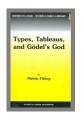
Formalization, Mechanization and Automation of Gödel's Proof of God's Existence

Christoph Benzmüller and Bruno Woltzenlogel Paleo

November 1, 2013



A gift to Priest Edvaldo and his church in Piracicaba, Brazil



Germany

- Telepolis & Heise
- Spiegel Online
- FA7
- Die Welt
- Berliner Morgenpost
- Hamburger Abendpost

- . . .

Austria

- Die Presse
- Wiener Zeitung
- ORF

Italy

- Repubblica
- Ilsussidario
- ٠...

India

- DNA India
- Delhi Daily News
- India Today
- . . .

US

- ABC News
- . . .

International

- Spiegel International
- Yahoo Finance
- United Press Intl.
- . .

Introduction — Quick answers to your most pressing questions!

SCIENCE NEWS

HOME / SCIENCE NEWS / RESEARCHERS SAY THEY USED MACBOOK TO PROVE GOEDEL'S GOD THEOREM

Researchers say they used MacBook to prove Goedel's God theorem

Oct. 23, 2013 | 8:14 PM | 1 comments

Are we in contact with Steve Jobs?

No

Do you really need a MacBook to obtain the results?

Nο

Is Apple sending us money?

No

(but maybe they should)

Introduction

Def: Ontological Argument/Proof

- * deductive argument
- * for the existence of god
- * starting from premises, which are justified by pure reasoning, i.e. they do not depend on observation in the world.

Existence of God: different types of arguments/proofs

 a posteriori (use experience/observation in the world
teleologicalcosmologicalmoral

 a priori (based on pure reasoning, independent)
— ontological argument
— definitional
modal

— other a priori arguments

Wohl eine jede Philosophie kreist um den ontologischen Gottesbeweis

(Adorno, Th. W.: Negative Dialektik. Frankfurt a. M. 1966, p.378)













Rich history on ontological arguments (pros and cons)



Anselm's notion of God:

"God is that, than which nothing greater can be conceived."

Gödel's notion of God:

"A God-like being possesses all 'positive' properties."

To show by logical reasoning:

"(Necessarily) God exists."

Different Interests in Ontological Arguments:

- Philosophical: Boundaries of Metaphysics & Epistemology
 - We talk about a metaphysical concept (God),
 - but we want to draw a conclusion for the real world.
 - Necessary Existence (NE): metaphysical NE vs. logical NE vs. modal NE
- Theistic: Successful argument should convince atheists.
- Our: Can computers (theorem provers) be used
 - to formalize the definitions and axioms?
 - to verify the arguments step-by-step?
 - to fully automate (sub-)arguments?

"Computer-assisted Theoretical Philosophy"

Introduction

Main challenge: No provers for Higher-order Modal Logic (HML)

Our solution: Embedding in Higher-order Classical Logic (HOL)

[BenzmüllerPaulson, Logica Universalis, 2013]

What we did (rough outline for remaining presentation!):

A: pen and paper: detailed natural deduction proof

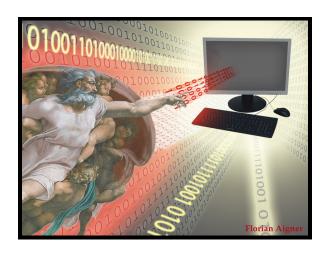
B: formalization: in classical higher-order logic (HOL) automation: theorem provers Leo-II and SATALLAX consistency: model finder NITPICK (NITROX)

C: step-by-step verification: proof assistant Coo

D: automation & verification: proof assistant Isabelle

Did we get new results?

Yes — let's discuss later!



