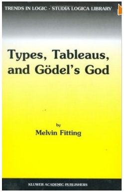
Formalization, Mechanization and Automation of Gödel's Proof of God's Existence

Christoph Benzmüller and **Bruno Woltzenlogel Paleo**

November 1, 2013



$$\dfrac{\dfrac{\text{Axiom 3}}{P(G)} \qquad \dfrac{\dfrac{\text{Theorem 1}}{\forall\varphi.[P(\varphi) \to \Diamond\exists x.\varphi(x)]}}{P(G) \to \Diamond\exists x.G(x)}\,\forall_E}{\Diamond\exists x.G(x)}\,\to_E$$

Introduction

SPIEGEL ONLINE WISSENSCHAFT

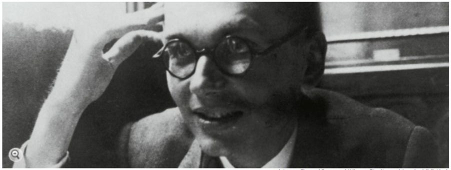
Login | Registrierung

Politik | Wirtschaft | Panorama | Sport | Kultur | Netzwelt | Wissenschaft | Gesundheit | einestages | Karriere | Uni | Schule | Reise | Auto

Nachrichten > Wissenschaft > Mensch > Mathematik > Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürter



picture-alliance/ Imagno/ Wiener Stadt- und Landesbibliothek

Kurt Gödel (um das Jahr 1935): Der Mathematiker hielt seinen Gottesbeweis jahrzehntelang geheim

Ein Wesen existiert, das alle positiven Eigenschaften in sich vereint. Das bewies der legendäre Mathematiker Kurt Gödel mit einem komplizierten Formelgebilde. Zwei Wissenschaftler haben diesen Gottesbeweis nun überprüft - und für gültig befunden.

Montag, 09.09.2013 – 12:03 Uhr

Drucken | Versenden | Merken

Jetzt sind die letzten Zweifel ausgeräumt: Gott existiert tatsächlich. Ein Computer hat es mit kalter Logik bewiesen - das MacBook des Computerwissenschaftlers Christoph Benzmüller von der Freien Universität Berlin.

Germany
- Telepolis & Heise
- Spiegel Online
- FAZ
- Die Welt
- Berliner Morgenpost
- Hamburger Abendpost
- ...

Austria
- Die Presse
- Wiener Zeitung
- ORF
- ...

Italy
- Repubblica
- Today.it
- Ilsussidario
- ...

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

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

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.

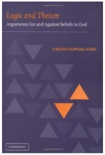
Existence of God: different types of arguments/proofs

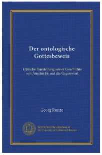
- a posteriori (use experience/observation in the world)
  - teleological
  - cosmological
  - moral
  - ...
- a priori (based on pure reasoning, independent)
  - ontological argument
    - definitional
    - modal
    - ...
  - other a priori arguments

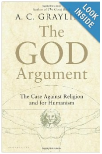
Wohl eine jede Philosophie kreist um den ontologischen Gottesbeweis

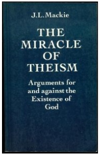
(Adorno, Th. W.: Negative Dialektik. Frankfurt a. M. 1966, p.378)

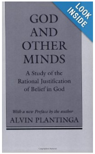
. . .

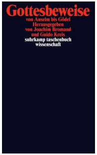












. . .

Rich history on ontological arguments (pros and cons)

… Anselm v. C. … Th. Aquinas …… Descartes Spinoza Leibniz … Hume Kant … Hegel … Frege … Hartshorne Malcolm Lewis Plantinga Gödel …

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"

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

Our idea: 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: axioms, defs, thms in HOL

proof automation: theorems provers Leo-II and Satallax

consistency: automatic verification with Nitpick

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

D: automation & verification: proof assistant Isabelle

Did we get new results? Yes — let's discuss later!

