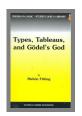
Formalization, Mechanization and Automation of Gödel's Proof of God's Existence

Christoph Benzmüller and Bruno Woltzenlogel Paleo

November 1, 2013



$$\underbrace{\frac{\text{Axiom } 3}{P(G)}}_{\text{$P(G)$}} \underbrace{\begin{array}{c} -\text{Theorem } 1 \\ \forall \varphi. [P(\varphi) \to \Diamond \exists x. \varphi(x)] \\ \hline P(G) \to \Diamond \exists x. G(x) \\ \\ \Diamond \exists x. G(x) \end{array}}_{\Rightarrow E} \forall_E$$

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.
- **4□ > 4周 > 4** = **>** 4 = **>** = **9** 9 9

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 Johs?

Vo.

Do you really need a MacBook to obtain the results?

Nο

Is Apple sending us money?

(but maybe they should)

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?

Nο

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 of 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		

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 of the world.

Existence of God: different types of arguments/proofs

tel	smological
—	
a prior	i (based on pure reasoning, independent)
on	tological argument
	 definitional
	– modal
	-
oth	ner 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."



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
- Ours: Can computers (theorem provers) be used . . .
 - ...to formalize the definitions, axioms and theorems?
 - ...to verify the arguments step-by-step?
 - ... to fully automate (sub-)arguments?

"Computer-assisted Theoretical Philosophy"



Introduction

Challenge: No provers for Higher-order Quantified Modal Logic (QML)

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 any new results?

Yes — let's discuss this later!



Part A:
Informal Proof and Natural Deduction Proof

Gödel's Manuscript: 1930's, 1941, 1946-1955, 1970

	Ontologischer Bereis	Feb 10, 1970
P(9)	19 is positive (is	9 EP.)
At. 1.2	191. P(Y) 5 P(9,4) At 2	Proto Pro
11	(4) - (4) P(9) 3 P(X)7	-3-00 / C . / .)
12	7 (4) [4 (1) 3/(3)](9(3) 3	Y(4)71 (Freum .)
p DN9	= N(pog) Neccosi	(y
At 2	P(q) > NP(q) } & & & & & & & & & & & & & & & & & &	cause it follows . The mature of The suspending
TA.	G(X) S G EM. X	supporty 1
Df. Ax3	E(x) = P[QEuxNJx q(x)] $P(E)$	mercesary Eristen
Th.	G(x) > N(71) G(x)	
h	" > N(39) e(3) " > N(39) e(3) " > N(39) e(3) (3*) e(4) > N(39) e(3)	M-pontheling
exclusive	or and for any munion of sum	nama ila

```
M (3x) F(x) means all pos prope is com-
   patible This is the because of:
    A+4: P(4). 92, 4: > P(4) which imple
     the SX=X is possitive in negative
    Dut if a yetem 5 of por. projo, veic incom
      It would mean, that the Aum prop. A (which
     uporitive) would be x + x
    Positive means positive in the moral acide
  sense (in departly of the accidental stynether of
  The avoild ). Only then the at time.
  also mean! "attribution at an opposed to privation
  (or contain y per vation) - This interpret for pla proof
   of a hundred (X) NABOX) CHANGE (X) 3 x+
      honce x + x positive por XEX of Techniq Ar-
X i.e. the formal forms in terms if eller plops "contained
Member without negation.
```

