A (Simplified) Supreme Being Necessarily Exists, says the Computer: Computationally Explored Variants of Gödel's Ontological Argument

Christoph Benzmüller
Freie Universität Berlin, Germany
c.benzmueller@fu-berlin.de

Abstract

An approach to universal (meta-)logical reasoning in classical higher-order logic is employed to explore and study simplifications of Kurt Gödel's modal ontological argument. Some argument premises are modified, others are dropped, modal collapse is avoided and validity is shown already in weak modal logics **K** and **T**. Key to the gained simplifications of Gödel's original theory is the exploitation of a link to the notions of filter and ultrafilter in topology.

The paper illustrates how modern knowledge representation and reasoning technology for quantified non-classical logics can contribute new knowledge to other disciplines. The contributed material is also well suited to support teaching of non-trivial logic formalisms in classroom.

1 Introduction

Variants of Kurt Gödel's (1970), resp. Dana Scott's (1972), modal ontological argument have previously been studied and verified on the computer by Benzmüller and Woltzenlogel Paleo (2014; 2016) and Benzmüller and Fuenmayor (2020), and some previously unknown issues were revealed in these works,[1] and it was shown that logic **KB**, instead of **S5**, is already sufficient to derive from Gödel's axioms that a supreme being necessarily exists.

In this paper simplified variants of Gödel's modal ontological argument are explored and assessed. These simplifications have been developed in interaction with the proof assistant Isabelle/HOL (Nipkow, Paulson, and Wenzel 2002) and by employing the shallow semantical embedding (SSE) approach (Benzmüller 2019) as enabling technology. This technology supports the reuse of automated theorem proving (ATP) and model finding tools for classical higher-order logic (HOL) to represent and reason with a wide range of non-classical logics and theories, including higher-order modal logics (HOMLs) and Gödel's modal ontological argument, which are in the focus of this paper.

One of the new, simplified modal arguments is as follows. The notion of being Godlike ($\mathcal{G}$) is exactly as in Gödel's original work. Thus, a Godlike entity, by definition, possesses all positive properties ($\mathcal{P}$ is an uninterpreted constant denoting positive properties):

$$\mathcal{G}\, x \quad \equiv \quad \forall Y.(\mathcal{P}\, Y \rightarrow Y\, x)$$

The three only axioms of the new theory which constrain the interpretation of Gödel's positive properties ($\mathcal{P}$) are:

A1' Self-difference is not a positive property.[2]

$$\neg\mathcal{P}(\lambda x.x \neq x)$$

A2' A property entailed or necessarily entailed by a positive property is positive.

$$\forall X\, Y.\, ((\mathcal{P}\, X \wedge (X \sqsubseteq Y \vee X \Rrightarrow Y)) \rightarrow \mathcal{P}\, Y)$$

A3 The conjunction of any collection of positive properties is positive.[3]

$$\forall \mathcal{Z}.(\mathcal{P}os\, \mathcal{Z} \rightarrow \forall X.(X{\textstyle\bigsqcap}\mathcal{Z} \rightarrow \mathcal{P}\, X))$$

(Technical reading: if $\mathcal{Z}$ is any set of positive properties, then the property X obtained by taking the conjunction of the properties in $\mathcal{Z}$ is positive.)

In these premises the following defined symbols are used, where $\forall$ is a possibilist second-order quantifier and where $\forall^{\mathrm{E}}$ is an actualist first-order quantifier for individuals:

$$\begin{aligned}
X \sqsubseteq Y \quad &\equiv \quad \forall^{\mathrm{E}} z.(X\, z \rightarrow Y\, z)\\
X \Rrightarrow Y \quad &\equiv \quad \Box(X \sqsubseteq Y)\\
\mathcal{P}os\, \mathcal{Z} \quad &\equiv \quad \forall X.(\mathcal{Z}\, X \rightarrow \mathcal{P}\, X)\\
X{\textstyle\bigsqcap}\mathcal{Z} \quad &\equiv \quad \Box\forall^{\mathrm{E}} u.(X\, u \leftrightarrow (\forall Y.\, \mathcal{Z}\, Y \rightarrow Y\, u))
\end{aligned}$$

From A1', A2' and A3 it follows, in a few argumentation steps in modal logic **K**, that a Godlike entity possibly and necessarily exists. Modal collapse, which expresses that there are no contingent truths and which thus eliminates the

[1] E.g., the theorem prover LEO-II detected that Gödel's (1970) variant of the argument is inconsistent; this inconsistency had, unknowingly, been fixed in the variant of Scott (1972); cf. (Benzmüller and Woltzenlogel Paleo 2016) for more details.

[2] An alternative to A1' would be: The empty property ($\lambda x.\bot$) is not a positive property.

[3] The third-order formalization of A3 as given here ($\mathcal{Z}$ is a third-order variable ranging over sets of properties), together with the definition of $\mathcal{G}$, implies that being Godlike is a positive property. Since supporting this inference is the only role this axiom plays in the argument, ($\mathcal{P}\,\mathcal{G}$) can be taken—and has been taken, cf. Scott (1972)—as an alternative to our A3; cf. also Fig. 7.

possibility of alternative possible worlds, does not follow from these axioms. Also monotheism is not implied. These observations should render the new theory interesting to theoretical philosophy and theology.

Compare the above with Gödel's premises of the modal ontological argument (in the consistent variant of Scott):

A1 One of a property or its complement is positive.

$$\forall X.(\neg(\mathcal{P}\ X) \leftrightarrow \mathcal{P}(\rightharpoondown X))$$

A2 A property necessarily entailed by a positive property is positive.

$$\forall X\ Y.((\mathcal{P}\ X \wedge (X \Rrightarrow Y)) \rightarrow \mathcal{P}\ Y)$$

A3 The conjunction of any collection of positive properties is positive (or, alternatively, being Godlike is a positive property).

$$\forall \mathcal{Z}.(\mathcal{P}os\ \mathcal{Z} \rightarrow \forall X.(X{\textstyle\prod}\mathcal{Z} \rightarrow \mathcal{P}\ X))$$

A4 Any positive property is necessarily positive.

$$\forall X.(\mathcal{P}\ X \rightarrow \Box(\mathcal{P}\ X))$$

A5 Necessary existence ($\mathcal{NE}$) is a positive property.

$$\mathcal{P}\ \mathcal{NE}$$

Furthermore, axiom B is added to ensure that we operate in logic **KB** instead of just **K**[4] (remember that logic **S5** is not needed):

$$\forall\varphi.(\varphi \rightarrow \Box\Diamond\varphi)$$

Essence ($\mathcal{E}$) and necessary existence ($\mathcal{NE}$) are defined as (other definitions are as before):

$$\begin{aligned}\mathcal{E}\ Y\ x \quad &\equiv \quad Y\ x \wedge \forall Z.(Z\ x \rightarrow (Y \Rrightarrow Z))\\ \mathcal{NE}\ x \quad &\equiv \quad \forall Y.(\mathcal{E}\ Y\ x \rightarrow \Box\exists^{\mathrm{E}}\ Y)\end{aligned}$$

Informally: Property Y is an essence $\mathcal{E}$ of an entity x if, and only if, (i) Y holds for x and (ii) Y necessarily entails every property Z of x. Moreover, an entity x has the property of necessary existence if, and only if, the essence of x is necessarily instantiated.

