Gödel's God in Isabelle/HOL

Christoph Benzmüller and Bruno Woltzenlogel Paleo

November 4, 2013

A1	Either a property or its negation is positive, but not 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:	$NE(x) \leftrightarrow \forall\phi[\phi\ ess.\ x \rightarrow \Box\exists y\phi(y)]$
A5	Necessary existence is a positive property:	$P(NE)$
T3	Necessarily, God exists:	$\Box\exists x G(x)$

1 Introduction

A formalization and (partial) automation of Dana Scott's version [11] of Goedel's ontological argument [8] in quantified modal logic KB (QML KB) is presented. QML KB is in turn modeled as a fragment of classical higher-order logic (HOL). Thus, the formalization is essentially a formalization in HOL. The employed embedding of QML KB in HOL is adapting the work of Benzmüller and Paulson [2, 1]. Note that the QML KB formalization employs quantification over individuals and quantification over sets of individuals (poperties).

The formalization presented here has been carried out and verified in the Isabelle/HOL proof assistant; for more information on this system see the textbook by Nipkow, Paulson, and Wenzel [10]. More recent tutorials on Isabelle can be found at: http://isabelle.in.tum.de.

Some further notes:

1. This LaTeX text document has been produced automatically from the Isabelle source code document at https://github.com/FormalTheology/GoedelGod/tree/master/Formalizations/Isabelle/GoedelGodSession with the Isabelle build tool.

2. The formalization presented here is related to the THF [13] and Coq [4] formalizations at https://github.com/FormalTheology/GoedelGod/tree/master/Formalizations/.

3. All reasoning gaps in Scott's proof script have been automated with Sledgehammer [5], performing remote calls to the higher-order automated theorem prover LEO-II [3]. These calls suggest the Metis [9] calls as given below. The Metis proofs are verified in Isabelle/HOL.

4. For consistency checking, the model finder Nitpick [6] has been employed.

2 An Embedding of QML KB in HOL

The types i for possible worlds (or states) and μ for individuals are introduced.

typedecl i — the type for possible worlds
typedecl μ — the type for indiviuals

Possible worlds are connected by an accessibility relation r.

consts $r :: i \Rightarrow i \Rightarrow \mathit{bool}$ (**infixr** r 70) — accessibility relation r

The B axiom (symmetry) for relation r is stated. B is needed only for proving theorem T3.

axiomatization where sym: $x\ r\ y \longrightarrow y\ r\ x$

QML formulas are identified with certain HOL terms of type $i \Rightarrow \mathit{bool}$. This type will be abbreviated in the remainder as σ.

type-synonym $\sigma = (i \Rightarrow \mathit{bool})$

The classical connectives $\neg, \wedge, \rightarrow$, and $\forall$ (over individuals and over sets of individuals) and $\exists$ (over individuals) are lifted to type σ. The lifted connectives are $m\neg$, $m\wedge$, $m\Rightarrow$, $\forall$, Π, and $\exists$. Further connectives could be introduced analogously. Definitions could be used instead of abbreviations.