Scott's Version of Gödel's Axioms, Definitions and Theorems

$\mathbf{A1}$	Either a property or its negation is positive, but not	both: $\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$
$\mathbf{A2}$	A property necessarily implied	
	by a positive property is positive: $\forall \phi \forall$	$\psi[(P(\phi) \land \Box \forall x [\phi(x) \to \psi(x)]) \to P(\psi)]$
T1	Positive properties are possibly exemplified:	$\forall \varphi [P(\varphi) \to \Diamond \exists x \varphi(x)]$
$\mathbf{D}1$	A God-like being possesses all positive properties:	$G(x) \leftrightarrow \forall \phi [P(\phi) \to \phi(x)]$
$\mathbf{A3}$	The property of being God-like is positive:	P(G)
\mathbf{C}	Possibly, God exists:	$\Diamond \exists x G(x)$
$\mathbf{A4}$	Positive properties are necessarily positive:	$\forall \phi [P(\phi) \to \Box P(\phi)]$
D2	An essence of an individual is	
	a property possessed by it and	
	necessarily implying any of its properties: ϕ ess. $x \leftarrow$	$\phi(x) \land \forall \psi(\psi(x) \to \Box \forall y(\phi(y) \to \psi(y)))$
T2	Being God-like is an essence of any God-like being:	$\forall x[G(x) \to G \ ess. \ x]$
D3	Necessary existence of an individual is	
	the necessary exemplification of all its essences:	$NE(x) \leftrightarrow \forall \phi [\phi \ ess. \ x \to \Box \exists y \phi(y)]$
$\mathbf{A5}$	Necessary existence is a positive property:	P(NE)
T3	Necessarily, God exists:	$\Box \exists x G(x)$



C1: $\Diamond \exists z. G(z)$

C1:
$$\Diamond \exists z. G(z)$$
 L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$

L2:
$$\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$$

C1: $\Diamond \exists z. G(z)$ L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$

L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$

C1: $\Diamond \exists z. G(z)$ L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$



C1: $\Diamond \exists z. G(z)$ L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$

C1: $\Diamond \exists z. G(z)$ L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$ T3: $\Box \exists x. G(x)$

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

C1: $\Diamond \exists z. G(z)$ L2: $\Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)$

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

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$

$$\begin{array}{c|c} P(E) \\ \hline L1: \exists z.G(z) \rightarrow \Box \exists x.G(x) & \S5 \\ \hline & \Diamond \exists z.G(z) \rightarrow \Diamond \Box \exists x.G(x) & \forall \xi.[\Diamond \Box \xi \rightarrow \Box \xi] \\ \hline & L2: \Diamond \exists z.G(z) \rightarrow \Box \exists x.G(x) \\ \hline \hline C1: \Diamond \exists z.G(z) & L2: \Diamond \exists z.G(z) \rightarrow \Box \exists x.G(x) \\ \hline \hline & T3: \Box \exists x.G(x) \\ \hline \end{array}$$

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

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$
 (cheating!)

$$\begin{array}{c|c} P(E) \\ \hline & L1: \exists z.G(z) \rightarrow \Box \exists x.G(x) \\ \hline & \Diamond \exists z.G(z) \rightarrow \Diamond \Box \exists x.G(x) \\ \hline & L2: \Diamond \exists z.G(z) \rightarrow \Box \exists x.G(x) \\ \hline & & \\ \hline & \\ \hline & \\ \hline & & \\ \hline & & \\ \hline & & \\ \hline & \\ \hline & & \\ \hline & & \\ \hline & \\ \hline & & \\ \hline & \\ \hline & \\$$

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

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$
 D3: $E(x) \equiv \forall \varphi. [\varphi \ ess \ x \rightarrow \Box \exists y. \varphi(y)]$

$$\begin{array}{c|c} \textbf{T2:} \ \forall y. [G(y) \rightarrow G \ ess \ y] & P(E) \\ \hline & \underline{ \begin{array}{c|c} \textbf{L1:} \ \exists z. G(z) \rightarrow \Box \exists x. G(x) \\ \hline & \Diamond \exists z. G(z) \rightarrow \Diamond \Box \exists x. G(x) \\ \hline & \textbf{L2:} \ \Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x) \\ \hline & \underline{ \begin{array}{c|c} \textbf{C1:} \ \Diamond \exists z. G(z) \\ \hline & \textbf{L2:} \ \Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x) \\ \hline & \textbf{T3:} \ \Box \exists x. G(x) \\ \hline \end{array} }$$