Part A:
Informal Proof and Natural Deduction Proof

Gödel's Manuscript (1970)

ToDo: Improve Resolution

Pop) quipositive (e q	EP.)
At- 1	P(9) P(4) 5 P(904) Az	Pro V Pro
1	$G(x) = (\varphi) [P(\varphi) \ni \varphi(x)]$	
12	$\varphi E_{M,x} = (\psi) [\psi(x) J_M(y)] \varphi(y) J_M(y) = M(y) [\psi(x) J_M(y)] \psi(y) J_M(y) $	V(4) 17 (Freum)
PDNg	= N(pog) Neconit	2
At 2	P(p) > NP(p) } bec ~P(p) > N~P(p) } from 6(x) > 6 Em. x	anse it follows. The nature of the
TA.	G(X) > GEM.X	shorthat 1
Df.	F(x) = potion	meremon Frish
Ax 3	P(E)	merchany Erist
7%.	e(x) > N(38) e(d)	
h_{2a,c_0}	(3x) G(x) > N(2x) G(x)	
h	M(39) e(3) = (3) e(3)	M= pontlete.

```
M (7x) F(x) means all pos, prope is: com-
   patible This is the because of !
    A+4: P(4). 93, 4: > P(4) which imply
     the { X=X is possitive 
X=X is negative
     Dat if a yetem 5 of pers, peops, vice incom
      It would mean, that the Aum prop. A (which
     upositive) would be x +x
    Positive means positive in the moral acoker
  sense (in departly of the accidental structure of
  The arold ) On y the the at time , It is
  Old means "attendation at an opposed to privation
  (or contain y per vation) - This interprets graphen perol
    3/ 9 pm ac at (X) N 7 80x) ( OACHTS = (K) 2 x =
      hance x + x position port x = x is the contrary Ar-
X i.e the promot from in terms if eller play "contained
Member without negation
```

Versions

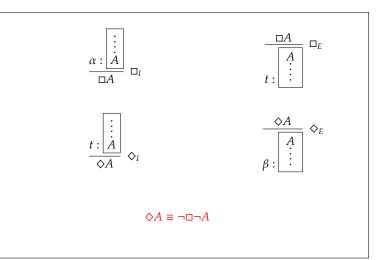
```
A1 Either a property is positive or its negation is (never both):
     \forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]
A2 A property necessarily implied by a positive property is
     positive:
                                   \forall \phi \forall \psi [(P(\phi) \land \Box \forall x [\phi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]
T1 Positive properties are possibly exemplified:
     \forall \varphi [P(\varphi) \rightarrow \Diamond \exists x \varphi(x)]
D1 A God-like being possesses all positive properties:
     G(x) \leftrightarrow \forall \phi [P(\phi) \to \phi(x)]
A3 The property of being God-like is positive:
                                                                                       P(G)
     Possibly, God exists:
                                                                                  \Diamond \exists x G(x)
A4 Positive properties are necessarily positive:
     \forall \phi [P(\phi) \rightarrow \Box P(\phi)]
D2 An essence of an individual is a property possessed by it and
     necessarily implying any of its properties:
     \phi \ ess \ x \leftrightarrow \phi(x) \land \forall \psi(\psi(x) \rightarrow \Box \forall \psi(\phi(y) \rightarrow \psi(y)))
T2 Being God-like is an essence of any God-like being:
     \forall x[G(x) \rightarrow G \ ess \ x]
D3 Necessary existence of an individual is
     the necessary exemplification of all its essences:
     E(x) \leftrightarrow \forall \phi [\phi \ ess \ x \rightarrow \Box \exists y \phi(y)]
A5 Necessary existence is a positive property:
                                                                                        P(E)
T3 Necessarily, God exists:
                                                                                  \Box \exists x G(x)
```

Proof Overview

D1:
$$G(x) \equiv \forall \varphi.[P(\varphi) \rightarrow \varphi(x)]$$

D2: $\varphi \ ess \ x \equiv \varphi(x) \land \forall \psi.(\psi(x) \rightarrow \Box \forall x.(\varphi(x) \rightarrow \psi(x)))$
D3: $E(x) \equiv \forall \varphi.[\varphi \ ess \ x \rightarrow \Box \exists y.\varphi(y)]$