Using Gödel's premises as stated it can be proved automatically that a Godlike entity possibly and necessarily exists (Benzmüller and Woltzenlogel Paleo 2016). However, modal collapse is still implied even in the weak logic **KB**.[5]

Own prior work recently showed that different modal ultrafilter properties can be deduced from Gödel's premises (Benzmüller and Fuenmayor 2020). These insights are key to the argument simplifications developed and studied in this paper: If Gödel's premises entail that positive properties form a modal ultrafilter, then why not turning things around, and start out with an axiom U1 postulating ultrafilter properties for $\mathcal{P}$? Then use U1 instead of other axioms for proving that a Godlike entity necessarily exists, and on the fly explore what further simplifications of the argument are triggered. This research plan worked out and it led to simplified argument variants as presented above and in the remainder.

The proof assistant Isabelle/HOL and its integrated ATP systems have supported our exploration work surprisingly well, despite the undecidability and high complexity of the underlying logic setting. As usual, we here only present the main steps of the exploration process, and various interesting eureka or frustration steps in between are dropped.

The structure of this paper is as follows: An SSE of HOML in HOL is introduced in Sect. 2. This section, parts of which have been adapted from Kirchner et al. (2019), ensures that the paper is sufficiently self-contained; readers familiar with the SSE approach may simply skip it. Modal filter and ultrafilter are defined in Sect. 3. Section 4 recaps, in some more detail, the Gödel/Scott variant of the modal ontological argument from above. Subsequently, an ultrafilter-based modal ontological argument is presented in Sect. 5. This new argument is further simplified in Sect. 6, leading to our proposal based on axioms A1', A2' and A3 as presented before. Further simplifications and modifications are studied in Sect. 7, and related work is discussed in Sect. 8.

Since we develop, explain and discuss our formal encodings directly in Isabelle/HOL, some familiarity with this proof assistant and its underlying logic HOL (Andrews 2002; Benzmüller and Andrews 2019) is assumed. The entire sources[6] of our formal encodings are presented and explained in detail in this paper.

The contributions of this paper are thus manifold. Besides the novel variants of the modal ontological argument that we contribute to metaphysics and theology, we demonstrate how the SSE technique, in combination with higher-order reasoning tools, can be employed in practical studies to explore new knowledge. Moreover, we contribute useful source encodings that can be reused and adapted to teach quantified modal logics in interdisciplinary lecture courses.

2 Modeling HOML in HOL

Various SSEs of quantified non-classical logics in HOL have been developed, studied and applied in related work, cf. Benzmüller (2019) and Kirchner et al. (2019) and the references therein. These contributions, among others, show that the standard translation from propositional modal logic to first-order (FO) logic can be concisely modeled (i.e., embedded) within HOL theorem provers, so that the modal operator $\Box$, for example, can be explicitly defined by the λ-term $\lambda\varphi.\lambda w.\forall v.(Rwv \rightarrow \varphi v)$, where R denotes the accessibility relation associated with $\Box$. Then one can construct FO formulas involving $\Box\varphi$ and use them to represent and prove theorems. Thus, in an SSE, the target logic is internally represented using higher-order (HO) constructs in a theorem proving system such as Isabelle/HOL. Own prior work developed an SSE that captures quantified extensions

[4]Symmetry of the accessibility relation **r** associated with the modal $\Box$-operator can be postulated alternatively in our meta-logical framework.

[5]For more information on modal collapse (in logic **S5**) consult Sobel (1987; 2004), Fitting (2002) and Kovač (2012); see also the references therein.

[6]See logikey.org → Computational-Metaphysics. The experiments reported in this paper were conducted on a standard notebook (2,5 GHz Intel Core i7, 16 GB memory).

of modal logic (Benzmüller and Paulson 2013). For example, if $\forall x.\phi x$ is shorthand in HOL for $\Pi(\lambda x.\phi x)$, then $\Box\forall x Px$ would be represented as $\Box\Pi'(\lambda x.\lambda w.Pxw)$, where Π' stands for the λ-term $\lambda\Phi.\lambda w.\Pi(\lambda x.\Phi xw)$, and the $\Box$ gets resolved as above.

To see how these expressions can be resolved to produce the right representation, consider the following series of reductions:

$$
\begin{aligned}
& \Box\forall x Px \\
\equiv\ & \Box\Pi'(\lambda x.\lambda w.Pxw) \\
\equiv\ & \Box((\lambda\Phi.\lambda w.\Pi(\lambda x.\Phi xw))(\lambda x.\lambda w.Pxw)) \\
\equiv\ & \Box(\lambda w.\Pi(\lambda x.(\lambda x.\lambda w.Pxw)xw)) \\
\equiv\ & \Box(\lambda w.\Pi(\lambda x.Pxw)) \\
\equiv\ & (\lambda\varphi.\lambda w.\forall v.(Rwv \to \varphi v))(\lambda w.\Pi(\lambda x.Pxw)) \\
\equiv\ & (\lambda\varphi.\lambda w.\Pi(\lambda v.Rwv \to \varphi v))(\lambda w.\Pi(\lambda x.Pxw)) \\
\equiv\ & (\lambda w.\Pi(\lambda v.Rwv \to (\lambda w.\Pi(\lambda x.Pxw))v)) \\
\equiv\ & (\lambda w.\Pi(\lambda v.Rwv \to \Pi(\lambda x.Pxv))) \\
\equiv\ & (\lambda w.\forall v.Rwv \to \forall x.Pxv) \\
\equiv\ & (\lambda w.\forall vx.Rwv \to Pxv)
\end{aligned}
$$

Thus, we end up with a representation of $\Box\forall x Px$ in HOL. Of course, types are assigned to each (sub-)term of the HOL language. We assign individual terms (such as variable x above) the type e, and terms denoting worlds (such as variable w above) the type i. From such base choices, all other types in the above presentation can actually be inferred.