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

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$
 D3: $E(x) \equiv \forall \varphi. [\varphi \ ess \ x \rightarrow \Box \exists y. \varphi(y)]$

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

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$
 D3: $E(x) \equiv \forall \varphi. [\varphi \ ess \ x \rightarrow \Box \exists y. \varphi(y)]$

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*:
$$E(x) \equiv \Box \exists y. G(y)$$
 D3: $E(x) \equiv \forall \varphi. [\varphi \ ess \ x \rightarrow \Box \exists y. \varphi(y)]$

$$\frac{A1b}{\forall \varphi. [\neg P(\varphi) \rightarrow P(\neg \varphi)]} \qquad \frac{A4}{\forall \varphi. [P(\varphi) \rightarrow \Box P(\varphi)]} \qquad A5$$

$$\underline{T2: \forall y. [G(y) \rightarrow G \ ess \ y]} \qquad P(E)$$

$$\underline{L1: \exists z. G(z) \rightarrow \Box \exists x. G(x)} \qquad \qquad \underbrace{S5}$$

$$\underline{\forall \xi. [\Diamond \Box \xi \rightarrow \Box \xi]}$$

$$\underline{L2: \Diamond \exists z. G(z) \rightarrow \Box \exists x. G(x)}$$

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

$$\underline{T3: \Box \exists x. G(x)}$$

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

D2: $\varphi \ ess \ x \equiv \varphi(x) \land \forall \psi. (\psi(x) \to \Box \forall x. (\varphi(x) \to \psi(x)))$

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

C1:
$$\Diamond \exists z. G(z)$$

D3*: $E(x) \equiv \Box \exists y. G(y)$

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

D2: φ ess $x \equiv \varphi(x) \land \forall \psi. (\psi(x) \to \Box \forall x. (\varphi(x) \to \psi(x)))$

$$\mathbf{D2} \cdot \mathbf{F}(x) = -\mathbf{D2} \cdot \mathbf{F}(x) + \mathbf{F}(x)$$

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$
 D3: $E(x) \equiv \forall \varphi. [\varphi \ ess \ x \rightarrow \Box \exists y. \varphi(y)]$

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

D2: φ ess $x \equiv \varphi(x) \land \forall \psi. (\psi(x) \to \Box \forall x. (\varphi(x) \to \psi(x)))$

D3*:
$$E(x) \equiv \Box \exists y. G(y)$$
 D3: $E(x) \equiv \forall \varphi. [\varphi \ ess \ x \rightarrow \Box \exists y. \varphi(y)]$

$$\begin{array}{c|c}
\hline{A3} \\
\hline{P(G)}
\end{array}$$

$$\begin{array}{c|c}
\hline{C1: \diamondsuit \exists z.G(z)}
\end{array}$$

$$\begin{array}{c|c}
\hline
A4 \\
\hline
\forall \varphi. [\neg P(\varphi) \to P(\neg \varphi)] & \forall \varphi. [P(\varphi) \to \Box P(\varphi)] \\
\hline
\hline
\hline
12: \forall y. [G(y) \to G ess y] & P(E)
\end{array}$$

$$\begin{array}{c|c}
\hline
L1: \exists z.G(z) \to \Box \exists x.G(x) \\
\hline
& & & & & & & \\
\hline
& & & \\
\hline
& & & & \\$$

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

D2: $\varphi \ ess \ x \equiv \varphi(x) \land \forall \psi. (\psi(x) \to \Box \forall x. (\varphi(x) \to \psi(x)))$