Natural Deduction Calculus



Natural Deduction Proofs T1 and C1

$$\frac{\frac{\mathbf{A2}}{\forall \varphi. \forall \psi. [(P(\varphi) \land \Box \forall x. [\varphi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]}}{\forall \psi. [(P(\rho) \land \Box \forall x. [\rho(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]} \forall_{E}$$

$$\frac{P(\varphi) \land \Box \forall x. [\rho(x) \rightarrow \neg \rho(x)]) \rightarrow P(\neg \rho)}{(P(\varphi) \land \Box \forall x. [\neg \rho(x)]) \rightarrow P(\neg \rho)} \forall_{E}$$

$$\frac{P(\varphi) \land \Box \forall x. [\neg \rho(x)]) \rightarrow P(\neg \rho)}{(P(\varphi) \land \Box \forall x. [\neg \rho(x)]) \rightarrow \neg P(\varphi)} \forall_{E}$$

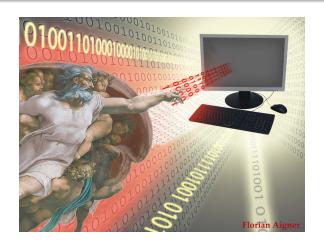
$$\frac{P(\varphi) \land \Box \forall x. [\neg \rho(x)]) \rightarrow \neg P(\varphi)}{(P(\varphi) \rightarrow \varphi \exists x. \varphi(x))} \forall_{E}$$

$$\frac{P(\varphi) \Rightarrow \varphi \exists x. \varphi(x)}{(\varphi, \varphi) \Rightarrow \varphi \Rightarrow \varphi} \forall_{E}$$

$$\frac{P(\varphi) \Rightarrow \varphi \Rightarrow \varphi \Rightarrow \varphi}{(\varphi, \varphi) \Rightarrow \varphi} \forall_{E}$$

$$\frac{\neg \varphi}{(\varphi, \varphi) \Rightarrow \varphi} \forall_{E}$$

Natural Deduction Proofs T2 (Partial)



Part B:

formalization: in classical higher-order logic (HOL) automation: theorem provers Leo-II and SATALLAX consistency checking: model finder NITPICK (NITROX)

Main challenge: No provers for Higher-order Modal Logic (HML)

Our solution: Embedding in *Higher-order Classical Logic* (HOL)

Then use existing HOL theorem provers for reasoning in HML

[BenzmüllerPaulson, Logica Universalis, 2013]

Previous empiricial findings:

Embedding of First-order Modal Logic in HOL works well
[BenzmüllerOttenRaths, ECAI, 2012]
[Benzmüller, LPAR, 2013]

$$\mathsf{HML} \quad \varphi, \psi \quad ::= \quad \dots \mid \neg \varphi \mid \varphi \land \psi \mid \varphi \rightarrow \psi \mid \Box \varphi \mid \Diamond \varphi \mid \forall x \varphi \mid \exists x \varphi \mid \forall P \varphi$$

Kripke style semantics (possible world semantics)

HOL

$$s,t ::= C |x| \lambda xs |st| \neg s |s \lor t| \forall x t$$

- meanwhile very well understood
- Henkin semantics vs. standard semantics
- various theorem provers do exists

```
interactive: Isabelle/HOL, HOL4, Hol Light, Coq/HOL, PVS, ... automated: TPS, LEO-II, Satallax, Nitpick, Isabelle/HOL, ...
```

$$\mathsf{HML} \quad \varphi, \psi \quad ::= \quad \dots \mid \neg \varphi \mid \varphi \land \psi \mid \varphi \rightarrow \psi \mid \Box \varphi \mid \Diamond \varphi \mid \forall x \varphi \mid \exists x \varphi \mid \forall P \varphi$$