An explicit encoding of HOML in Isabelle/HOL, following the above ideas, is presented in Fig. 1.[7] In lines 4–5 the base types i and e are declared (text passages in red are comments). Note that HOL comes with an inbuilt base type bool, the bivalent type of Booleans. No cardinality constraints are associated with types i and e, except that they must be non-empty. To keep our presentation concise, useful type synonyms are introduced in lines 6–9. σ abbreviates the type i$\Rightarrow$bool ($\Rightarrow$ is the function type constructor in HOL), and terms of type σ can be seen to represent world-lifted propositions, i.e., truth-sets in Kripke's modal relational semantics (Garson 2018). The explicit transition from modal propositions to terms (truth-sets) of type σ is a key aspect of the SSE technique, and in the remainder of this article we use phrases such as "world-lifted" or "σ-type" terms to emphasize this conversion in the SSE approach. γ, which stands for e$\Rightarrow\sigma$, is the type of world-lifted, intensional properties. μ and ν, which abbreviate $\sigma\Rightarrow\sigma$ and $\sigma\Rightarrow\sigma\Rightarrow\sigma$, are the types associated with unary and binary modal logic connectives.

The modal logic connectives are defined in lines 12–25. In line 16, for example, the definition of the world-lifted $\lor$-connective of type ν is given; explicit type information is presented after the ::-token for 'c5', which is the ASCII-denominator for the (right-associative) infix-operator $\lor$ as

[7] In Isabelle/HOL explicit type information can often be omitted due the system's internal type inference mechanism. This feature is exploited in our formalization to improve readability. However, for *new* abbreviations and definitions, we often explicitly declare the types of the freshly introduced symbols. This supports a better intuitive understanding, and it also reduces the number of polymorphic terms in the formalization (heavy use of polymorphism may generally lead to decreased proof automation performance).

```
1  theory HOML imports Main
2  begin
3  (*Type declarations and type abbreviations*)
4  typedecl i (*Possible worlds*)
5  typedecl e (*Individuals*)
6  type_synonym σ = "i⇒bool" (*World-lifted propositions*)
7  type_synonym γ = "e⇒σ" (*Lifted predicates*)
8  type_synonym μ = "σ⇒σ" (*Unary modal connectives*)
9  type_synonym ν = "σ⇒σ⇒σ" (*Binary modal connectives*)
10
11 (*Modal logic connectives (operating on truth-sets)*)
12 abbreviation c1::σ ("⊥") where "⊥ ≡ λw. False"
13 abbreviation c2::σ ("⊤") where "⊤ ≡ λw. True"
14 abbreviation c3::μ ("¬_") where "¬φ ≡ λw.¬(φ w)"
15 abbreviation c4::ν ("_∧_") where "φ∧ψ ≡ λw.(φ w)∧(ψ w)"
16 abbreviation c5::ν ("_∨_") where "φ∨ψ ≡ λw.(φ w)∨(ψ w)"
17 abbreviation c6::ν ("_→_") where "φ→ψ ≡ λw.(φ w)⟶(ψ w)"
18 abbreviation c7::ν ("_↔_") where "φ↔ψ ≡ λw.(φ w)⟷(ψ w)"
19 consts R::"i⇒i⇒bool" ("_r_")  (*Accessibility relation*)
20 abbreviation c8::μ ("□_") where "□φ ≡ λw.∀v.(wrv)⟶(φ v)"
21 abbreviation c9::μ ("◇_") where "◇φ ≡ λw.∃v.(wrv)∧(φ v)"
22 abbreviation c10::"γ⇒γ" ("⇁_")
23                        where "⇁Φ ≡ λx.λw.¬(Φ x w)"
24 abbreviation c11::"e⇒e⇒σ" ("_=_") where "x=y ≡ λw.(x=y)"
25 abbreviation c12::"e⇒e⇒σ" ("_≠_") where "x≠y ≡ λw.(x≠y)"
26
27 (*Polymorphic possibilist quantification*)
28 abbreviation q1::"('a⇒σ)⇒σ" ("∀")
29                                where "∀Φ ≡ λw.∀x.(Φ x w)"
30 abbreviation q2 (binder"∀"[8]9) where "∀x. φ(x) ≡ ∀φ"
31 abbreviation q3::"('a⇒σ)⇒σ" ("∃")
32                                where "∃Φ ≡ λw.∃x.(Φ x w)"
33 abbreviation q4 (binder"∃"[8]9) where "∃x. φ(x) ≡ ∃φ"
34
35 (*Actualist quantification for individuals*)
36 consts existsAt::γ ("_@_")
37 abbreviation q5::"γ⇒σ" ("∀E")
38             where "∀EΦ ≡ λw.∀x.(x@w)⟶(Φ x w)"
39 abbreviation q6 (binder"∀E"[8]9) where "∀Ex. φ(x) ≡ ∀Eφ"
40 abbreviation q7::"γ⇒σ" ("∃E")
41             where "∃EΦ ≡ λw.∃x.(x@w)∧(Φ x w)"
42 abbreviation q8 (binder"∃E"[8]9) where "∃Ex. φ(x) ≡ ∃Eφ"
43
44 (*Meta-logical predicate for global validity*)
45 abbreviation g1::"σ⇒bool" ("⌊_⌋") where "⌊ψ⌋ ≡  ∀w. ψ w"
46
47 (*Consistency, Barcan and converse Barcan formula*)
48 lemma True nitpick[satisfy] oops  (*Model found by Nitpick*)
49 lemma "⌊(∀Ex.□(φ x)) → □(∀Ex.(φ x))⌋" nitpick oops (*Ctm*)
50 lemma "⌊(□(∀Ex.(φ x))) → ∀Ex.□(φ x)⌋" nitpick oops (*Ctm*)
51 lemma "⌊(∀x.□(φ x)) → □(∀x. φ x)⌋" by simp
52 lemma "⌊(□(∀x.(φ x))) → ∀x.□(φ x)⌋" by simp
53
54 (*unimportant*) nitpick_params[user_axioms,show_all]
55 (*unimportant*) declare [[smt_solver=cvc4,smt_oracle]]
56 end
```

Figure 1: SSE of HOML in HOL.

introduced in parenthesis shortly after. $\varphi_\sigma \lor \psi_\sigma$ is then defined as abbreviation for the truth-set $\lambda w.(\varphi_\sigma w) \lor (\psi_\sigma w)$, respectively. In the remainder we generally use bold-face symbols for world-lifted connectives (such as $\boldsymbol{\lor}$) in order to rigorously distinguish them from their ordinary counterparts (such as $\lor$) in meta-logic HOL.