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

D3*: $E(x) \equiv \Box \exists y. G(y)$

D3*: $E(x) \equiv \Box \exists y. G(y)$

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

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

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

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

A2

Natural Deduction Calculus

$$\frac{A}{A} \quad \overline{B}$$

$$\vdots \quad \vdots \quad \vdots$$

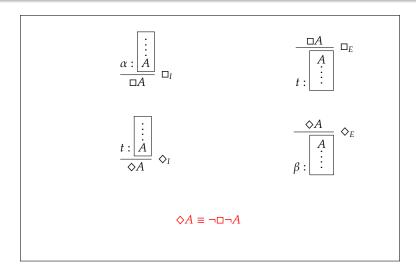
$$\frac{A \vee B \quad \overline{C} \quad \overline{C}}{C} \vee_{E} \qquad \frac{A}{A \wedge B} \wedge_{I} \qquad \frac{B}{A \to B} \rightarrow_{I}^{n}$$

$$\frac{A}{A \vee B} \vee_{I_{1}} \qquad \frac{A \wedge B}{A} \wedge_{E_{1}} \qquad \frac{B}{A \to B} \rightarrow_{I}$$

$$\frac{B}{A \vee B} \vee_{I_{2}} \qquad \frac{A \wedge B}{B} \wedge_{E_{2}} \qquad \frac{A}{B} \rightarrow_{E}$$

$$\frac{A[\alpha]}{\forall x.A[x]} \forall_{I} \qquad \frac{\forall x.A[x]}{A[t]} \forall_{E} \qquad \frac{A[t]}{\exists x.A[x]} \exists_{I} \qquad \frac{\exists x.A[x]}{A[\beta]} \exists_{E}$$

$$\neg A \equiv A \to \bot$$



Natural Deduction Proofs T1 and C1

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

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

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

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

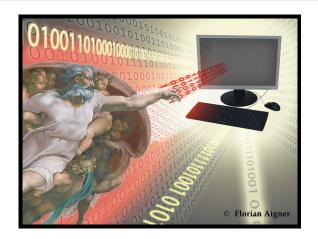
$$\frac{P(\rho) \rightarrow \Diamond \exists x. \rho(x)}{\underline{T1:} \forall \varphi. [P(\varphi) \rightarrow \Diamond \exists x. \varphi(x)]} \forall_{E}$$

$$\frac{A3}{\underline{P(G)}} \frac{\nabla \varphi. [P(\varphi) \rightarrow \Diamond \exists x. \varphi(x)]}{\underline{P(G) \rightarrow \Diamond \exists x. G(x)}} \forall_{E}$$

$$\Diamond \exists x. G(x)$$

Natural Deduction Proofs T2 (Partial)

$$\begin{array}{c|c} & \square P(\psi)^7 & \square_E & \square_3 & \square_7 & \square_8 \\ \hline P(\psi) & \square_E & P(\psi) \rightarrow \forall x. (\overrightarrow{G}(x) \rightarrow \psi(x)) & \rightarrow_E \\ \hline P(\psi) & \square \forall x. (G(x) \rightarrow \psi(x)) & \square_1 & \square_2 & \square_3 & \square_4 & \square_4 \\ \hline P(\psi) & \square \forall x. (G(x) \rightarrow \psi(x)) & \square_1 & \square_2 & \square_3 & \square_4 & \square_4 \\ \hline \square P(\psi) & \square \forall x. (G(x) \rightarrow \psi(x)) & \square_1 & \square_2 & \square_4 & \square_4 & \square_4 & \square_4 & \square_4 \\ \hline \square P(\psi) & \square \forall x. (G(x) \rightarrow \psi(x)) & \rightarrow_E & \square_4 \\ \hline \square P(\psi) & \square \forall x. (G(x) \rightarrow \psi(x)) & \rightarrow_E & \square_4 &$$



Part B:

Formalization: Automation: Consistency: in classical higher-order logic (HOL) theorem provers Leo-II and Satallax model finder Nitpick (Nitrox)

Challenge: No provers for Higher-order Quantified Modal Logic (QML)

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

Then use existing HOL theorem provers for reasoning in QML

[BenzmüllerPaulson, Logica Universalis, 2013]

Previous empiricial findings:

Embedding of *First-order Modal Logic* in HOL works well

[BenzmüllerOttenRaths, ECAI, 2012]

