



Searching for Wormholes in Proof-Space

David Fuenmayor¹² and Christoph Benzmüller¹²

¹ Otto-Friedrich-Universität Bamberg, Germany

² Freie Universität Berlin, Germany

1 Motivation

The present ‘position paper’ for presentation at the IWIL-2023 workshop aims at further deepening the discussion initiated in [BK01] and recently expanded in [BFSS23], regarding the intrinsic advantages of reasoning in higher-order logics and the role of ‘cut’ lemmata and definitions in finding ‘too-long-for-cut-avoiding-first-order-reasoners’ proofs. The leitmotif of this work can thus be suggestively summarized as:

definitions and cut-lemmata can behave like wormholes in proof-space,

where the proof-space analogy has been charmingly illustrated in the foreword (by Dolph Ulrich) of the book “Automated Reasoning and the Discovery of Missing and Elegant Proofs” by Larry Wos & Gail W. Pieper [WP03]. We quote:

“The overriding difficulty met at every turn was the unimaginably vast size of the space of proofs, a space in which all proofs solving a particular problem at hand might well be as unreachable as the farthest stars in the most distant galaxies. Consideration of quite short proofs suffices to illustrate this combinatorial explosion: even for systems of logic of the sort studied in this book that have just one axiom, for instance, there can be more 10-step proofs than kilometers in a light year, more 15-step proofs than stars in a trillion Milky Ways.”

In previous work [BFSS23] a thought experiment was presented as a motivating example: Folbert and Holly (waiting at the gates of heaven) become engaged in a theorem proving contest in which they have to pose first-order proof problems to each other, and the one whose ATP solves the given problem the faster will be admitted to heaven. Folbert goes for first-order (FO) ATPs and Holly for higher-order ATPs. We quote:

“Key to Holly’s advantage are the (hyper-)exponentially shorter proofs that are possible as one moves up the ladder of expressiveness from first-order logic to second-order logic, to third-order logic, and so on [Gö36]. The fact that the proof problems are stated in FO logic does not matter. When stating the same problem in the same FO way but in higher-order logic, much shorter proofs are possible, some of which might even be (hyper-)exponentially shorter than the proofs that can be found with comparatively inexpressive FO ATPs. A very prominent example of such a short proof is that of Boolos’ Curious Inference [Boo87].”

In this presentation we will discuss three examples of what we see as ‘wormholes’ in proof-space. The first one draws upon *Boolos’ Curious Inference* [Boo87], as discussed in [BFSS23], and has a motivational character; it serves to illustrate the astronomical magnitudes involved. The second one concerns the area of *correspondence theory* for modal logics. A third one, which is still work-in-progress, concerns applications in formalized mathematics, and involves the notion of compactness in topology (and to some extent also in logic).

2 Case Studies

We encode and analyze our case studies using the proof assistant Isabelle/HOL [NPW02], which we employ mainly as a frontend (with sophisticated editor/syntax capabilities) in order to invoke concurrently many different first- and higher-order backend ATPs, via the integrated meta-prover Sledgehammer [BKPU16]. The encoding has been carried out using ‘vanilla’ simple type theory, avoiding library definitions and lemmata as much as possible (theory ‘Main’ is imported for technical reasons). No Isabelle-specific extensions (locales, type classes, etc.) have been employed. The idea is that our proof problems shall remain self-contained and as close as possible to THF syntax [SB10], so our experiments can be easily replicated directly on the concerned provers (as done in [BFSS23]).

The corresponding Isabelle/HOL source files can be consulted under:¹

<https://github.com/davfuenmayor/IWIL-2023/tree/main/sources>

Boolos’ Curious Inference (BCI) We encode the original BCI problem (file `BCI.thy`) employing algebraic notions and discuss the role of cut-definitions in enabling fully automatic proofs. Subsequently, we present a natural algebraic generalization of BCI (file `BCIgen.thy`) and show how this can also be solved automatically by ATPs. We also show that this general ‘cut-lemma’ can be automatically employed by ATPs to solve BCI-like problems (e.g. involving fast-growing Ackermann-like functions).

Modal Correspondence Theory We present an algebraically-flavoured theory (cf. file `modal-correspondence.thy`) allowing for automatically proving correspondences between axiomatic conditions on relations and their respective modal operations (‘box’ and ‘diamond’). This theory is much less ambitious than others in the modal correspondence market (e.g. [CP12] and references therein) but is simple enough to nicely illustrate how ATPs can cleverly exploit cut-lemmata to reason in modal and non-classical contexts.

Compactness in Topology We encode the notion of a compact set wrt. a topology (file `topology-compactness.thy`) employing definitions of different complexities. We then use ATPs to show under which circumstances those definitions become equivalent. In order to illustrate the prospects of introducing ATPs in formalized mathematics, we explore several ‘cut-lemmata’ that should allow us to prove, fully automatically, a common classroom example, namely that the continuous image of a compact set is compact.²

¹Because of space and time constraints we cannot discuss here the contents of these case studies. We encourage the reader to consult the provided source files which have been carefully commented and –so we believe– are of more pedagogic value as an interactive document than as L^AT_EX prose.

²This is, however, still work-in-progress (and part of a larger project, cf. [FS22]).

References

- [BFSS23] Christoph Benzmüller, David Fuenmayor, Alexander Steen, and Geoff Sutcliffe. Who Finds the Short Proof? *Logic Journal of the IGPL*, 01 2023. jzac082.
- [BK01] Christoph Benzmüller and Manfred Kerber. A lost proof. In *Proceedings of the IJCAR 2001 Workshop: Future Directions in Automated Reasoning*, pages 13–24, Siena, Italy, 2001.
- [BKPU16] Jasmin Christian Blanchette, Cezary Kaliszyk, Lawrence C Paulson, and Josef Urban. Hammering towards qed. *Journal of Formalized Reasoning*, 9(1):101–148, 2016.
- [Boo87] George Boolos. A curious inference. *J. Philos. Log.*, 16(1):1–12, 1987.
- [CP12] Willem Conradie and Alessandra Palmigiano. Algorithmic correspondence and canonicity for distributive modal logic. *Annals of Pure and Applied Logic*, 163(3):338–376, 2012.
- [FS22] David Fuenmayor and Fabián Fernando Serrano Suárez. Formalising basic topology for computational logic in simple type theory. In Kevin Buzzard and Temur Kutsia, editors, *Intelligent Computer Mathematics - 15th International Conference, CICM 2022, Tbilisi, Georgia, September 19-23, 2022, Proceedings*, volume 13467 of *Lecture Notes in Computer Science*, pages 56–74. Springer, 2022.
- [Gö36] Kurt Gödel. Über die Länge von Beweisen. In K. Menger, K. Gödel, and A. Wald, editors, *Ergebnisse eines mathematischen Kolloquiums: Heft x7: 1934-1935*, page 23–24. Franz Deuticke, Vienna, 1936.
- [NPW02] Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. *Isabelle/HOL - A Proof Assistant for Higher-Order Logic*, volume 2283 of *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, 2002.
- [SB10] Geoff Sutcliffe and Christoph Benzmüller. Automated reasoning in higher-order logic using the TPTP THF infrastructure. *Journal of Formalized Reasoning*, 3(1):1–27, 2010.
- [WP03] Larry Wos and Gail W Pieper. *Automated reasoning and the discovery of missing and elegant proofs*. Rinton Press, 2003.