Further modal logic connectives, such as $\boldsymbol{\bot}$, $\boldsymbol{\top}$, $\boldsymbol{\neg}$, $\boldsymbol{\to}$, $\boldsymbol{\leftrightarrow}$, are introduced analogously. The operator $\rightharpoondown$, introduced in lines 22–23, is inverting properties of types γ; this operation occurs in Gödel's axiom A1. $\boldsymbol{=}$ and $\boldsymbol{\neq}$ are defined in lines 24–25 as world-independent, syntactical (in-)equality.

The world-lifted modal □-operator is introduced in lines 19–20; accessibility relation R is now synonymously named **r** in infix notation. The definition of ◇ is analogous.

The world-lifted (polymorphic) possibilist quantifier ∀ as discussed before is introduced in line 28–29. In line 30, user-friendly binder-notation for ∀ is additionally defined. Instead of distinguishing between ∀ and Π′ as in our illustrating example, ∀-symbols are overloaded here. The introduction of the possibilist ∃-quantifier is analogous.

Additional actualist quantifiers, $\forall^E$ and $\exists^E$, are introduced in lines 36–42. Their definition is guarded by an explicit, possibly empty, existsAt (@) predicate, which encodes whether an individual object actually "exists" at a particular world, or not. The actualist quantifiers are declared non-polymorphic, and they support quantification over individuals only. In the remainder we will indeed apply ∀ and ∃ for different types in the type hierarchy of HOL, while $\forall^E$ and $\exists^E$ exclusively quantify over individuals only.

Global validity of a world-lifted formula ψ_σ, denoted as $\lfloor\psi\rfloor$, is introduced in line 45 as an abbreviation for $\forall w_\mathrm{i}.\psi w$.

Consistency of the introduced concepts is confirmed by the model finder *nitpick* (Blanchette and Nipkow 2010) in line 48. Since only abbreviations and no axioms have been introduced so far, the consistency of the Isabelle/HOL theory HOML as displayed in Fig. 1 is actually evident.

In line 49–52 it is studied whether instances of the Barcan and the converse Barcan formulas are implied. As expected, both principles are valid only for possibilist quantification, while they have countermodels for actualist quantification.

Lines 54–55 declare some specific parameter settings for some of the reasoning tools that we employ.

Theorem 1. *The SSE of HOML in HOL is faithful (for **K**).*

Proof. Analogous to (Benzmüller and Paulson 2013). □

Theory HOML thus models base logic **K** in Isabelle/HOL. Axiom B, see above, can be postulated to arrive at logic **KB**.

3 Modal Filter and Ultrafilter

Theory MFilter, for "modal filter", see Fig. 2, imports theory HOML and adapts the topological notions of filter and ultrafilter to our modal logic setting. For an introduction to the notions of filter and ultrafilter see the literature, e.g., (Burris and Sankappanavar 1981) or also (Odifreddi 2000).

Our notion of modal ultrafilter is introduced in lines 20–21 as a world-lifted characteristic function of type $(\gamma{\Rightarrow}\sigma){\Rightarrow}\sigma$. A modal ultrafilter is thus a world-dependent set of intensions of γ-type properties; in other words, a σ-subset of the σ-powerset of γ-type property extensions. A modal ultrafilter ϕ is defined as a modal filter satisfying an additional maximality condition: $\forall\varphi.\varphi \in \phi \vee ({}^{-1}\varphi) \in \phi$, where $\in$ is elementhood of γ-type objects in σ-sets of γ-type objects (see lines 3–4), and where ${}^{-1}$ is the relative set complement operation on sets of entities (line 11).

A modal filter ϕ, see lines 14–17, is required to

1. be large: $\mathbf{U} \in \phi$, where $\mathbf{U}$ denotes the full set of γ-type objects we start with,

```
 1 theory MFilter imports HOML
 2 begin (*Some abbreviations for auxiliary operations*)
 3 abbreviation a::"γ⇒(γ⇒σ)⇒σ" ("_∈_")
 4                                   where "x∈S ≡ S x"
 5 abbreviation b::γ ("∅") where "∅ ≡ λx. ⊥"
 6 abbreviation c::γ ("U") where "U ≡ λx. ⊤"
 7 abbreviation d::"γ⇒γ⇒σ" ("_⊆_")
 8              where "φ⊆ψ ≡ ∀x.((φ x) → (ψ x))"
 9 abbreviation e::"γ⇒γ⇒γ" ("_⊓_")
10              where "φ⊓ψ ≡ λx.((φ x) ∧ (ψ x))"
11 abbreviation f::"γ⇒γ" ("-1_") where "-1ψ ≡ λx. ¬(ψ x)"
12
13 (*Definition of modal filter*)
14 abbreviation g::"(γ⇒σ)⇒σ" ("Filter")
15  where "Filter Φ ≡ (((U∈Φ) ∧ ¬(∅∈Φ))
16          ∧ (∀φ ψ. (((φ∈Φ) ∧ (φ⊆ψ)) → (ψ∈Φ))))
17          ∧ (∀φ ψ. (((φ∈Φ) ∧ (ψ∈Φ)) → ((φ⊓ψ)∈Φ)))"
18
19 (*Definition of modal ultrafilter *)
20 abbreviation h::"(γ⇒σ)⇒σ" ("UFilter") where
21  "UFilter Φ ≡ (Filter Φ)∧(∀φ.((φ∈Φ) ∨ ((-1φ)∈Φ)))"
22
23 (*Modal Filter and ultrafilter are consistent*)
24 lemma "⌊∀Φ φ.((UFilter Φ) → ¬((φ∈Φ) ∧ ((-1φ)∈Φ)))⌋"
25   by force
26 end
```

Figure 2: Definition of filter and ultrafilter (for possible worlds).

2. exclude the empty set: $\emptyset \notin \phi$, where $\emptyset$ is the world-lifted empty set of γ-type objects,
3. be closed under supersets: $\forall\varphi\,\psi.(\varphi \in \phi \wedge \varphi \subseteq \psi) \rightarrow \psi \in \phi$ (the world-lifted $\subseteq$-relation is defined in lines 7–8), and
4. be closed under intersections: $\forall\varphi\,\psi.(\varphi \in \phi \wedge \psi \in \phi) \rightarrow (\varphi \sqcap \psi) \subseteq \phi$ (where $\sqcap$ is defined in lines 9–10).

Own prior work (Benzmüller and Fuenmayor 2020) studied two different notions of modal ultrafilter (termed γ- and δ-ultrafilter), which are defined on intensions and extensions of properties, respectively. This distinction is not needed in this paper; what we call modal ultrafilter here corresponds to our prior notion of γ-ultrafilter.