[Benzmüller, LPAR, 2013]

QML
$$\varphi, \psi ::= \dots | \neg \varphi | \varphi \land \psi | \varphi \rightarrow \psi | \Box \varphi | \Diamond \varphi | \forall x \varphi | \exists x \varphi | \forall P \varphi$$

Kripke style semantics (possible world semantics)

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

- 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, ...
```

QML
$$\varphi, \psi ::= \ldots |\neg \varphi| \varphi \land \psi | \varphi \rightarrow \psi | \Box \varphi | \diamond \varphi | \forall x \varphi | \exists x \varphi | \forall P \varphi$$
HOL $s,t ::= C | x | \lambda xs | st | \neg s | s \lor t | \forall x t$

QML in HOL: QML formulas φ are mapped to HOL predicates $\varphi_{\iota \to 0}$

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

The equations in Ax are given as axioms to the HOL provers! (Remark: Note that we are here dealing with constant domain quantification)

QML
$$\varphi, \psi ::= \ldots |\neg \varphi| \varphi \land \psi |\varphi \rightarrow \psi| \Box \varphi | \diamond \varphi | \forall x \varphi | \exists x \varphi | \forall P \varphi$$
HOL $s,t ::= C |x| \lambda xs |st| \neg s |s \lor t| \forall x t$

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

The equations in Ax are given as axioms to the HOL provers! (Remark: Note that we are here dealing with constant domain quantification)

QML
$$\varphi, \psi ::= \ldots |\neg \varphi| \varphi \land \psi |\varphi \rightarrow \psi| \Box \varphi | \diamond \varphi | \forall x \varphi | \exists x \varphi | \forall P \varphi$$
HOL $s,t ::= C |x| \lambda xs |st| \neg s |s \lor t| \forall x t$

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

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

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

(Remark: Note that we are here dealing with constant domain quantification)

Example

QML formula

QML formula in HOL expansion, $\beta\eta$ -conversion expansion, $\beta\eta$ -conversion expansion, $\beta\eta$ -conversion

$\Diamond \exists x G(x)$

 $\forall a \text{lid} (\Diamond \exists x G(x))_{t \to 0} \\ \forall w_t (\Diamond \exists x G(x))_{t \to 0} w \\ \forall w_t \exists u_t (rwu \land (\exists x G(x))_{t \to 0} u) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t \exists u_t (rwu \land \exists x Gxu) \\ \forall w_t (rwu \land x Gxu) \\ \forall$

What are we doing?

In order to prove that φ is valid in QML,

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

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

Example

QML 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 QML,

-> we instead prove that valid $\varphi_{t\rightarrow 0}$ can be derived from Ax in HOL.

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

Example

QML formula QML 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 QML, —> we instead prove that valid $\varphi_{t \to o}$ can be derived from Ax in HO

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

Example

```
QML formula QML 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 QML,

-> we instead prove that valid $arphi_{t o o}$ can be derived from Ax in HOL.

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

Example

```
QML formula QML 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 QML,

–> we instead prove that valid $arphi_{t o o}$ can be derived from Ax in HOL.

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

Example

```
QML formula QML 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 QML,

-> we instead prove that valid $\varphi_{t\to 0}$ can be derived from Ax in HOL.

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

Example

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

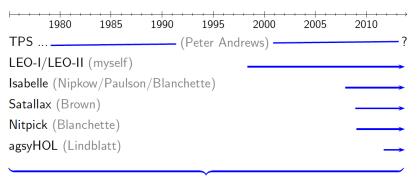
What are we doing?

