Model-theoretic Conservative Extension for HOL with Ad-hoc Overloading

Arve Gengelbach Tjark Weber
Uppsala University, Uppsala, Sweden
arve.gengelbach@it.uu.se tjark.weber@it.uu.se

Definitions of new symbols should merely abbreviate expressions in logical frameworks, and any new definition should only imply new facts that contain the definiendum. On the semantic side, this means any model of a theory of definitions should be extensible to accommodate new definitions, while preserving validity of propositions that contain only previously defined symbols. Generally, this property of a theory extension is called *model-theoretic conservativity*, and is expected to hold for extension by definitions.

In earlier work [1], we presented a notion of model-theoretic conservativity for higher-order logic (HOL) with ad-hoc overloading for eager ground semantics [3]. Here, we adapt this earlier work to lazy ground semantics and different dependency tracking [4], as explained below. We also mechanise our results in the HOL4 theorem prover.

HOL is widely used for formal verification in several proof assistants. Its most popular implementation is the Isabelle/HOL theorem prover, where users may define both types and constants. Additionally, Isabelle/HOL supports *ad-hoc overloading*: polymorphic constants may be defined differently at different (non-overlapping) types. If u is a constant, the definition $u \equiv t$ asserts equality between u and the term t. If u is a type, the definition introduces an isomorphism between u and the (non-empty) subset of a host type described by the predicate t. We require that definitions do not overlap.

For HOL with ad-hoc overloading, the study of its meta-logical properties is complex. It is a long story to prove that every definitional theory has a model [3, 4], and thereby proving the framework *consistent*, i. e. False is not derivable. Consistency is a weaker property than (model-theoretic) conservativity.

Studying model-theoretic conservativity in this setting, we face the challenge that signature extensions are distinct from theory extensions: any symbol may be referred to prior to its definition. For instance, consider a model $\mathscr{M}$ of a definitional theory $D = \{c_\alpha \equiv d_\alpha\}$ that interprets c_{bool} as true. Clearly, $\mathscr{M}$ interprets c_{bool} and d_{bool} equally. The theory extension $D \cup \{d_{\mathsf{bool}} \equiv \mathsf{False}\}$ cannot have $\mathscr{M}$ as a model. In our earlier work [1] we solved this challenge by tracking definitional *dependencies* between symbols. For instance, in the previous example, the definition $c_\alpha \equiv d_\alpha$ entails that c_{bool} depends on d_{bool}. We showed that HOL with ad-hoc overloading satisfies a notion of model-theoretic conservativity where the interpretation of these dependencies is permitted to change through theory extension.

Our work used the dependency relation that Kunčar and Popescu [3] introduced to identify cyclic dependencies, and also their eager ground semantics for HOL. Recently, Åman Pohjola and Gengelbach [4], while mechanising the consistency argument of Kunčar and Popescu, noticed issues both with this dependency relation and the eager semantics. First, the dependency relation of Kunčar and Popescu erroneously did not track dependencies of function types on their arguments. Åman Pohjola and Gengelbach address this by introducing, for any type constructor k, dependencies of $(\alpha_1, \ldots, \alpha_n)k$ on $\alpha_1, \ldots, \alpha_n$. Second, in eager ground semantics a formula holds if all of its ground instances evaluate to true. Because type variables are instantiated before term variables, eager ground semantics inadvertently constrains the interpretation of distinct term variables that have the same name, such as x_α and x_{bool}. Åman Pohjola

Submitted to:
LFMTP 2020

and Gengelbach address this by introducing *lazy* ground semantics, where type instantiation is delayed.

These solutions also affect our earlier proofs of model-theoretic conservativity, and add further proof obligations, e. g. to cover the larger dependency relation. We distinguish between built-in and user-defined symbols. The former have fixed semantics and contain, e. g. the function type constructor (of arity 2), and the type constructor bool (of arity 0). A *(signature) fragment* consists of a set of user-defined symbols such that the type of each constant in the fragment can be constructed by repeated application of built-in type constructors to types in the fragment. Given a definitional theory and a user-defined symbol u, the *u-independent fragment* F_u is defined as the set of all user-defined symbols that do not depend on any type instance of u. The u-independent fragment plays a key role in our notion of model-theoretic conservativity for HOL with ad-hoc overloading [1].

One new result is that the u-independent fragment indeed is a fragment w. r. t. the enlarged dependency relation. The proof needs to carefully establish dependencies through the larger relation to rectify the earlier use of an erroneous lemma. We have mechanised this result in HOL4 on top of the formalisation by Åman Pohjola and Gengelbach, and have proven several properties of the independent fragment.[1]

Next, we show model-theoretic conservativity for the lazy ground semantics. The main result of our earlier work [1, Theorem 3.3] mentions the u-independent fragment F_u, its types Type^{F_u} and its terms Term^{F_u}, but we now understand the notion of model $\mathscr{M} \models D$ w. r. t. the lazy semantics.

Theorem 3.3 (Model-theoretic Conservativity)**.** *Let $\mathscr{M}$ be a model of a well-formed definitional theory D, i. e. $\mathscr{M} \models D$. Moreover, let $D' := D \cup \{u \equiv t\}$ be a well-formed extension of D. There exists a model $\mathscr{M}'$ of the extended theory D' with the following property: the models $\mathscr{M}$ and $\mathscr{M}'$ agree on the interpretation of all types and terms in $\mathsf{Type}^{F_u} \cup \mathsf{Term}^{F_u}$.*

We maintain its earlier proof idea: given $\mathscr{M}$, a model $\mathscr{M}'$ for the extended theory is constructed by well-founded recursion over the (now enlarged) dependency relation, with symbols in the u-independent fragment as base case. The interpretation of symbols that depend on any type instance of u is changed, to ensure validity of all definitions. We adjust our earlier proof to the different semantics, and treat type isomorphisms specially. In contrast, both Kunčar and Popescu [3] and Åman Pohjola and Gengelbach [4] construct a specific model for a definitional theory by well-founded recursion over the full dependency relation. We are currently working on a mechanisation of Theorem 3.3 in HOL4.

The syntactic counterpart of model-theoretic conservativity is called *proof-theoretic conservativity* [2]. A future mechanisation should be possible in the same framework, although differences to the syntactic counterpart of our given model-theoretic conservativity remain to be studied. For any proposition from a symbol-independent fragment, if the proposition is derivable from a theory, is it also derivable from the theory without the symbol's definition?

References

[1] Arve Gengelbach & Tjark Weber (2017): *Model-Theoretic Conservative Extension for Definitional Theories*. In: *LSFA 2017*, *ENTCS* 338, Elsevier, pp. 133–145, doi:10.1016/j.entcs.2018.10.009.

[2] Ondrej Kuncar & Andrei Popescu (2018): *Safety and conservativity of definitions in HOL and Isabelle/HOL*. *Proc. ACM Program. Lang.* 2(POPL), pp. 24:1–24:26, doi:10.1145/3158112.

[3] Ondrej Kuncar & Andrei Popescu (2019): *A Consistent Foundation for Isabelle/HOL*. *J. Autom. Reasoning* 62(4), pp. 531–555, doi:10.1007/s10817-018-9454-8.

[4] Johannes Åman Pohjola & Arve Gengelbach (2020): *A Mechanised Semantics for HOL with Ad-hoc Overloading*. *CoRR* abs/2002.10212. Available at `https://arxiv.org/abs/2002.10212`. Accepted at LPAR 2020.

[1] Our mechanisation is available at `http://user.it.uu.se/~arvge836/lfmtp2020.html`.