# Gödel’s God in Isabelle/HOL

Christoph Benzmüller and Bruno Woltzenlogel Paleo

September 13, 2023

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Either a property or its negation is positive, but not both:</td>
</tr>
<tr>
<td>A2</td>
<td>A property necessarily implied by a positive property is positive:</td>
</tr>
<tr>
<td>T1</td>
<td>Positive properties are possibly exemplified:</td>
</tr>
<tr>
<td>D1</td>
<td>A God-like being possesses all positive properties:</td>
</tr>
<tr>
<td>A3</td>
<td>The property of being God-like is positive:</td>
</tr>
<tr>
<td>C</td>
<td>Possibly, God exists:</td>
</tr>
<tr>
<td>A4</td>
<td>Positive properties are necessarily positive:</td>
</tr>
<tr>
<td>D2</td>
<td>An essence of an individual is a property possessed by it and necessarily implying any of its properties:</td>
</tr>
<tr>
<td>T2</td>
<td>Being God-like is an essence of any God-like being:</td>
</tr>
<tr>
<td>D3</td>
<td>Necessary existence of an individual is the necessary exemplification of all its essences:</td>
</tr>
<tr>
<td>A5</td>
<td>Necessary existence is a positive property:</td>
</tr>
<tr>
<td>T3</td>
<td>Necessarily, God exists:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>Possibly, God exists:</td>
</tr>
</tbody>
</table>

Figure 1: Scott’s version of Gödel’s ontological argument [12].

## 1 Introduction

Dana Scott’s version [12] (cf. Fig. 1) of Gödel’s proof of God’s existence [8] is formalized in quantified modal logic KB (QML KB) within the proof assistant Isabelle/HOL. QML KB is 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 (properties).

The gaps in Scott’s proof have been automated with Sledgehammer [5], performing remote calls to the higher-order automated theorem prover LEO-II [3]. Sledgehammer suggests the Metis [9] calls, which result in proofs that are verified by Isabelle/HOL. For consistency checking, the model finder Nitpick [6] has been employed. The successful calls to Sledgehammer are deliberately kept as comments in the file for demonstration purposes (normally, they are automatically eliminated by Isabelle/HOL).
Isabelle is described in the textbook by Nipkow, Paulson, and Wenzel [10] and in tutorials available at: http://isabelle.in.tum.de.

1.1 Related Work


An older ontological argument by Anselm was formalized in PVS by John Rushby [11].

2 An Embedding of QML KB in HOL

The types $i$ for possible worlds and $\mu$ for individuals are introduced.

- **typedecl** $i$ — the type for possible worlds
- **typedecl** $\mu$ — the type for individuals

Possible worlds are connected by an accessibility relation $r$.

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

QML formulas are translated as HOL terms of type $i \Rightarrow bool$. This type is abbreviated as $\sigma$.

- **type-synonym** $\sigma = (i \Rightarrow bool)$

The classical connectives $\neg, \land, \rightarrow$, and $\forall$ (over individuals and over sets of individuals) and $\exists$ (the latter two are modeled as constant symbols). Other connectives can be introduced analogously. We exemplarily do this for $\lor$, $\equiv$, and $\mathsf{Leibniz}$ equality on individuals.

- **abbreviation** $m\neg :: \sigma \Rightarrow \sigma$ where $m\neg \varphi \equiv (\lambda w. m\neg \varphi w)$
- **abbreviation** $m\land :: \sigma \Rightarrow \sigma \Rightarrow \sigma$ where $m\land \varphi \equiv (\lambda w. m\land \varphi w)$
- **abbreviation** $m\lor :: \sigma \Rightarrow \sigma \Rightarrow \sigma$ where $m\lor \psi \equiv (\lambda w. m\lor \psi w)$
- **abbreviation** $m\rightarrow :: \sigma \Rightarrow \sigma \Rightarrow \sigma$ where $m\rightarrow \psi \equiv (\lambda w. \forall x. m\rightarrow \psi x)$
- **abbreviation** $m\equiv :: \sigma \Rightarrow \sigma \Rightarrow \sigma$ where $m\equiv \varphi \equiv (\lambda w. \forall x. m\equiv \varphi x)$
- **abbreviation** $m\mathsf{Leibniz} :: \mu \Rightarrow \mu \Rightarrow \sigma$ where $m\mathsf{Leibniz} x y \equiv (\lambda \varphi. x m\rightarrow \varphi y)$
- **abbreviation** $m\mathsf{box} :: \sigma \Rightarrow \sigma$ where $m\mathsf{box} \varphi \equiv (\lambda w. \exists v. m\mathsf{box} \varphi v)$
- **abbreviation** $m\mathsf{dia} :: \sigma \Rightarrow \sigma$ where $m\mathsf{dia} \varphi \equiv (\lambda w. \exists v. m\mathsf{dia} \varphi v)$

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

- **abbreviation** valid $:: \sigma \Rightarrow bool$ 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 $A2$: $\forall \phi \forall \psi [P(\phi) \land \Box \forall x [\phi(x) \rightarrow \psi(x)]] \rightarrow P(\psi)]$ (A property necessarily implied by a positive property is positive).