In order to prove that φ is valid in QML,

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

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

Automated Theorem Provers and Model Finders for HOL



- all accept TPTP THF Syntax [SutcliffeBenzmüller, J.Form.Reas, 2009]
 - can be called remotely via SystemOnTPTP at Miami
 - they significantly gained in strength over the last years
 - they can be bundled into a combined prover HOL-P

Exploit HOL with Henkin semantics as metalogic

Automate other logics (& combinations) via semantic embeddings

— HOL-P becomes a Universal Reasoner —

Proof Automation and Consistency 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 Migmi 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

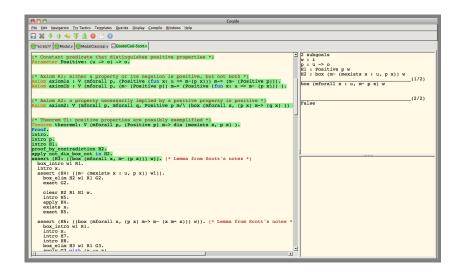
- Goal: verification of the natural deduction proof
 - Step-by-step formalization
 - Almost no automation (intentionally!)
- Interesting facts:
 - Embedding is transparent to the user
 - Embedding gives labeled calculus for free

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

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

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

Cog Proof





Part D:

automation & verification: proof assistant Isabelle











Installation Documentation

Community Site Mirrors: Combridge (.uk)

calculus. Isabelle is developed at University of Cambridge (Larry Paulson), Technische Universität München (Tobias Nipkow) and Université Paris-Sud (Makarius Wenzel). See the Isabelle overview for a brief introduction.

What is Isabelle?

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

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

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 subscribe or see the archive (also available via Google groups and Narkive).
- . isabelle-dev@in.tum.de covers the Isabelle development process, including intermediate repository versions, and administrative issues concerning the website or testing infrastructure. Early adopters of repository versions should subscribe or see the archive (also available at mail-archive.com or gmane.org).

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

Automation & Verification in Proof Assistant Isabelle/HOL

Isabelle/HOL (Cambridge University/TU Munich)

- HOL instance of the generic Isabelle proof assistant
- User interaction and proof automation
- Automation is supported by Sledgehammer tool
- Verification of the proofs in Isabelle/HOL's small proof kernel

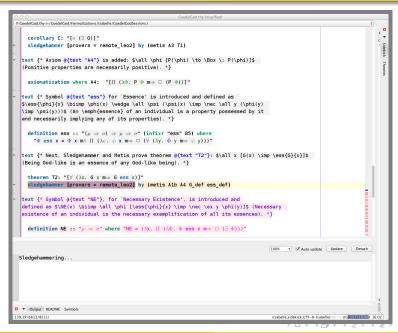
What we did?

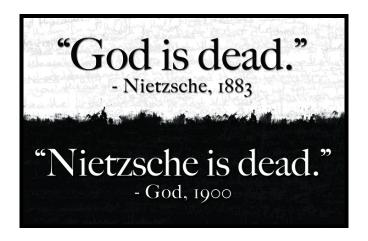
- Proof automation of Gödel's proof script (Scott version)
- SLEDGEHAMMER makes calls to remote THF provers in Miami
- These calls the suggest respective calls to the Metis prover
- Metis proofs are verified in Isabelle/HOL's proof kernel

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



Automation & Verification in Proof Assistant Isabelle/HOL





Part E: Criticisms

$\forall P. [\Diamond \Box P \rightarrow \Box P]$

If something is possibly necessary, then it is necessary.

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M.M \models F \longrightarrow \Box B$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

 $\Diamond \Box \Diamond \Diamond F$

weak intuitions \Rightarrow dozens of modal logics

S5 is considered adequate



$$\forall P. [\Diamond \Box P \rightarrow \Box P]$$

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box B$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

