS.]

References