**axiomatization where**

$A1a$: $\forall (\lambda \Phi. P (\lambda x. \ m \neg (\Phi \ x)) \ m \rightarrow \ m \neg (P \ \Phi))]$ and $A1b$: $\forall (\lambda \Phi. \ m \neg (P \ \Phi) \ m \rightarrow P (\lambda x. \ m \neg (\Phi \ x)))]$ and $A2$: $\forall (\lambda \Phi. \ \forall (\lambda \Psi. (P \ \Phi \ m \wedge \ \Box (\forall (\lambda x. \ \Phi \ x \ m \rightarrow \Psi \ x)))) \ m \rightarrow P (\Psi))]$

We prove theorem $T1$: $\forall \phi [P(\phi) \rightarrow \Diamond \exists x \phi(x)]$ (Positive properties are possibly exemplified). $T1$ is proved directly by Sledgehammer with command `sledgehammer [provers = remote_leo2]`. Sledgehammer suggests to call Metis with axioms $A1a$ and $A2$. Metis successfully generates a proof object that is verified in Isabelle/HOL’s kernel.

**theorem** $T1$: $\forall (\lambda \Phi. P \ \Phi \ m \rightarrow \Diamond (\exists \Phi))]$

− sledgehammer [provers = remote_leo2]

```proof```

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 $G = (\lambda x. \forall (\lambda \Phi. P \ \Phi \ m \rightarrow \Phi \ x))$

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]

```proof```

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

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

Symbol $ess$ for ‘Essence’ is introduced and defined as

$\phi \ ess \ x \leftrightarrow \phi(x) \land \forall \psi (\psi(x) \rightarrow \Box \forall y (\phi(y) \rightarrow \psi(y)))$

(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

$\Phi \ ess \ x = \Phi \ x \ m \land \forall (\lambda \Psi. \Psi \ x \ m \rightarrow \Box (\forall (\lambda y. \Phi \ y \ m \rightarrow \Psi \ y)))$

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]

```proof```

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 \( NE = (\lambda x. \forall (\lambda \Phi. \Phi \ ess \ x \ m \rightarrow \Box (\exists \ \Phi)) \)

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

axiomatization where \( A5: \ [P \ NE] \)

The \( B \) axiom (symmetry) for relation \( r \) is stated. \( B \) is needed only for proving theorem T3 and for corollary C2.

axiomatization where \( \text{sym}: x \ r \ y \rightarrow y \ r \ x \)

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]

\( \langle \text{proof} \rangle \)

Surprisingly, the following corollary can be derived even without the \( T \) axiom (reflexivity).

corollary \( C2: [\exists \ G] \)

— sledgehammer [provers = remote_leo2]

\( \langle \text{proof} \rangle \)

The consistency of the entire theory is confirmed by Nitpick.

lemma \( \text{True nitpick} \ [\text{satisfy, user-axioms, expect = genuine}] \langle \text{proof} \rangle \)

4 Additional Results on Gödel’s God.

Gödel’s God is flawless: (s)he does not have non-positive properties.

theorem \( \text{Flawlessness:} \ [\forall (\lambda \Phi. \forall (\lambda x. (G x m \rightarrow (m \neg (P \ \Phi) \ m \rightarrow m \neg (\Phi x)))))] \)

— sledgehammer [provers = remote_leo2]

\( \langle \text{proof} \rangle \)

There is only one God: any two God-like beings are equal.

theorem \( \text{Monotheism:} \ [\forall (\lambda x. \forall (\lambda y. (G x m \rightarrow (G y m \rightarrow (x mL= y)))))] \)

— sledgehammer [provers = remote_leo2]

\( \langle \text{proof} \rangle \)

5 Modal Collapse

Gödel’s axioms have been criticized for entailing the so-called modal collapse. The prover Satallax [7] confirms this. However, sledgehammer is not able to determine which axioms, definitions and previous theorems are used by Satallax: hence it suggests to call Metis using everything, but this (unsurprisingly) fails. Attempting to use ‘Sledgehammer min’ to minimize Sledgehammer’s suggestion does not work. Calling Metis with \( T2, T3 \) and \( \text{ess-def} \) also does not work.

lemma \( \text{MC:} \ [\forall (\lambda \Phi. (\Phi m \rightarrow (\Box \ \Phi)))] \)

— sledgehammer [provers = remote_satallax]

— by \( \text{metis T2 T3 ess_def} \)

\( \langle \text{proof} \rangle \)
Acknowledgments: Nik Sultana, Jasmin Blanchette and Larry Paulson provided very important help on issues related to consistency checking in Isabelle. Jasmin Blanchette instructed us on producing Isabelle sessions and he showed us some useful tricks in Isabelle.

References


5