abbreviation $\mathit{mnot} :: \sigma \Rightarrow \sigma$ $(m\neg)$ **where** $m\neg\ \varphi \equiv (\lambda w.\ \neg\ \varphi\ w)$
abbreviation $\mathit{mand} :: \sigma \Rightarrow \sigma \Rightarrow \sigma$ (**infixr** $m\wedge$ 79) **where** $\varphi\ m\wedge\ \psi \equiv (\lambda w.\ \varphi\ w \wedge \psi\ w)$
abbreviation $\mathit{mimplies} :: \sigma \Rightarrow \sigma \Rightarrow \sigma$ (**infixr** $m\Rightarrow$ 74) **where** $\varphi\ m\Rightarrow\ \psi \equiv (\lambda w.\ \varphi\ w \longrightarrow \psi\ w)$
abbreviation $\mathit{mforall\text{-}ind} :: (\mu \Rightarrow \sigma) \Rightarrow \sigma$ $(\forall)$ **where** $\forall\ \Phi \equiv (\lambda w.\ \forall x.\ \Phi\ x\ w)$
abbreviation $\mathit{mforall\text{-}indset} :: ((\mu \Rightarrow \sigma) \Rightarrow \sigma) \Rightarrow \sigma$ (Π) **where** $\Pi\ P \equiv (\lambda w.\ \forall x.\ P\ x\ w)$
abbreviation $\mathit{mexists\text{-}ind} :: (\mu \Rightarrow \sigma) \Rightarrow \sigma$ $(\exists)$ **where** $\exists\ \Phi \equiv (\lambda w.\ \exists x.\ \Phi\ x\ w)$
abbreviation $\mathit{mbox} :: \sigma \Rightarrow \sigma$ $(\Box)$ **where** $\Box\ \varphi \equiv (\lambda w.\ \forall v.\ \neg\ w\ r\ v \vee \varphi\ v)$
abbreviation $\mathit{mdia} :: \sigma \Rightarrow \sigma$ $(\Diamond)$ **where** $\Diamond\ \varphi \equiv (\lambda w.\ \exists v.\ w\ r\ v \wedge \varphi\ v)$

For grounding lifted formulas, the meta-predicate *valid* is introduced.

abbreviation $\mathit{valid} :: \sigma \Rightarrow \mathit{bool}$ ([-]) **where** $[p] \equiv \forall w.\ p\ w$

3 Gödel's Ontological Argument

Constant symbol P (Gödel's 'Positive') is declared.

consts $P :: (\mu \Rightarrow \sigma) \Rightarrow \sigma$

The meaning of P is restricted by axioms $A1(a/b)$: $\forall\phi[P(\neg\phi) \leftrightarrow \neg P(\phi)]$ (Either a property or its negation is positive, but not both.) and $A2$: $\forall\phi\forall\psi[(P(\phi) \wedge \Box\forall x[\phi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]$ (A property necessarily implied by a positive property is positive).

axiomatization where
A1a: $[\Pi\ (\lambda\Phi.\ P\ (\lambda x.\ m\neg\ (\Phi\ x))\ m\Rightarrow\ m\neg\ (P\ \Phi))]$ **and**
A1b: $[\Pi\ (\lambda\Phi.\ m\neg\ (P\ \Phi)\ m\Rightarrow\ P\ (\lambda x.\ m\neg\ (\Phi\ x)))]$ **and**
A2: $[\Pi\ (\lambda\Phi.\ \Pi\ (\lambda\psi.\ (P\ \Phi\ m\wedge\ \Box\ (\forall\ (\lambda x.\ \Phi\ x\ m\Rightarrow\ \psi\ x)))\ m\Rightarrow\ P\ \psi))]$

We prove theorem T1: $\forall\varphi[P(\varphi) \rightarrow \Diamond\exists x\varphi(x)]$ (Positive properties are possibly exemplified). T1 is proved directly by Sledghammer with command *sledgehammer* [*provers* = *remote-leo2*]. This successful attempt then suggests to instead try the Metis call in the line below. This Metis call generates a proof object that is verified in Isabelle/HOL's kernel.

theorem *T1*: $[\Pi\ (\lambda\Phi.\ P\ \Phi\ m\Rightarrow\ \Diamond\ (\exists\ \Phi))]$
sledgehammer [*provers* = *remote-leo2*]
by (*metis A1a A2*)

Next, the symbol G for 'God-like' is introduced and defined as $G(x) \leftrightarrow \forall\phi[P(\phi) \rightarrow \phi(x)]$ (A God-like being possesses all positive properties).

definition $G :: \mu \Rightarrow \sigma$ **where** $G = (\lambda x.\ \Pi\ (\lambda\Phi.\ P\ \Phi\ m\Rightarrow\ \Phi\ x))$