4 Gödel/Scott Variant

We start out in Fig. 3 with the introduction of some basic abbreviations and definitions for the Gödel/Scott variant of the modal ontological argument. This theory file, which is termed BaseDefs and which imports HOML, is reused without modification also in all other variants as explored in this paper later on. In line 4 the uninterpreted constant symbol $\mathcal{P}$, for "positive properties", is declared; it has type $\gamma{\Rightarrow}\sigma$. $\mathcal{P}$ thus denotes an intensional, world-depended concept. In lines 7–11 abbreviations for the previously discussed relations and predicates $\sqsubseteq$, $\Rrightarrow$, $\prod$ and $\mathcal{P}os$ are introduced. In lines 14–15, Gödel's notion of "being Godlike" ($\mathcal{G}$) is defined, and in lines 18–21 the previously discussed definitions for Essence ($\mathcal{E}$) and Necessary Existence ($\mathcal{NE}$) are given.

The full formalization of Scott's variant of Gödel's argument is presented as theory ScottVariant in Fig. 4. This theory imports and relies on the previously introduced notions from theory files HOML, MFilter and BaseDefs.

The premises of Gödel's argument, as already discussed

```
theory BaseDefs imports HOML
begin
(*Positive properties*)
consts posProp::"γ⇒σ" ("𝒫")

(*Basic definitions for modal ontological argument*)
abbreviation a ("_⊑_") where "X⊑Y ≡ ∀ᴱz.((X z) → (Y z))"
abbreviation b ("_⇛_") where "X⇛Y ≡ □(X⊑Y)"
abbreviation c ("𝒫os") where "𝒫os Z ≡ ∀X.((Z X) → (𝒫 X))"
abbreviation d ("_⨅_")
   where "X⨅𝒵 ≡ □(∀ᴱu.((X u) ↔ (∀Y.((𝒵 Y) → (Y u)))))"

(*Definition of Godlike*)
definition G ("𝒢")
          where "𝒢 x ≡ ∀Y.((𝒫 Y) → (Y x))"

(*Definitions of Essence and Necessary Existence*)
definition E ("ℰ")
          where "ℰ Y x ≡ (Y x) ∧ (∀Z.((Z x) → (Y⇛Z)))"
definition NE ("𝒩ℰ")
          where "𝒩ℰ x ≡ ∀Y.((ℰ Y x) → □(∃ᴱ Y))"
end
```

Figure 3: Definitions for all variants discussed in the remainder.

earlier, are stated in lines 4–10. In line 12 a semantical counterpart B' (symmetry of the accessibility relation **r** associated with the □-operator) of the B axiom is proved.

An abstract level "proof net" for theorem T6, the necessary existence of a Godlike entity, is presented in lines 15–25. Following the literature, the proof goes as follows: From A1 and A2 infer T1: positive properties are possibly exemplified. From A3 and the defn. of $\mathcal{G}$ obtain T2: being Godlike is a positive property (Scott actually directly postulated T2). Using T1 and T2 show T3: possibly a Godlike entity exists. Next, use A1, A4, the defns. of $\mathcal{G}$ and $\mathcal{E}$ to infer T4: being Godlike is an essential property of any Godlike entity. From this, A5, B' and the defns. of $\mathcal{G}$, and $\mathcal{NE}$ have T5: the possible existence of a Godlike entity implies its necessary existence. T5 and T3 then imply T6.

The five subproofs and their dependencies have been automatically proved using ATP systems integrated with Isabelle/HOL via its *sledgehammer* tool (Blanchette et al. 2013); sledgehammer identified and returned the abstract level proof justifications as displayed here, e.g. "using T1 T2 by simp". The mentioned proof engines/tactics blast, metis, and simp are trustworthy components of Isabelle/HOL's, since they internally reconstruct and check each (sub-)proof in the proof assistants small and trusted proof kernel. The smt method, which relies on an external satisfiability modulo solver (CVC4 in our case), is less trusted, but we nevertheless use it here since it was the only Isabelle/HOL method that was able to close this subproof in a single step (we want to avoid displaying longer interactive proofs due to space restrictions). Using the defns. from Sect. 2, one can generally reconstruct and verify all presented proofs with pen and paper directly in meta-logic HOL. Moreover, reconstruction of modal logic proofs from such proof nets within direct proof calculi for quantified modal logics, cf. Kanckos and Woltzenlogel Paleo (2017) or Fitting (2002) is also possible.

The presented theory is consistent, which is confirmed in line 31 by model finder *nitpick*; *nitpick* reports a model consisting of one world and one Godlike entity.

```
1  theory ScottVariant imports HOML MFilter BaseDefs
2  begin
3  (*Axioms of Scott's variant*)
4  axiomatization where
5   A1: "⌊∀X.((¬(𝒫 X)) ↔ (𝒫(⇁X)))⌋" and
6   A2: "⌊∀X Y.(((𝒫 X) ∧ (X⇛Y)) → (𝒫 Y))⌋" and
7   A3: "⌊∀𝒵.((𝒫os 𝒵) → (∀X.((X⨅𝒵) → (𝒫 X))))⌋" and
8   A4: "⌊∀X.((𝒫 X) → □(𝒫 X))⌋" and
9   A5: "⌊𝒫 𝒩ℰ⌋" and
10  B:  "⌊∀φ.(φ → □◇φ)⌋" (*Logic KB*)
11
12 lemma B': "∀x y. ¬(xry) ∨ (yrx)" using B by fastforce
13
14 (*Necessary existence of a Godlike entity*)
15 theorem T6: "⌊□(∃ᴱ 𝒢)⌋"
16 proof -
17  have T1: "⌊∀X.((𝒫 X) → ◇(∃ᴱ X))⌋"
18          using A1 A2 by blast
19  have T2: "⌊𝒫 𝒢⌋" by (metis A3 G_def)
20  have T3: "⌊◇(∃ᴱ 𝒢)⌋" using T1 T2 by simp
21  have T4: "⌊∀ᴱx.((𝒢 x)→(ℰ 𝒢 x))⌋"
22          by (metis A1 A4 G_def E_def)
23  have T5: "⌊(◇(∃ᴱ𝒢))→ □(∃ᴱ𝒢)⌋"
24          by (smt A5 G_def B' NE_def T4)
25  thus ?thesis using T3 by blast qed
26
27 (*Existence of a Godlike entity*)
28 lemma "⌊∃ᴱ 𝒢⌋" using A1 A2 B' T6 by blast
29
30 (*Consistency*)
31 lemma True nitpick[satisfy] oops (*Model found*)
32
33 (*Modal collapse: holds*)
34 lemma MC: "⌊∀Φ.(Φ → □Φ)⌋"
35 proof - {fix w fix Q
36  have 1: "∀x.((𝒢 x w) ⟶
37        (∀Z.((Z x) → □(∀ᴱz.((𝒢 z) → (Z z))))) w)"
38        by (metis A1 A4 G_def)
39  have 2: "(∃x. 𝒢 x w)⟶((Q → □(∀ᴱz.((𝒢 z) → Q))) w)"
40        using 1 by force
41  have 3: "(Q → □Q) w" using B' T6 2 by blast}
42  thus ?thesis by auto qed
43
44 (*Analysis of positive properties using ultrafilters*)
45 theorem U1: "⌊UFilter 𝒫⌋" sledgehammer (*Proof found*)
46 proof -
47  have 1: "⌊(U∈𝒫) ∧ ¬(∅∈𝒫)⌋"
48          using A1 A2 by blast
49  have 2: "⌊∀φ ψ.(((φ∈𝒫)∧(φ⊆ψ))→(ψ∈𝒫))⌋"
50          by (smt A2 B' MC)
51  have 3: "⌊∀φ ψ.(((φ∈𝒫)∧(ψ∈𝒫))→((φ⊓ψ)∈𝒫))⌋"
52          by (metis A1 A2 G_def B' T6)
53  have 4: "⌊∀φ.((φ∈𝒫) ∨ ((⁻¹φ)∈𝒫))⌋"
54          using A1 by blast
55  thus ?thesis using 1 2 3 4 by simp qed
56
57 lemma L1: "⌊∀X Y.((X⇛Y) → (X⊑Y))⌋"
58          by (metis A1 A2 MC)
59 lemma L2: "⌊∀X Y.(((𝒫 X) ∧ (X⊑Y)) → (𝒫 Y))⌋"
60          by (smt A2 B' MC)
61
62 (*Set of supersets of X, we call this HF X*)
63 abbreviation HF where "HF X ≡ λY.(X⊑Y)"
64
65 (*HF 𝒢 is a filter; hence, HF 𝒢 is Hauptfilter of 𝒢*)
66 lemma F1: "⌊Filter (HF 𝒢)⌋" by (metis A2 B' T6 U1)
67 lemma F2: "⌊UFilter (HF 𝒢)⌋" by (smt A1 F1 G_def)
68
69 (*T6 follows directly from F1*)
70 theorem T6again: "⌊□(∃ᴱ 𝒢)⌋" using F1 by simp
71 end
```

