CATEGORICAL MODEL THEORY AND THE SEMANTIC VIEW OF THEORIES

ANDREI RODIN

1. Categorical Model theory

Today's Categorical Model theory (CMT) stems from the functorial semantics of algebraic theories proposed by Lawvere in his thesis back in 1963 [4]. This theory uses a family of concepts of model none of which can be called today fairly standard. This fact is evidenced by the continuing discussion in the Homotopy Type theory (HoTT) [5] where presently there is no full agreement among the researchers in the field as to what counts as a model of this theory and what does not.

One approch relies on the concept of *classifying category* T freely generated from the syntax of the given theory. Then a model M is a functor $T \to C$ into the category of sets ($C = Set$) or another appropriate category. This functorial setting has an important universal property: up to the categorical equivalence T can be identified with the initial object in the functor category of T-models. This property allows one to think of a theory in this setting as being a "generic model" (Lawvere). Using this approach Awodey [1] defines for HoTT the concept of *natural* model .

Voevodsky [7] pursues a different approach, which involves the concept of *contextual category* (more recently - in a modified form of C-system) earlier proposed by Cartmell [2]. The idea behind the concept of contextual category is that of a category, which fully encodes all relevant algebraic features of the given syntax. According to this approach those and only those categories, which fall under the corresponding definition of contextual category, qualify as models of given theory T. In this case the initiality property of the syntactic category $S(T)$ is not implied by any general theorem. The initiality conjecture for HoTT still stands open.

Finally, there is yet another approac in CMT, which involves the concept of *internal language* (aka *internal logic*) of a given category. It has been recently proposed to think of internal languages and syntactic categories in terms of adjoint functors between a category of theories and a category of categories as shown on the diagram below:

$$Categories \underset{Synt}{\overset{Lang}{\rightleftarrows}} Theories$$

In this setting a model of given theory T in a certain ground category C is a functor (a morphism in the category of theories) of the form

$$M : T \to Lang(C)$$

which expresses the idea of representation of a given theory in the language of some other theory (such as a representation of some geometrical theory in the language of arithmetic).

These and other technical advances of CMT so far have no generally accepted epistemological underpinning, which might help one to orientate among multiple developments. It remains, generally, unclear whether or not the classical Tarskian notion of model based on the T-schema and its standard epistemological understanding can be helpful in CMT. In what follows I show that the classical Tarskian concept of model is not adequate for accounting for the model theory of HoTT in its existing form and propose a remedy. Then I argue that the proposed non-standard understanding of concepts of theory and model can be used for supporting a new version of the semantic of view of theories, which may help to bridge the persisting gap between the notion of model as it is used in logic, on the one hand, and the colloquial notion of model used elsewhere in science, on the other hand.

2. Modeling HoTT

I shall consider HoTT *without* the univalence axiom. In this case the syntax of HoTT is that of (the intensional version of) Martin-Lof's Constructive Type theory (MLTT). HoTT also involves a semi-formal interpretation of its syntax in the Homotopy theory: types are interpreted as spaces (more precisely, infinite-dimensional fundamental groupoids of such spaces) and terms are interpreted as points of these spaces. This interpretation helped to reveal a feature of MLTT's syntax, which earlier remained hidden. Namely, it has been observed that types in MLTT are stratified into the so-called *homotopy levels*. It is important to stress that this stratification is a robust mathematical fact but not just a matter of one's favorite informal interpretation of the given calculus. This stratification necessitates a revision of the informal "propositions-as-types paradigm" , which is popular in the Computer Science. It shows that only types of certain homotopic level (namely, of level (- 1) as defined in [5]) can be identified with propositions while the higher types should be interpreted differently. This revision implies, in particular, that HoTT cannot be coherently interpreted as a system of propositions or sentences; correspondingly, the Tarskian notion of model based on the T-schema and the satisfaction relation applies only to propositional types (and the corresponding rules) of HoTT but not to this theory as a whole.

MLTT is a system of formal rules without axioms. In the case of propositional types these rules can be called *logical* rules in the usual sense. When these rules are applied to the higher types they should be thought of as rules for constructing non-propositional objects. A *model* of MLTT-HOTT is an implementation of this system of rules in some background, where higher-order constructions play the role of truth-makers for their associated propositions. (A proposition associated with a given higher type T is obtained from T via its (-1)*truncation*). This basic interpretation agrees with all existing models of HoTT disregarding the subtleties mentioned above. An interesting epistemological question is this. Does the epistemic role of higher-order constructions in HoTT reduce to their role as truth-makers or there is something more to it? Since the truncation of higher types to propositional types, generally, involves a significant loss of structure, HoTT rather supports the second answer (unless one assumes that a major part of this theory is epistemically insignificant). In the following concluding section I provide an independent argument, which supports the same conclusion and explains the epistemic value of higher non-propositional structures in HoTT.

3. Semantic View of theories: a constructive perspective

P. Suppes [6] argued that a typical scientific theory should be identified not with any particular class of statements (formal or contentual) but rather with a certain class of models . On this basis Suppes and his followers designed a Bourbaki-style format of formal presentation where a scientific theory is presented through an appropriate class of itsset-theoretic models.Albeit such a Bourbaki-style presentation can be useful for purposes of logical and structural analysis, it appears to be useless as a practical tool, which may help working scientists to formulate and develop their theories in a formal setting [3].

Such a limitation is hardly surprising given that the standard set-theoretic semantics of theories provides no formal means for building and operating with models other than by referring to the fact that a model in question satisfies such-and-such propositions. Differences in epistemological views on the roles of syntax and semantics affect the style of formal presentation but not its architecture. This is why in practice the usual *non-statement* aka *semantic* approach to the formalization of scientific theories demonstrates the same limitations as its syntactically oriented rival.

HoTT and its model theory provides a novel notion of theory, which does not reduce to a class of propositions but has a further higher-order non-propositional structure. The axiomatic basis of such a theory consists of a system of rules, which apply both at the propositional and non-propositional levels. I believe that such a broader concept of theory and its model better fits the colloquial counterparts of these notions in the scientific practice than the standard Tarskian notions. The main reason isthat a typical scientific theory involves a lot of *procedural* content, which is used in modeling; such procedures may comprise but typically do not reduce to the procedures of logical inference (if by the logical inference one understands here a procedure which inputs and outputs sentences).

Thus HoTT and CMT provide the semantic view of theories with new formal techniques; the renewed semantic view, in its turn, provides an epistemological background for possible applications of these techniques in science and Knowledge Representation.

References

[1] S. Awodey. *Natural models of homotopy type theory.* arXiv:1406.3219, 2015.

[2] J. Cartmell. Generalised algebraic theories and contextual categories. *Annals of Pure and Applied Logic*, 32:209–243, 1986.

[3] H. Halvorson. Scientific theories. *The Oxford Handbook of Philosophy of Science (forthcoming); available at*, 32, 2015.

[4] F.W. Lawvere. *Functorial Semantics of Algebraic Theories. Ph.D.* Columbia University, 1963.

[5] Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* Institute for Advanced Study (Princeton); available at http://homotopytypetheory.org/book/, 2013.

[6] P. Suppes. *Representation and invariance of scientific structures.* Stanford, Calif. : CSLI Publications, 2002.

[7] V. Voevodsky. *Martin-Lof identity types in the C-systems defined by a universe category.* arXiv:1505.06446, 2015.