

# Secondary Publication



Kirchner, Daniel

## Computer-Verified Foundations of Metaphysics

Date of secondary publication: 24.01.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-1060980

### Primary publication

Kirchner, Daniel (2024): Computer-Verified Foundations of Metaphysics, in: Künstliche Intelligenz : KI ; Forschung, Entwicklung, Erfahrungen, Berlin: Springer, Vol. 38, Nr. 1–2, pp. 95–98, doi: 10.1007/s13218-024-00834-z.

### Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>



# Computer-Verified Foundations of Metaphysics

Daniel Kirchner<sup>1</sup>

Received: 17 June 2023 / Accepted: 5 January 2024 / Published online: 10 February 2024  
© The Author(s) 2024

## 1 Introduction

We report on recent successes in the application of computational methods and automated reasoning techniques to a foundational theory of metaphysics: in [13], we utilize and extend the method of *shallow semantic embeddings* (SSEs) in classical higher-order logic (HOL) to construct a custom theorem proving environment for *abstract object theory* (AOT) on the basis of Isabelle/HOL.

Isabelle/HOL [17] is the most developed object logic in the generic proof assistant Isabelle and implements a version of classical higher-order logic. It provides access to sophisticated automation tools like *sledgehammer* [20] and model finding tools like *nitpick* [9]. These tools in turn interface with first-order provers like E [22], SPASS [24] and Vampire [21], SMT solvers like veriT [10] and Z3 [16], as well as higher-order provers like Zipperposition [23].

SSEs are a means for universal logical reasoning by translating a target logic to HOL using a representation of its semantics (see [1]). They form a cornerstone of the more recent development of the LogiKEY framework that focuses on the application of the method to ethical reasoning, normative theories and deontic logics [3, 5].

The target theory of our work, AOT, is a foundational metaphysical theory, developed by Edward Zalta, that explains the objects presupposed by the sciences as *abstract objects* that reify property patterns. The most recent canonical presentation of AOT can be found in a regularly updated online monograph called *Principia Logico-Metaphysica* (PLM) [33]. The theory is also the subject of two books [28, 29] and various scientific articles.

Our work in [13] focuses in particular on AOT's aspiration to provide a philosophically grounded basis for the construction and analysis of the objects of mathematics.

We can support this claim by verifying Uri Nodelman's and Edward Zalta's reconstruction of Frege's theorem: we confirm that the Dedekind-Peano postulates for natural numbers are consistently derivable in AOT using Frege's method. Furthermore, we can suggest and discuss generalizations and variants of the construction and can thereby provide theoretical insights into, and contribute to the philosophical justification of, the construction.

In the process, we can demonstrate that our method allows for a nearly transparent exchange of results between traditional pen-and-paper-based reasoning and the computerized implementation, which in turn can retain the automation mechanisms available for our meta-logic Isabelle/HOL.

During our work, we could significantly contribute to the evolution of our target theory itself, while simultaneously solving the technical challenge of using an SSE to implement a theory that is based on logical foundations that significantly differ from the meta-logic HOL.

In general, our results demonstrate the fruitfulness of the practice of Computational Metaphysics, i.e. the application of computational methods to metaphysical questions and theories (see [14]). Furthermore, our method allows for the transfer of existing technology for automated reasoning to new domains and for the computer-aided experimental investigation of variations and different axiomatizations of complex foundational logical theories with large bodies of theorems.

## 2 Challenges and Contributions

Prior to our work, SSEs were primarily applied to isolated logical arguments [4, 7, 11] and logics that are closely related to the meta-logic HOL like higher-order modal logic [2, 6, 8]. In contrast, our target, AOT, is a full logical theory that positions *itself* as a foundational framework and involves hyperintensional relations, a free logic governing its complex terms, and an additional mode of predication besides classical exemplification called *encoding*. Furthermore, it is based on a *relational type theory* that was previously argued

✉ Daniel Kirchner  
daniel.kirchner@uni-bamberg.de

<sup>1</sup> AI Systems Engineering, Otto-Friedrich-Universität Bamberg, Bamberg, Germany

to be incompatible with, and not representable in, systems based on functional type theory like Isabelle/HOL (see [19]).