ToDo: Show Goedel's Manuscript

- 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: $\forall\phi\forall\psi[(P(\phi) \wedge \Box\forall x[\phi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]$
- 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: $G(x) \leftrightarrow \forall\phi[P(\phi) \rightarrow \phi(x)]$
- A3 The property of being God-like is positive: $P(G)$
- C 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: $\phi\ ess\ x \leftrightarrow \phi(x) \wedge \forall\psi(\psi(x) \rightarrow \Box\forall y(\phi(y) \rightarrow \psi(y)))$
- 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

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

$$\textbf{D2: } \varphi\ ess\ x \equiv \varphi(x) \land \forall\psi.(\psi(x) \rightarrow \Box\forall x.(\varphi(x) \rightarrow \psi(x)))$$

$$\textbf{D3: } E(x) \equiv \forall\varphi.[\varphi\ ess\ x \rightarrow \Box\exists y.\varphi(y)]$$

$$\dfrac{\dfrac{\textbf{A3}}{P(G)} \qquad \dfrac{\dfrac{\textbf{A2}}{\forall\varphi.\forall\psi.[(P(\varphi) \land \Box\forall x.[\varphi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]} \qquad \dfrac{\textbf{A1a}}{\forall\varphi.[P(\neg\varphi) \rightarrow \neg P(\varphi)]}}{\textbf{T1: } \forall\varphi.[P(\varphi) \rightarrow \Diamond\exists x.\varphi(x)]}}{\textbf{C1: } \Diamond\exists x.G(x)}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\textbf{A1b}}{\forall\varphi.[\neg P(\varphi) \rightarrow P(\neg\varphi)]} \qquad \dfrac{\textbf{A4}}{\forall\varphi.[P(\varphi) \rightarrow \Box\, P(\varphi)]}}{\textbf{T2: } \forall y.[G(y) \rightarrow G\ ess\ y]} \qquad \dfrac{\textbf{A5}}{P(E)}}{\textbf{L1: } \exists z.G(z) \rightarrow \Box\exists x.G(x)}}{\Diamond\exists z.G(z) \rightarrow \Diamond\Box\exists x.G(x)} \qquad \dfrac{\textbf{S5}}{\forall\xi.[\Diamond\Box\xi \rightarrow \Box\xi]}}{\textbf{L2: } \Diamond\exists z.G(z) \rightarrow \Box\exists x.G(x)}$$

$$\dfrac{\textbf{C1: } \Diamond\exists x.G(x) \qquad \textbf{L2: } \Diamond\exists z.G(z) \rightarrow \Box\exists x.G(x)}{\textbf{T3: }}$$

$$\dfrac{A \lor B \quad \begin{matrix}\overline{A}\\ \vdots \\ C\end{matrix} \quad \begin{matrix}\overline{B}\\ \vdots \\ C\end{matrix}}{C}\ \lor_E \qquad \dfrac{A \quad B}{A \land B}\ \land_I \qquad \dfrac{\begin{matrix}\overline{A}^{\,n}\\ \vdots \\ B\end{matrix}}{A \longrightarrow B}\ \longrightarrow_I^n$$

$$\dfrac{A}{A \lor B}\ \lor_{I_1} \qquad \dfrac{A \land B}{A}\ \land_{E_1} \qquad \dfrac{B}{A \longrightarrow B}\ \longrightarrow_I$$

$$\dfrac{B}{A \lor B}\ \lor_{I_2} \qquad \dfrac{A \land B}{B}\ \land_{E_2} \qquad \dfrac{A \quad A \longrightarrow B}{B}\ \longrightarrow_E$$

$$\dfrac{A[\alpha]}{\forall x.A[x]}\ \forall_I \qquad \dfrac{\forall x.A[x]}{A[t]}\ \forall_E \qquad \dfrac{A[t]}{\exists x.A[x]}\ \exists_I \qquad \dfrac{\exists x.A[x]}{A[\beta]}\ \exists_E$$