 $\Diamond \Box \Diamond \Diamond F$

weak intuitions \Rightarrow dozens of modal logics

S5 is considered adequate



$$\forall P. [\Diamond \Box P \rightarrow \Box P]$$

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

ogical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

 $\Diamond \Box \Diamond \Diamond F$

weak intuitions \Rightarrow dozens of modal logics

S5 is considered adequate



$$\forall P. [\Diamond \Box P \rightarrow \Box P]$$

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

 $\Diamond \Box \Diamond \Diamond F$

weak intuitions \Rightarrow dozens of modal logics

S5 is considered adequate

$$\forall P. [\Diamond \Box P \rightarrow \Box P]$$

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

 $\Diamond \Box \Diamond \Diamond F$

weak intuitions \Rightarrow dozens of modal logics

S5 is considered adequate

$$\forall P. [\lozenge \Box P \to \Box P]$$

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

$$\Diamond \Box \Diamond \Diamond F$$

weak intuitions \Rightarrow dozens of modal logics

S5 is considered adequate



$$\forall P. [\Diamond \Box P \rightarrow \Box P]$$

$$\Diamond \Box (A \lor \neg A) \qquad \Box (A \lor \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

$$\Diamond \Box \Diamond \Diamond F$$

weak intuitions ⇒ dozens of modal logics

S5 is considered adequate



$$\forall P. [\lozenge \Box P \to \Box P]$$

$$\diamondsuit_c \square_c (A \vee \neg A) \qquad \square_c (A \vee \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists $M, M \models F \longrightarrow \Diamond F$

What about iterations?

$$\Diamond \Box \Diamond \Diamond F$$

weak intuitions ⇒ dozens of modal logics

S5 is considered adequate



$$\forall P. [\Diamond \Box P \rightarrow \Box P]$$

If something is possibly necessary, then it is necessary.

$$\diamondsuit_c \square_c (A \vee \neg A) \qquad \square_c (A \vee \neg A)$$

logical necessity ~ validity

logical possibility ~ satisfiability

for all $M, M \models F \longrightarrow \Box F$

exists
$$M, M \models F \longrightarrow \Diamond F$$

What about iterations?

$$\Diamond \Box \Diamond \Diamond F$$

weak intuitions ⇒ dozens of modal logics

S5 is considered adequate

(But KB is sufficient!)



$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths".

Everything is determined.

There is no free will.

$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths".

Everything is determined.

There is no free will

$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths".

everything is determined.

There is no free will.

$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths". Everything is determined.

There is no free will.

$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths". Everything is determined. There is no free will.

$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths". Everything is determined. There is no free will.

$$\forall P.[P \rightarrow \Box P]$$

Follows from T2, T3 and D2.

There are no contingent "truths".

Everything is determined.

There is no free will.

$$\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$$

Are the following properties positive or negative?

$$\lambda x.G(x)$$
 $\lambda x.E(x)$ $\lambda x.x = x$ $\lambda x.T$

$$\lambda x.blue(x)$$
 $\lambda x.white(x)$ $\lambda x.human(x)$

Solution:

"... positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the ax. true. ..."

$$\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$$

Are the following properties positive or negative?

$$\lambda x.G(x)$$
 $\lambda x.E(x)$ $\lambda x.x = x$ $\lambda x.\top$

$$\lambda x.blue(x)$$
 $\lambda x.white(x)$ $\lambda x.human(x)$

Solution:

"... positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the ax. true. ..."

$$\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$$

Are the following properties positive or negative?

$$\lambda x.G(x)$$
 $\lambda x.E(x)$ $\lambda x.x = x$ $\lambda x.\top$

$$\lambda x.blue(x)$$
 $\lambda x.white(x)$ $\lambda x.human(x)$

Solution:

"... positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the ax. true. ..."

Gödel, 1970

$$\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$$

Are the following properties positive or negative?

$$\lambda x.G(x)$$
 $\lambda x.E(x)$ $\lambda x.x = x$ $\lambda x.\top$

$$\lambda x.blue(x)$$
 $\lambda x.white(x)$ $\lambda x.human(x)$

Solution:

"... positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the ax. true. ..."

$$\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$$

Are the following properties positive or negative?

$$\lambda x.G(x)$$
 $\lambda x.E(x)$ $\lambda x.x = x$ $\lambda x.\top$

$$\lambda x.blue(x)$$
 $\lambda x.white(x)$ $\lambda x.human(x)$

Solution:

"... positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the ax. true. ..."

$$\forall \phi [P(\neg \phi) \leftrightarrow \neg P(\phi)]$$

Are the following properties positive or negative?

$$\lambda x.G(x)$$
 $\lambda x.E(x)$ $\lambda x.x = x$ $\lambda x.\top$

$$\lambda x.blue(x)$$
 $\lambda x.white(x)$ $\lambda x.human(x)$

Solution:

"... positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the ax. true. ..."