In successfully implementing AOT, our work shows that the SSE approach scales to challenging logics, even if they extend classical logic significantly, and that the approach can be used to represent full theories by faithfully reconstructing their axioms and deductive system. We demonstrate that not only the logical structure of a theory can be technically implemented, but also that its syntax and reasoning flow can be preserved. To that end, we utilize Isabelle's mechanisms for constructing custom inner syntax and, further, by extending Isabelle's outer syntax with customized commands that allow a concise reproduction of the meta-theoretical mechanisms employed in the presentation of the target theory. This includes, for example, specialized commands for introducing new definitions and restricted variables, for which AOT imposes specific inferential roles and restrictions that otherwise could not be concisely captured by the existing infrastructure Isabelle provides.<sup>1</sup>

As a result, we arrive at a customized theorem proving environment for AOT that guarantees to only accept valid reasoning according to AOT's own deductive system and, in particular, avoids artifactual theorems. This is achieved by constructing a so-called *abstraction layer* that restricts automation to only rely on the reconstruction of the axioms and derivation rules of AOT, while barring the derivation of theorems that are mere artifacts of the underlying model construction. This enables the safe use of tools like *sledgehammer* for proving theorems of our target theory.

At the same time, we use a semantic representation in HOL as back-end, that is based on a newly developed extension of the partial, set-theoretic Aczel models (see [30]) of AOT. This allows for the experimental exploration of variations and extensions of the theory and its axiom system, while guaranteeing consistency relative to the meta-logic HOL.

This setup allowed us to significantly contribute to the evolution of our target theory, including the extension of its free logic, that was previously restricted to its individual terms, to relation terms, and the refinement of its axiom system, leading to an improved understanding of the conditions under which relations exist in the theory, which enabled new theoretical results in the construction of natural numbers.

By implementing Nodelman and Zalta's construction of natural numbers as given in chapter 14 of [26] and verifying their derivation of the Dedekind-Peano postulates, we

demonstrate that our method allows one to capture even complex constructions and reasoning in the target system.

The presentation of our work in [13] utilizes Isabelle's document generation mechanism (see [25]), and thereby ensures that formal statements (apart from a few exceptions that are explicitly marked) are verified against the computerized implementation.

The technical implementation used as a basis for the thesis is preserved as an entry in the *archive of formal proofs* [12], which ensures that it will remain up-to-date with future changes to Isabelle/HOL and allows it to continue to be used as a basis for future explorations of the target theory.

### 3 Implications and Future Work

Our work supports the use of SSEs as a powerful mechanism for universal logical reasoning that enables the application of existing automation infrastructure to new formalisms. We show that this remains feasible even if the target theory is based on logical foundations that significantly differ from the used meta-logic. Our method is general and can be applied to a multitude of logical systems, paving the way for future computational explorations of formalisms that so far may not have received the same attention in the development of automation infrastructure as classical approaches.

Furthermore, we strengthen the position of our target theory AOT as an expressive foundational theory that may serve as a philosophically grounded and logically structured basis for knowledge representation.

In this context, an interesting open question is the feasibility of an implementation of the generalized typed version of object theory, which not only involves abstract objects, but also abstract relations of arbitrary order and arity. This generalization is of particular interest, since it allows one to represent and analyze existing independent domain theories, like mathematical theories as ZF, themselves as abstract objects in a shared reasoning framework (see [15, 31, 32] and chapter 15 of PLM [33]).<sup>2</sup>

However, since the relative strength of this typed version of object theory may exceed that of ZF,<sup>3</sup> and thereby that of HOL, a provably sound implementation using an SSE is likely to require the use of a stronger extension of Isabelle/HOL, as for example HOLZF [18] or beyond, as a basis and so far remains an unsolved challenge for future research.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

<sup>1</sup> For example, AOT involves so-called *definitions by equivalence* that introduce new hyperintensional entities that the system requires to be necessarily equivalent, but not identical to the definiens. For similar reasons, we introduce custom proving infrastructure like a specialized substitution method that accounts for the hyperintensionality of the theory.

<sup>2</sup> The chapter numbering may change in future revisions of PLM, while the version at the time of writing will remain available as [27].

<sup>3</sup> Even without the *importation* of ZF as a domain theory.

**Data availability** The data that support the findings of this study are openly available in the Archive of Formal Proofs at <https://www.isa-afp.org/entries/AOT.html>.