$$\neg A \equiv A \longrightarrow \bot \qquad \dfrac{\neg\neg A}{A}\ \neg\neg_E$$

$$
\dfrac{
  \dfrac{
    \dfrac{
      \dfrac{
        \dfrac{\textbf{A2}}{\forall\varphi.\forall\psi.[(P(\varphi) \wedge \Box\forall x.[\varphi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]}
      }{\forall\psi.[(P(\rho) \wedge \Box\forall x.[\rho(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]}\;\forall_E
    }{(P(\rho) \wedge \Box\forall x.[\rho(x) \rightarrow \neg\rho(x)]) \rightarrow P(\neg\rho)}\;\forall_E
  }{(P(\rho) \wedge \Box\forall x.[\neg\rho(x)]) \rightarrow P(\neg\rho)}
  \qquad
  \dfrac{
    \dfrac{\textbf{A1a}}{\forall\varphi.[P(\neg\varphi) \rightarrow \neg P(\varphi)]}
  }{P(\neg\rho) \rightarrow \neg P(\rho)}\;\forall_E
}{
  \dfrac{
    \dfrac{(P(\rho) \wedge \Box\forall x.[\neg\rho(x)]) \rightarrow \neg P(\rho)}{P(\rho) \rightarrow \Diamond\exists x.\rho(x)}
  }{\forall\varphi.[P(\varphi) \rightarrow \Diamond\exists x.\varphi(x)]}\;\forall_I
}
$$

$$
\dfrac{
  \dfrac{\textbf{A3}}{P(G)}
  \qquad
  \dfrac{
    \dfrac{\textbf{T1}}{\forall\varphi.[P(\varphi) \rightarrow \Diamond\exists x.\varphi(x)]}
  }{P(G) \rightarrow \Diamond\exists x.G(x)}\;\forall_E
}{\Diamond\exists x.G(x)}\;\rightarrow_E
$$

$$\frac{\alpha : \boxed{\begin{array}{c}\vdots\\ A\end{array}}}{\Box A}\ \Box_I \qquad\qquad \frac{\Box A}{t : \boxed{\begin{array}{c}A\\ \vdots\end{array}}}\ \Box_E$$

$$\frac{t : \boxed{\begin{array}{c}\vdots\\ A\end{array}}}{\Diamond A}\ \Diamond_I \qquad\qquad \frac{\Diamond A}{\beta : \boxed{\begin{array}{c}A\\ \vdots\end{array}}}\ \Diamond_E$$

$$\Diamond A \equiv \neg A$$

$$\dfrac{\dfrac{\dfrac{\psi(x)^6 \qquad \dfrac{\Pi_2}{\psi(x) \to \Box P(\psi)}}{\Box P(\psi)}\to_E \qquad \dfrac{\dfrac{\dfrac{\dfrac{\Box P(\psi)^7}{P(\psi)}\Box_E \qquad \dfrac{\Pi_3}{P(\psi) \to \forall x.(G(x) \to \psi(x))}}{\forall x.(G(x) \to \psi(x))}\to_E}{\Box\forall x.(G(x) \to \psi(x))}\Box_I}{\Box P(\psi) \to \Box\forall x.(G(x) \to \psi(x))}\to_I^7}{\Box\forall x.(G(x) \to \psi(x))}\to_E}{\psi(x) \to \Box\forall x.(G(x) \to \psi(x))}\to_I^6$$

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

Our idea: 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]

Theorem Provers for Higher-order Modal Logic?

HML $\varphi, \psi \quad ::= \quad \ldots \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid \Box\varphi \mid \Diamond\varphi \mid \forall x\, \varphi \mid \exists x\, \varphi \mid \forall P\, \varphi$

- Kripke style semantics (possible world semantics)

HOL $s, t \quad ::= \quad C \mid x \mid \lambda x s \mid s\, t \mid \neg s \mid s \vee t \mid \forall x\, t$

- 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, . . .