Part F: Conclusions

Summary of Results

The (new) insights we gained from experiments include:

- Logic K sufficient for T1, C and T2
- Logic S5 not needed for T3
- Logic KB sufficient for T3 (not well known)
- We found a simpler new proof of C
- Gödel's 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

Our novel contributions to the theorem proving community include

- Powerful infrastructure for reasoning with QML
- A new natural deduction calculus for higher-order modal logic
- Difficult new benchmarks problems for HOL provers
- Huge media attention

What have we achieved

- Verification of Gödel's ontological argument with HOL provers
 - exact parameters known: constant domain quantification, Henkin Semantics
 - experiments with different parameters could be performed
- Gained some novel results and insights
- Major step towards Computer-assisted Theoretical Philosophy
 - see also Ed Zalta's Computational Metaphysics project at Stanford University
 - see also John Rushby's recent verification of Anselm's proof in PVS
 - remember Leibniz' dictum Calculemus!
- Interesting bridge between CS, Philosophy and Theology

Ongoing and future work

- Formalize and verify literature on ontological arguments
 ...in particular the criticism and improvements to Gödel
- Own contributions supported by theorem provers

What have we achieved

- Verification of Gödel's ontological argument with HOL provers
 - exact parameters known: constant domain quantification, Henkin Semantics
 - experiments with different parameters could be performed
- Gained some novel results and insights
- Major step towards Computer-assisted Theoretical Philosophy
 - see also Ed Zalta's Computational Metaphysics project at Stanford University
 - see also John Rushby's recent verification of Anselm's proof in PVS
 - remember Leibniz' dictum Calculemus!
- Interesting bridge between CS, Philosophy and Theology

Ongoing and future work

- Formalize and verify literature on ontological arguments
 - ...in particular the criticism and improvements to Gödel
- Own contributions supported by theorem provers

Some Comments and Reactions

I'm sure that God would be impressed with your proof, if only he existed :-)

Larry

Die Philosophen können so schön staunen.

Sie packen Dinge in Begriffe (gucken dabei in die Luft) werfen die Begriffe dann in ihre Philosophiekiste, schütteln ganz dolle, und freuen sich, dass ganz genau rauskommt, was sie vorher reingetan haben. Und das geht sogar, wenn eine Maschine die Kiste schüttelt.

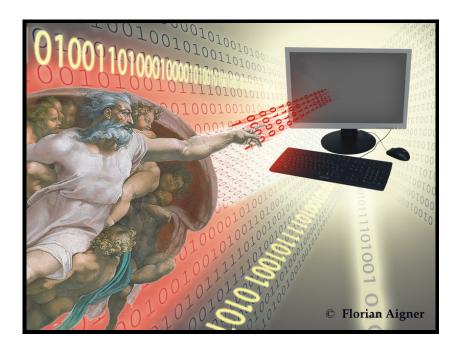
Unerstaunt 2017cp

60. Suchlauf

souveränsatt 09.09.2013

man kann auch auf andere Weise in diesem Zusammenhang methodisch vorgehen: bei einer längeren Autofahrt das Radio auf automatischen Suchlauf stellen. Nach zwei Tagen sieht man Gott

... find more on the internet ...





The following images used in these slides were obtained in commons.wikimedia.org and are licensed as follows:

CC-BY-SA:

ReligiousSymbols, PaganReligiousSymbols, NoGod.

Public Domain: TrifidNebula