Figure 4: Gödel's modal ontological argument; Scott's variant.

```
1  theory UFilterVariant imports HOML MFilter BaseDefs
2  begin
3  (*Axiom's of ultrafilter variant*)
4  axiomatization where
5   U1: "⌊UFilter 𝒫⌋" and
6   A2: "⌊∀X Y.(((𝒫 X) ∧ (X⇛Y)) → (𝒫 Y))⌋" and
7   A3: "⌊∀𝒵.((𝒫os 𝒵) → (∀X.((X⨅𝒵) → (𝒫 X))))⌋"
8
9  (*Necessary existence of a Godlike entity*)
10 theorem T6: "⌊□(∃ᴱ 𝒢)⌋" sledgehammer (*Proof found*)
11 proof -
12  have T1: "⌊∀X.((𝒫 X) → ◇(∃ᴱ X))⌋" by (metis A2 U1)
13  have T2: "⌊𝒫 𝒢⌋" by (metis A3 G_def)
14  have T3: "⌊◇(∃ᴱ 𝒢)⌋" using T1 T2 by simp
15  have T5: "⌊(◇(∃ᴱ 𝒢)) → □(∃ᴱ 𝒢)⌋"
16           by (metis A2 G_def T2 U1)
17  thus ?thesis using T3 by blast qed
18
19 (*Consistency*)
20 lemma True nitpick[satisfy] oops  (*Model found*)
21
22 (*Modal collapse*)
23 lemma MC: "⌊∀Φ.(Φ → □Φ)⌋" nitpick oops (*Countermodel*)
24 end
```

Figure 5: Ultrafilter variant.

Validity of modal collapse (MC) is confirmed in lines 34–42; a proof net displaying the proofs main idea is shown.

Most relevant for this paper is that the ATP systems were able to quickly prove that Gödel's notion of positive properties $\mathcal{P}$ constitutes a modal ultrafilter, cf. lines 45–55. This was key to the idea of taking the modal ultrafilter property of $\mathcal{P}$ as an axiom U1 of the theory; see the next section.

In lines 57–60 some further relevant lemmata are proved. And in line 62–70 we hint at a much simpler, alternative proof argument: Take the set HF$\mathcal{G}$ of all supersets of $\mathcal{G}$; it follows from Gödel's theory that this set is a modal filter (line 66) resp. modal ultrafilter (line 67), i.e., HF$\mathcal{G}$ is the modal Hauptfilter of $\mathcal{G}$. The necessary existence of a Godlike entity now becomes a simple corollary of this result (see line 70, where T6Again is proved *exclusively* from F1). In Sect. 7 we will later present an argument variant that is based on this observation.[8]

5 Ultrafilter Variant

Taking U1 ($\mathcal{P}$ is an ultrafilter) as an axiom for Gödel's theory in fact leads to a significant simplification of the modal ontological argument; this is shown in lines 10–17 of the theory file UFilterVariant in Fig. 5: not only Gödel's axiom A1 can be dropped, but also axioms A4 and A5, together with defns. $\mathcal{E}$ and $\mathcal{NE}$. Even logic **KB** can be given up, since **K** is now sufficient for verifying the proof argument.

The proof is similar to before: Use U1 and A2 to infer T1 (positive properties are possibly exemplified). From A3 and defn. of $\mathcal{G}$ have T2 (being Godlike is a positive property). T1 and T2 imply T3 (a Godlike entity possibly exists). From U1, A2, T2 and the defn. of $\mathcal{G}$ have T5 (possible existence of a Godlike entity implies its necessary existence). Use T5 and T3 to conclude T6 (necessary existence of a Godlike entity).

[8] Manfred Droste from the University of Leipzig pointed me to this proof argument alternative in an email communication.

Consistency of the theory is confirmed in line 20; again a model with one world and one Godlike entity is reported.

Most interestingly, modal collapse MC now has a simple countermodel as *nitpick* informs us in line 23. This countermodel consists of a single entity e_1 and two worlds i_1 and i_2 with accessibility relation $\mathsf{r} = \{\langle i_1, i_1\rangle, \langle i_2, i_1\rangle, \langle i_2, i_2\rangle\}$. Trivially, formula Φ is such that Φ holds in i_2 but not in i_1, which invalidates MC at world i_2. e_1 is the Godlike entity in both worlds, i.e., $\mathcal{G}$ is the property that holds for e_1 in i_1 and i_2, which we may denote as $\lambda e.\lambda w.e{=}e_1 \wedge (w{=}i_1 \vee w{=}i_2)$. Using tuple notation we may write $\mathcal{G} = \{\langle e_1, i_1\rangle, \langle e_1, i_2\rangle\}$.