HML $\quad \varphi, \psi \quad ::= \quad \ldots \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid \Box\varphi \mid \Diamond\varphi \mid \forall x\,\varphi \mid \exists x\,\varphi \mid \forall P\,\varphi$

HOL $\quad s, t \quad ::= \quad C \mid x \mid \lambda xs \mid s\,t \mid \neg s \mid s \vee t \mid \forall x\,t$

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

$$\begin{array}{lcl}
\neg & = & \lambda\varphi_{\iota\to o}\lambda s_\iota \neg\varphi s \\
\wedge & = & \lambda\varphi_{\iota\to o}\lambda\psi_{\iota\to o}\lambda s_\iota(\varphi s \wedge \psi s) \\
\rightarrow & = & \lambda\varphi_{\iota\to o}\lambda\psi_{\iota\to o}\lambda s_\iota(\neg\varphi s \vee \psi s) \\
\Box & = & \lambda\varphi_{\iota\to o}\lambda s_\iota\forall u_\iota\,(\neg rsu \vee \varphi u) \\
\forall & = & \lambda h_{\mu\to(\iota\to o)}\lambda s_\iota\forall d_\mu\,hds \\
\exists & = & \lambda h_{\mu\to(\iota\to o)}\lambda s_\iota\forall d_\mu\,hds \\
\forall & = & \lambda H_{(\mu\to(\iota\to o))\to(\iota\to o)}\lambda s_\iota\forall d_\mu\,Hds \\
 & & \\
\text{valid} & = & \lambda\varphi_{\iota\to o}\forall s_\iota\varphi s
\end{array} \qquad \text{Ax}$$

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

Example

HML formula	$\Diamond \exists x G(x)$
HOL formula	$\text{valid}(\Diamond \exists x G(x))_{\iota \to o}$
expansion, $\beta\eta$-conversion	$\forall S_\iota (\Diamond \exists x G(x))_{\iota \to o}$

In order to prove that φ is valid in HML

we instead prove Ax $\vdash$ valid $\varphi_{\iota \to o}$ in HOL

- 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

Isabelle Proof

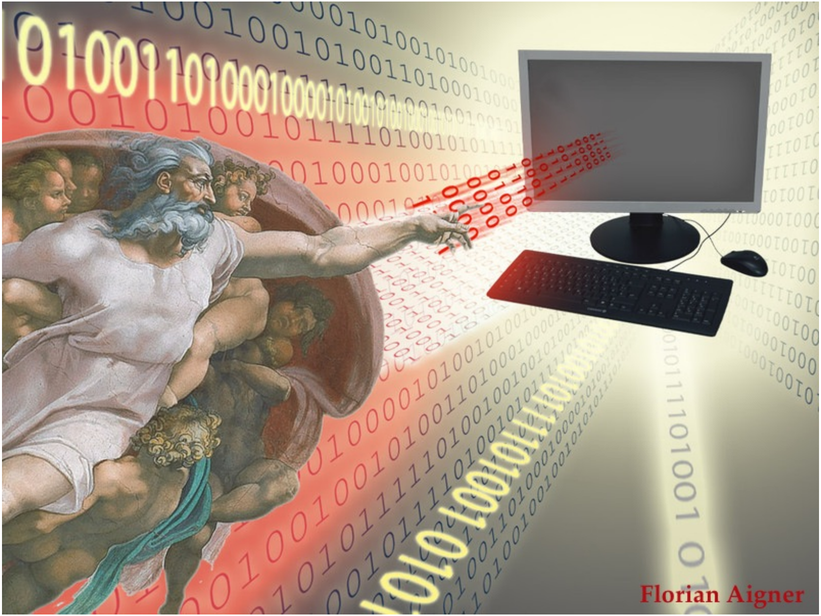
todo

Summary of Results

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

Conclusion

todo


Florian Aigner