HOL

$$s,t ::= C |x| \lambda xs |st| \neg s |s \lor t| \forall x t$$

HML in HOL: HML formulas φ are mapped to HOL predicates $\varphi_{\iota \to o}$

$$\begin{array}{lll} & = & \lambda \varphi_{l \to o} \lambda s_t \neg \varphi s \\ & \wedge & = & \lambda \varphi_{l \to o} \lambda \psi_{l \to o} \lambda s_t (\varphi s \wedge \psi s) \\ & \to & = & \lambda \varphi_{l \to o} \lambda \psi_{l \to o} \lambda s_t (\neg \varphi s \vee \psi s) \\ & \square & = & \lambda \varphi_{l \to o} \lambda s_t \forall u_t (\neg rsu \vee \varphi u) \\ & \diamondsuit & = & \lambda \varphi_{l \to o} \lambda s_t \exists u_t (rsu \wedge \varphi u) \\ & \forall & = & \lambda h_{\mu \to (\iota \to o)} \lambda s_t \forall d_\mu \, hds \\ & \exists & = & \lambda h_{\mu \to (\iota \to o)} \lambda s_t \exists d_\mu \, hds \\ & \forall & = & \lambda H_{(\mu \to (\iota \to o)) \to (\iota \to o)} \lambda s_t \forall d_\mu \, Hds \\ \\ & \text{valid} & = & \lambda \varphi_{l \to o} \forall w_t \varphi w \\ \end{array}$$

The equations in Ax are given as axioms to the HOL provers!

Example

```
HML formula HML formula in HOL expansion, \beta\eta-conversion expansion, \beta\eta-conversion expansion, \beta\eta-conversion
```

What are we doing?

In order to prove that φ is valid in HML,

-> we instead prove that valid $\varphi_{\iota \to o}$ can be derived from Ax in HOL.

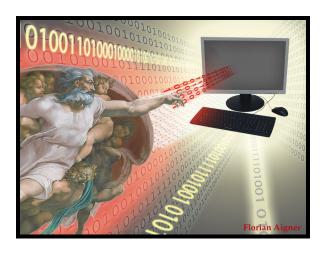
This can be done with interactive or automated HOL theorem provers.

Expansion: user or prover may flexibly choose expansion depth

Proof Automation and Consistencey Checking: Demo!

```
Terminal - bash - 125×32
MacBook-Chris %
MacBook-Chris %
MacBook-Chris % ./call_tptp.sh T3.p
Asking various HOL-ATPs in Miami remotely (thanks to Geoff Sutcliffe)
MacBook-Chris % agsyHOL---1.0 : T3.p ++++++ RESULT: SOT_7L4x_Y - agsyHOL---1.0 says Unknown - CPU = 0.00 WC = 0.02
LEO-II---1.6.0 : T3.p ++++++ RESULT: SOT_E4SCha - LEO-II---1.6.0 says Theorem - CPU = 0.03 WC = 0.09
Satallax---2.7 : T3.p ++++++ RESULT: SOT_kVZ1cB - Satallax---2.7 says Theorem - CPU = 0.00 WC = 0.14
Isabelle---2013: T3.p ++++++ RESULT: SOT_xa0aEp - Isabelle---2013 says Theorem - CPU = 14.06 WC = 17.73 SolvedBy = auto
TPS---3.12060151b : T3.p ++++++ RESULT: SOT ROEgsg - TPS---3.12060151b says Unknown - CPU = 33.56 WC = 41.57
Nitrox---2013 : T3.p ++++++ RESULT: S0T WGY1Tx - Nitrox---2013 says Unknown - CPU = 75.55 WC = 49.24
MacRook-Chris %
MacBook-Chris % ./call_tptp.sh Consistency.p
Asking various HOL-ATPs in Miami remotely (thanks to Geoff Sutcliffe)
MacBook-Chris % agsyHOL---1.0 : Consistency p ++++++ RESULT: SOT_ZtY_7o - agsyHOL---1.0 says Unknown - CPU = 0.00 WC = 0.00
Nitrox---2013 : Consistency.p ++++++ RESULT: SOT_HUZ10C - Nitrox---2013 says Satisfiable - CPU = 6.56 WC = 8.50
TPS---3.120601S1b : Consistency.p ++++++ RESULT: SOT_fpJxTM - TPS---3.120601S1b says Unknown - CPU = 43.00 WC = 49.42
Isabelle---2013 : Consistency.p ++++++ RESULT: SOT_6Tpp9i - Isabelle---2013 says Unknown - CPU = 69.96 WC = 72.62
LEO-II---1.6.0 : Consistency.p ++++++ RESULT: SOT_dY10si - LEO-II---1.6.0 says Timeout - CPU = 90 WC = 89.86
Satallax---2.7 : Consistency,p ++++++ RESULT: SOT_09WSLf - Satallax---2.7 says Timeout - CPU = 90 WC = 90.50
MacBook-Chris %
```

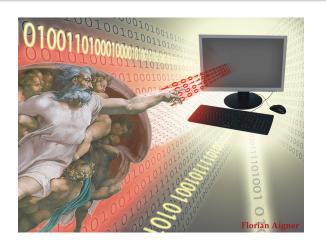
Provers are called remotely in Miami — no local installation needed!