Axiom *A3* is added: $P(G)$ (The property of being God-like is positive). Sledgehammer and Metis then prove corollary C: $\Diamond\exists x G(x)$ (Possibly, God exists).

axiomatization where *A3*: $[P\ G]$

corollary *C*: $[\Diamond\ (\exists\ G)]$
sledgehammer [*provers* = *remote-leo2*] **by** (*metis A3 T1*)

Axiom *A4* is added: $\forall\phi[P(\phi) \rightarrow \Box\ P(\phi)]$ (Positive properties are necessarily positive).

axiomatization where *A4*: $[\Pi\ (\lambda\Phi.\ P\ \Phi\ m\Rightarrow\ \Box\ (P\ \Phi))]$

Symbol *ess* for 'Essence' is introduced and defined as $\phi\ ess.\ x \leftrightarrow \phi(x)\wedge\forall\psi(\psi(x) \rightarrow \Box\forall y(\phi(y) \rightarrow \psi(y)))$ (An *essence* of an individual is a property possessed by it and necessarily implying any of its properties).

definition *ess* $:: (\mu \Rightarrow \sigma) \Rightarrow \mu \Rightarrow \sigma$ (**infixr** *ess 85*) **where**
$\Phi\ ess\ x = \Phi\ x\ m\wedge\ \Pi\ (\lambda\psi.\ \psi\ x\ m\Rightarrow\ \Box\ (\forall\ (\lambda y.\ \Phi\ y\ m\Rightarrow\ \psi\ y)))$

Next, Sledgehammer and Metis prove theorem *T2*: $\forall x[G(x) \rightarrow G\ ess.\ x]$ (Being God-like is an essence of any God-like being).

theorem *T2*: $[\forall\ (\lambda x.\ G\ x\ m\Rightarrow\ G\ ess\ x)]$
sledgehammer [*provers* = *remote-leo2*] **by** (*metis A1b A4 G-def ess-def*)

Symbol *NE*, for 'Necessary Existence', is introduced and defined as $NE(x) \leftrightarrow \forall\phi[\phi\ ess.\ x \rightarrow \Box\exists y\phi(y)]$ (Necessary existence of an individual is the necessary exemplification of all its essences).

definition $NE :: \mu \Rightarrow \sigma$ **where** $NE = (\lambda x.\ \Pi\ (\lambda\Phi.\ \Phi\ ess\ x\ m\Rightarrow\ \Box\ (\exists\ \Phi)))$

Moreover, axiom *A5* is added: $P(NE)$ (Necessary existence is a positive property).

axiomatization where *A5*: $[P\ NE]$

Finally, Sledgehammer and Metis prove the main theorem *T3*: $\Box\exists x G(x)$ (Necessarily, God exists).

theorem *T3*: $[\Box\ (\exists\ G)]$

sledgehammer [*provers* = *remote-leo2*] **by** (*metis A5 C T2 sym G-def NE-def*)

corollary *C2*: [$\exists$ *G*]
sledgehammer [*provers* = *remote-leo2*](*T1 T3 G-def sym*) **by** (*metis T1 T3 G-def sym*)

The consistency of the entire theory is checked with Nitpick.

lemma *True* **nitpick** [*satisfy*, *user-axioms*, *expect* = *genuine*] **oops**

It has been critisized that Gödel's ontological argument implies what is called the modal collapse. The prover Satallax [7] can indeed show this, but verification with Metis still fails.

lemma *MC*: [p $m\Rightarrow$ ($\Box$ p)]
using *T2 T3 ess-def sym* **sledgehammer** [*provers* = *remote-satallax*] **oops**

4 Further results on Gödel's God.

Lifted Leibniz equality is introduced.

abbreviation *mequals* :: $\mu \Rightarrow \mu \Rightarrow \sigma$ (**infixr** *m=* *90*) **where** x *m=* $y \equiv \Pi$ (λ φ.(φ x $m\Rightarrow$ φ y))