Remember that $\mathcal{P}$, which is of type $\gamma{\Rightarrow}\sigma$, is an intensional, world-depended concept. In our countermodel for MC in line 23 the extension of $\mathcal{P}$ for world i_1 has the above $\mathcal{G}$ and $\lambda e.\lambda w.e{=}e_1 \wedge w{=}i_1$ as its elements, while in world i_2 we have $\mathcal{G}$ and $\lambda e.\lambda w.e{=}e_1 \wedge w{=}i_2$. Using tuple notation we may note $\mathcal{P}$ as

$$\{\langle\{\langle e_1, i_1\rangle, \langle e_1, i_2\rangle\}, i_1\rangle, \langle\{\langle e_1, i_1\rangle\}, i_1\rangle,\\ \langle\{\langle e_1, i_1\rangle, \langle e_1, i_2\rangle\}, i_2\rangle, \langle\{\langle e_1, i_2\rangle\}, i_2\rangle\}$$

In order to verify that $\mathcal{P}$ is a modal ultrafilter we have to check whether the respective modal ultrafilter conditions are satisfied in both worlds. $\mathbf{U} \in \mathcal{P}$ in i_1 and also in i_2, since both $\langle\{\langle e_1, i_1\rangle, \langle e_1, i_2\rangle\}, i_1\rangle$ and $\langle\{\langle e_1, i_1\rangle, \langle e_1, i_2\rangle\}, i_2\rangle$ are in $\mathcal{P}$; $\emptyset \notin \mathcal{P}$ in i_1 and also in i_2, since both $\langle\{\}, i_1\rangle$ and $\langle\{\}, i_2\rangle$ are not in $\mathcal{P}$. It is also easy to verify that $\mathcal{P}$ is closed under supersets and intersection in both worlds.

Note that in our countermodel for MC, also Gödel's axiom A4 is invalidated. Consider $X = \lambda e.\lambda w.e{=}e_1 \wedge w{=}i_2$, i.e., X is true for e_1 in i_2, but false for e_1 in i_1. We have $\mathcal{P}\,X$ in i_2, but we do not have $\Box(\mathcal{P}\,X)$ in i_1, since $\mathcal{P}\,X$ does not hold in i_1, which is reachable in $\mathbf{r}$ from i_2.

nitpick is capable of computing different *partial* modal ultrafilters as part of its countermodel exploration: out of 512 candidates, *nitpick* identifies 32 structures of form $\langle F, i\rangle$, for $i \in \{i_1, i_2\}$, in which F satisfies the ultrafilter conditions in the specified world i. An example for such an $\langle F, i\rangle$ is

$$\langle\{\langle\{\langle e_1, i_1\rangle\}, i_1\rangle, \langle\{\}, i_1\rangle,\\ \langle\{\langle e_1, i_1\rangle, \langle e_1, i_2\rangle\}, i_2\rangle, \langle\{\langle e_1, i_2\rangle\}, i_2\rangle\}, i_2\rangle$$

F is not a proper modal ultrafilter, since F fails to be an ultrafilter in world i_1.

6 Simplified Variant

What modal ultrafilters properties of $\mathcal{P}$ are actually needed to support T6? Which ones can be dropped? Experiments with our framework, as displayed in theory file SimpleVariant in Fig. 6, confirm that only the filter conditions from Sect. 3 must be upheld for $\mathcal{P}$; maximality can be dropped. However, it is possible to merge filter condition 3 (closed under supersets) for $\mathcal{P}$ with Gödel's A2 into axiom A2' as shown in line 6 of Fig. 6. Moreover, instead of requiring that the empty set $\emptyset = \lambda x.\bot$ must not be a positive property, we postulate that self-difference $\lambda x.x{\neq}x$ is not (line 5); note that self-difference is extensionally equal to $\emptyset$. Self-identity and self-difference have been used frequently in the history of the ontological argument, which is part of the mo-

```
1  theory SimpleVariant imports HOML MFilter BaseDefs
2  begin
3  (*Axiom's of new, simplified variant*)
4  axiomatization where
5   A1': "⌊¬(𝒫(λx.(x≠x)))⌋" and
6   A2': "⌊∀X Y.(((𝒫 X) ∧ ((X⊑Y) ∨ (X⇛Y))) → (𝒫 Y))⌋" and
7   A3:  "⌊∀𝒵.((𝒫os 𝒵) → (∀X.((X⨅𝒵) → (𝒫 X))))⌋"
8
9  lemma T2: "⌊𝒫 𝒢⌋" by (metis A3 G_def)
10 lemma L1: "⌊𝒫(λx.(x=x))⌋" by (metis A3)
11
12 (*Necessary existence of a Godlike entity*)
13 theorem T6: "⌊□(∃ᴱ 𝒢)⌋" sledgehammer (*Proof found*)
14 proof -
15  have T1: "⌊∀X.((𝒫 X) → ◇(∃ᴱ X))⌋" by (metis A1' A2')
16  have T3: "⌊◇(∃ᴱ 𝒢)⌋" using T1 T2 by simp
17  have T5: "⌊(◇(∃ᴱ 𝒢)) → □(∃ᴱ 𝒢)⌋" by (metis A1' A2' T2)
18   thus ?thesis using T3 by blast qed
19
20 (*Modal collapse and monotheism: not implied*)
21 lemma MC: "⌊∀Φ.(Φ → □Φ)⌋" nitpick oops (*Countermodel*)
22 lemma MT: "⌊∀x y.(((𝒢 x) ∧ (𝒢 y)) → (x=y))⌋"
23           nitpick oops (*Countermodel*)
24
25 (*Gödel's A1, A4, A5: not implied anymore*)
26 lemma A1: "⌊∀X.((¬(𝒫 X))↔(𝒫(⇁X)))⌋" nitpick oops (*Ctm*)
27 lemma A4: "⌊∀X.((𝒫 X) → □(𝒫 X))⌋" nitpick oops (*Ctm*)
28 lemma A5: "⌊𝒫 𝒩ℰ⌋" nitpick oops (*Countermodel*)
29
30 (*Checking filter and ultrafilter properties*)
31 theorem F1: "⌊Filter 𝒫⌋" sledgehammer oops (*Proof found*)
32 theorem U1: "⌊UFilter 𝒫⌋" nitpick oops (*Countermodel*)
33
34 lemma True nitpick[satisfy] oops (*Consistency: model found*)
35 end
```

Figure 6: Simplified variant.

tivation for this switch. As intended, filter condition 4 is now implied by the theory (see theorem F1 proved in line 31), as well as positiveness of self-identity (line 10). The essential idea of the theory SimpleVariant in Fig. 6 is to show that it actually suffices, in combination with A3, to postulate that $\mathcal{P}$ is a modal filter, and this is what our simplified axioms do.