## Declarations

**Conflict of interest** Christoph Benz Müller was the thesis supervisor of [13] and the author is part of his working group at Otto-Friedrich-Universität Bamberg.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Benz Müller C (2019) Universal (Meta-)Logical Reasoning: Recent Successes. *Sci Comp Program* 172:48–62. <https://doi.org/10.1016/j.scico.2018.10.008>
- Benz Müller C, Claus M, Sultana N (2015) Systematic verification of the modal logic cube in Isabelle/HOL. In: C. Kaliszky, A. Paskevich (eds.) *PxTP 2015*, vol. 186, pp. 27–41. EPTCS, Berlin, Germany. <https://doi.org/10.4204/EPTCS.186.5>
- Benz Müller C, Farjami A, Fuenmayor D, Meder P, Parent X, Steen A, van der Torre L, Zahoransky V (2020) LogiKEY workbench: Deontic logics, logic combinations and expressive ethical and legal reasoning (Isabelle/HOL dataset). *Data Brief* 33:106409. <https://doi.org/10.1016/j.dib.2020.106409>
- Benz Müller C, Fuenmayor D (2020) Computer-supported analysis of positive properties, ultrafilters and modal collapse in variants of Gödel's ontological argument. *Bulletin of the Section of Logic* 49(2), 127–148. <https://doi.org/10.18778/0138-0680.2020.08>
- Benz Müller C, Parent X, van der Torre L (2020) Designing normative theories for ethical and legal reasoning: Logikey framework, methodology, and tool support. *Artif Intell* 287:103348. <https://doi.org/10.1016/j.artint.2020.103348>
- Benz Müller C, Paulson L (2013) Quantified multimodal logics in simple type theory. *Logica Universalis* 7(1):7–20. <https://doi.org/10.1007/s11787-012-0052-y>
- Benz Müller C, Woltzenlogel Paleo B (2014) Automating Gödel's ontological proof of God's existence with higher-order automated theorem provers. In: T. Schaub, G. Friedrich, B. O'Sullivan (eds.) *ECAI 2014, Frontiers in Artificial Intelligence and Applications*, vol. 263, pp. 93–98. IOS Press. <https://doi.org/10.3233/978-1-61499-419-0-93>
- Benz Müller C, Woltzenlogel Paleo B (2015). Higher-Order Modal Logics: Automation and Applications. [https://doi.org/10.1007/978-3-319-21768-0\\_2](https://doi.org/10.1007/978-3-319-21768-0_2)
- Blanchette JC, Nipkow T (2010) Nitpick: A counterexample generator for higher-order logic based on a relational model finder. In: M. Kaufmann, L.C. Paulson (eds.) *Interactive Theorem Proving*, pp. 131–146. Springer Berlin Heidelberg, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-14052-5\\_11](https://doi.org/10.1007/978-3-642-14052-5_11)
- Bouton T, Caminha B. de Oliveira D, Déharbe D, Fontaine P (2009) veriT: An Open, Trustable and Efficient SMT-Solver. In: R.A. Schmidt (ed.) *Automated Deduction – CADE-22*, pp. 151–156. Springer Berlin Heidelberg, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-02959-2\\_12](https://doi.org/10.1007/978-3-642-02959-2_12)
- Fuenmayor D, Benz Müller C (2020) A case study on computational hermeneutics: E. J. Lowe's modal ontological argument. In: R. Silvestre, B. Göcke, J.Y. Béziau, P. Bilimoria (eds.) *Beyond Faith and Rationality: Essays on Logic, Religion and Philosophy, Sophia Studies in Cross-cultural Philosophy of Traditions and Cultures*, vol. 34, chap. 12. Springer Nature Switzerland AG. [https://doi.org/10.1007/978-3-030-43535-6\\_12](https://doi.org/10.1007/978-3-030-43535-6_12)
- Kirchner D (2022) Abstract object theory. *Archive of Formal Proofs*. <https://isa-afp.org/entries/AOT.html>, Formal proof development
- Kirchner D (2022) Computer-Verified Foundations of Metaphysics and an Ontology of Natural Numbers in Isabelle/HOL. Dissertation. <https://doi.org/10.17169/refubium-35141>
- Kirchner D, Benz Müller C, Zalta EN (2019) Computer science and metaphysics: A cross-fertilization. *Open Philosophy* 2(1):230–251. <https://doi.org/10.1515/opphil-2019-0015>