Part C: Formalization and Verification in Coq

Coq Proof

- Goal: verification of the natural deduction proof
 - Step-by-step formalization
 - Almost no automation (intentionally!)
- Interesting facts to note:
 - Embedding is transparent to the user
 - Embedding gives labeled calculus for free



Part D: automation & verification: proof assistant Isabelle





Overview

Installation

Community
Site Mirrors:

Combridge (Luk)

Isabelle is a generic proof assistant. It allows mathematical formulas to be expressed in a formal language and provides tools for proving those formulas in a logical calculus. Isabelle is developed at University of Cambridge (Lany Paulson), Technische Universität München (Tobias Nipkon) and Universitä Paris-Sud (Makarius Winzel). See the Isabelle overvierge for a brief introduction.

Now available: Isabelle2013



Download for Linux - Download for Windows

Some highlights:

- . Improvements of Isabelle/Scala and Isabelle/iEdit Prover IDE.
- Advanced build tool based on Isabelle/Scala.
- Undated manuals: isar-ref, implementation, system.
- · Pure: improved support for block-structured specification contexts.
- HOL tool enhancements: Sledgehammer, Lifting, Quickcheck.
- HOL library enhancements: HOL-Library, HOL-Probability, HOL-Cardinals.
- HOL: New BNF-based (co)datatype package.
- Improved performance thanks to Poly/ML 5.5.0.

See also the cumulative NEWS.

Distribution & Support

Isabelle is distributed for free under the BSD license. It includes source and binary packages and documentation, see the detailed installation instructions. A vast collection of isabelle examples and applications is available from the Archive of Formal Proofs.

Support is available by ample documentation, the Isabelle community Wiki, and the following mailing lists:

- isabelle-users@cl.cam.ac.uk provides a forum for Isabelle users to discuss problems, exchange information, and make announcements. Users of official
- Isabelle releases should <u>subscribe</u> or see the <u>archive</u> (also available via <u>Google groups</u> and <u>Narkive</u>).

 <u>isabelle dev@in.tum.de</u> covers the isabelle development process, including intermediate repository versions, and administrative issues concerning the website or testing infrastructure. Early adopters of propository versions should subscribe or see the archive (also available at mail-archive.com or gmane.cm).

Last updated: 2013-03-09 12:21:39

Automation & Verification in Proof Assistant Isabelle/HOL

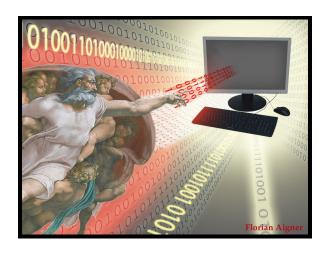
Isabelle

bla

What have we done

bla

See the handout (generated from the Isabelle source file).

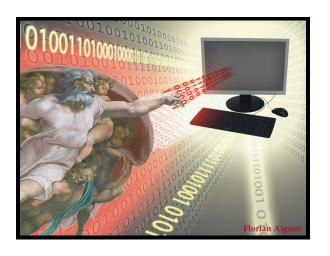


Part E: Criticisms

Criticisms S5

Criticisms Modal Collapse

Criticisms No Neutral Properties



Part F: Conclusions

Summary of Results about the Ontological Proof

- K sufficient for T1, C1 and T2
- S5 not needed for T3
- KB sufficient for T3
- A simpler new proof of C1
- Gödel's original axioms (without conjunct $\phi(x)$ in D2) are inconsistent
- Scott's axioms are consistent
- For T1, only half of A1 (A1a) is needed
- For T2, the other half (A1b) is needed

Summary of Results for Logic

- Infra-structure for reasoning with modal logic using existing proof assistants and higher-order automated theorem provers
- A new natural deduction calculus for higher-order modal logic