Gödel's God is flawless, that is, he has no negative properties.

theorem *Flawless*: [Π ($\lambda\varphi$. $\forall$ (λx. (G x $m\Rightarrow$ ($m\neg$ (P φ) $m\Rightarrow$ $m\neg$ (φ x))))))]
sledgehammer [*provers* = *remote-leo2*] **by** (*metis A1b G-def*)

Moreover, it can be shown that any two God-like beings are equal, that is, there is only one God-like being.

theorem *Monotheism*: [$\forall$ (λx. $\forall$ (λy. ($G(x)$ $m\Rightarrow$ ($G(y)$ $m\Rightarrow$ (x *m=* y)))))]
sledgehammer [*provers* = *remote-leo2*] **by** (*metis Flawless G-def*)

Add-on: We briefly show that lifted Leibniz equality indeed denotes equality.

lemma *eqtest1*: $x = y \Longrightarrow$ [x *m=* y]
sledgehammer [*provers* = *remote-leo2*] **by** *metis*

lemma *eqtest2*: [x *m=* y] $\Longrightarrow x = y$
sledgehammer [*provers* = *remote-satallax*] **oops**

Acknowledgments: Nik Sultana, Jasmin Blanchette and Larry Paulson provided very important help wrt consistency checking in Isabelle. Jasmin Blanchette instructed us on how to produce latex documents from Isabelle sources, and he showed us useful tricks in Isabelle.

References

[1] C. Benzmüller and L.C. Paulson. Exploring properties of normal multimodal logics in simple type theory with LEO-II. In *Festschrift in Honor of Peter B. Andrews on His 70th Birthday*, pp. 386–406. College Publications.

[2] C. Benzmüller and L.C. Paulson. Quantified multimodal logics in simple type theory. *Logica Universalis (Special Issue on Multimodal Logics)*, 7(1):7–20, 2013.

[3] C. Benzmüller, F. Theiss, L. Paulson, and A. Fietzke. LEO-II - a cooperative automatic theorem prover for higher-order logic. In *Proc. of IJCAR 2008*, volume 5195 of *LNAI*, pp. 162–170. Springer, 2008.

[4] Y. Bertot and P. Casteran. *Interactive Theorem Proving and Program Development.* Springer, 2004.

[5] J.C. Blanchette, S. Böhme, and L.C. Paulson. Extending Sledgehammer with SMT solvers. *Journal of Automated Reasoning*, 51(1):109–128, 2013.

[6] J.C. Blanchette and T. Nipkow. Nitpick: A counterexample generator for higher-order logic based on a relational model finder. In *Proc. of ITP 2010*, LNCS 6172, pp. 131–146. Springer, 2010.

[7] C.E. Brown. Satallax: An automated higher-order prover. In *Proc. of IJCAR 2012*, LNAI 7364, pp. 111 – 117. Springer, 2012.

[8] K. Gödel. *Appendix A. Notes in Kurt Gödel's Hand*, pp. 144–145. In [12], 2004.

[9] J. Hurd. First-order proof tactics in higher-order logic theorem provers. In *Design and Application of Strategies/Tactics in Higher Order Logics, NASA Tech. Rep. NASA/CP-2003-212448*, 2003.

[10] T. Nipkow, L.C. Paulson, and M. Wenzel. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic.* LNCS 2283. Springer, 2002.

[11] D. Scott. *Appendix B. Notes in Dana Scott's Hand*, pp. 145–146. In [12], 2004.

[12] J.H. Sobel. *Logic and Theism: Arguments for and Against Beliefs in God.* Cambr. U. Press, 2004.

[13] G. Sutcliffe and C. Benzmüller. Automated reasoning in higher-order logic using the TPTP THF infrastructure. *Journal of Formalized Reasoning*, 3(1):1–27, 2010.