- Leitgeb H, Nodelman U, Zalta EN. A Defense of Logicism. <http://mally.stanford.edu/papers/logicism.pdf>. In preparation
- de Moura L, Bjørner N (2008) Z3: An Efficient SMT Solver. In: C.R. Ramakrishnan, J. Rehof (eds.) *Tools and Algorithms for the Construction and Analysis of Systems*, pp. 337–340. Springer Berlin Heidelberg, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-78800-3\\_24](https://doi.org/10.1007/978-3-540-78800-3_24)
- Nipkow T, Paulson LC, Wenzel M (2002) Isabelle/HOL — A Proof Assistant for Higher-Order Logic, *LNCS*, vol. 2283. Springer. <https://doi.org/10.1007/3-540-45949-9>
- Obua S (2006) Partizan Games in Isabelle/HOLZF. In: K. Barkaoui, A. Cavalcanti, A. Cerone (eds.) *Theoretical Aspects of Computing - ICTAC 2006*, pp. 272–286. Springer Berlin Heidelberg, Berlin, Heidelberg. [https://doi.org/10.1007/11921240\\_19](https://doi.org/10.1007/11921240_19)
- Oppenheimer PE, Zalta EN (2011) Relations versus functions at the foundations of logic: Type-theoretic considerations. *J Logic Comput* 21:351–374. <https://doi.org/10.1093/logcom/exq017>
- Paulson L (2012) Three Years of Experience with Sledgehammer, a Practical Link between Automatic and Interactive Theorem Provers. In: R.A. Schmidt, S. Schulz, B. Konev (eds.) *PAAR-2010: Proceedings of the 2nd Workshop on Practical Aspects of Automated Reasoning, EPiC Series in Computing*, vol. 9, pp. 1–10. EasyChair. <https://doi.org/10.29007/tnfd>
- Riazanov A, Voronkov A (2002) The design and implementation of VAMPIRE. *AI Commun.* 15:91–110
- Schulz S, Cruanes S, Vukmirović P (2019) Faster, higher, stronger: E 2.3. In: P. Fontaine (ed.) *Proc. of the 27th CADE, Natal, Brasil*, no. 11716 in *LNAI*, pp. 495–507. Springer. [https://doi.org/10.1007/978-3-030-29436-6\\_29](https://doi.org/10.1007/978-3-030-29436-6_29)
- Vukmirović P, Bentkamp A, Blanchette J, Cruanes S, Nummelin V, Tournet S (2021) Making Higher-Order Superposition Work. In: Platzer A, Sutcliffe G (eds) *Automated Deduction - CADE 28*. Springer International Publishing, Cham, pp 415–432
- Weidenbach C (2001) Combining Superposition, Sorts and Splitting. In: A. Robinson, A. Voronkov (eds.) *Handbook of Automated Reasoning, Handbook of Automated Reasoning*, pp. 1965–2013. North-Holland, Amsterdam. <https://doi.org/10.1016/B978-044450813-3/50029-1>
- Wenzel M. The Isabelle system manual. <https://isabelle.in.tum.de/website-Isabelle2021-1/dist/Isabelle2021-1/doc/system.pdf>
- Zalta EN. *Principia Logico-Metaphysica*. <http://mally.stanford.edu/principia-2021-10-13.pdf>. [Draft/Excerpt; dated October 13, 2021]

27. Zalta EN. Principia Logico-Metaphysica. <http://mally.stanford.edu/principia-2023-06-04.pdf>. [Draft/Excerpt; dated June 4, 2023]
28. Zalta EN (1983) Abstract Objects: An Introduction to Axiomatic Metaphysics. Synthese Library Springer. <https://doi.org/10.1007/978-94-009-6980-3>
29. Zalta EN (1988) Intensional Logic and the Metaphysics of Intentionality. MIT Press, A Bradford book
30. Zalta EN (1999) Natural Numbers and Natural Cardinals as Abstract Objects: A Partial Reconstruction of Frege's Grundgesetze in Object Theory. J Philos Logic 28(6):619–660. <https://doi.org/10.1023/A:1004330128910>
31. Zalta EN (2000) Neo-Logicism? An Ontological Reduction of Mathematics to Metaphysics. Erkenntnis 53(1):219–265. <https://doi.org/10.1023/A:1005614102033>
32. Zalta EN (2023) Mathematical Pluralism. Noûs. <https://doi.org/10.1111/nous.12451>
33. Zalta EN (2023) Principia Logico-Metaphysica. <http://mally.stanford.edu/principia.pdf>. [Draft/Excerpt; continuously updated current version]