From the definition of $\mathcal{G}$ and the axioms A1', A2' and A3 (lines 5–7) theorem T6 immediately follows: in line 13 several theorem provers integrated with *sledgehammer* quickly report a proof ($\leq$ 1sec). Moreover, a more detailed and more intuitive "proof net" is presented in lines 14–18; the proof argument is analogous to what has been discussed before.

In lines 21–23, countermodels for modal collapse MC (similar to the one discussed before) and for monotheism MT are reported.

Further questions are answered experimentally (lines 26–38): neither A1, nor A4 or A5, of the premises we dropped from Gödel's theory are implied anymore, all have countermodels. In lines 31–32 we see that $\mathcal{P}$ is still a filter, but not an ultrafilter. Since some of these axioms, e.g. Gödel's strong A1, have been discussed controversially in the history of Gödel's argument, and since MC and MT are independent, we have arrived at a philosophically and theologically potentially relevant simplification of Gödel's work.

7 Further Simplified Variants

Postulating T2 instead of A3 Instead of working with third-order axiom A3 to infer T2 as in theory SimpleVariant,

```
1  theory SimpleVariantPG imports HOML MFilter BaseDefs
2  begin
3  (*Axiom's of simplified variant with A3 replaced*)
4  axiomatization where
5   A1': "⌊¬(𝒫(λx.(x≠x)))⌋" and
6   A2': "⌊∀X Y.(((𝒫 X) ∧ ((X⊑Y)∨(X⇛Y))) → (𝒫 Y))⌋" and
7   T2:  "⌊𝒫 𝒢⌋"
8
9  (*Necessary existence of a Godlike entity*)
10 theorem T6:  "⌊□(∃ᴱ 𝒢)⌋" sledgehammer (*Proof found*)
11 proof -
12  have T1: "⌊∀X.((𝒫 X) → ◇(∃ᴱ X))⌋" by (metis A1' A2')
13  have T3: "⌊◇(∃ᴱ 𝒢)⌋" using T1 T2 by simp
14  have T5: "⌊(◇(∃ᴱ 𝒢)) → □(∃ᴱ 𝒢)⌋" by (metis A1' A2' T2)
15  thus ?thesis using T3 by blast qed
16
17 lemma True nitpick[satisfy] oops (*Consistency: model found*)
18
19 (*Modal collapse and Monotheism: not implied*)
20 lemma MC: "⌊∀Φ.(Φ → □Φ)⌋" nitpick oops (*Countermodel*)
21 lemma MT: "⌊∀x y.(((𝒢 x)∧(𝒢 y))→(x=y))⌋"
22           nitpick oops (*Countermodel*)
23 end
```

Figure 7: Simplified variant with axiom T2.

```
1  theory SimpleVariantSE imports HOML MFilter BaseDefs
2  begin
3  (*Axiom's of new variant based on ultrafilters*)
4  axiomatization where
5   A1': "⌊¬(𝒫(λx.(x≠x)))⌋" and
6   A2'': "⌊∀X Y.(((𝒫 X) ∧ (X⊑Y)) → (𝒫 Y))⌋" and
7   T2:  "⌊𝒫 𝒢⌋"
8
9  (*Necessary existence of a Godlike entity*)
10 theorem T6: "⌊□(∃ᴱ 𝒢)⌋" using A1' A2'' T2 by blast
11 theorem T7: "⌊∃ᴱ 𝒢⌋" using A1' A2'' T2 by blast
12
13 (*Possible existence of a Godlike: has counterodel*)
14 lemma T3: "⌊◇(∃ᴱ 𝒢)⌋" nitpick oops (*Counterodel*)
15
16 lemma T3': assumes T: "⌊∀φ.((□φ) → φ)⌋"
17            shows "⌊◇(∃ᴱ 𝒢)⌋"
18            using A1' A2'' T2 T by metis
19 end
```

Figure 8: Simplified variant with simple entailment in logic **K**.

we directly postulate T2 as an axiom in theory SimpleVariantPG; cf. the new axiom T2 in line 7 of in Fig. 7.

Theorem T6 can be proved essentially as before (lines 10–15), and MC and MT still have countermodels (lines 20–22).

Simple Entailment in Axiom A2' Instead of using a disjunction of simple entailment and necessary entailment in axiom A2' we may in fact only require simple entailment in A2'; see axiom A2" (line 6) of the theory file SimpleVariantSE displayed in Fig. 8. Proofs for T6, the necessary existence of a Godlike entity, and also T7, existence of a Godlike entity, can still be quickly found (lines 10–11).

However, after replacing A2' by A2", T3 (the possible existence of a Godlike entity) is no longer implied; see line 14. As nitpick informs us, T3 now has an undesired countermodel consisting of one single world that is not connected to itself. By assuming modal axiom T (what is necessary true is true in the given world) this countermodel can be eliminated so that T3 is implied as desired (lines 16–18).

```
1  theory SimpleVariantSEinT imports HOML MFilter BaseDefs
2  begin
3  (*Axiom's of new variant based on ultrafilters*)
4  axiomatization where
5   A1': "⌊¬(𝒫(λx.(x≠x)))⌋" and
6   A2'': "⌊∀X Y.(((𝒫 X) ∧ (X⊑Y)) → (𝒫 Y))⌋" and
7   T2:   "⌊𝒫 𝒢⌋"
8
9  (*Modal Logic T*)
10 axiomatization where T: "⌊∀φ.((□φ) → φ)⌋"
11 lemma T': "⌊∀φ.(φ → (◇φ))⌋" by (metis T)
12
13 (*Necessary existence of a Godlike entity*)
14 theorem T6: "⌊□(∃ᴱ 𝒢)⌋" sledgehammer (*Proof found*)
15 proof -
16  have T1: "⌊∀X.((𝒫 X)→(◇(∃ᴱ X)))⌋" by (metis A1' A2'' T')
17  have T3: "⌊◇(∃ᴱ 𝒢)⌋" by (metis T1 T2)
18  have T5: "⌊(◇(∃ᴱ 𝒢)) → □(∃ᴱ 𝒢)⌋" by (metis A1' A2'' T2)
19  thus ?thesis using T3 by simp qed
20
21 (*T6 again, with an alternative, simpler proof*)
22 theorem T6again: "⌊□(∃ᴱ 𝒢)⌋"
23 proof -
24  have L1: "⌊(∃X.((𝒫 X)∧¬(∃ᴱX)))→(𝒫(λx.(x≠x)))⌋"
25           by (smt A2'')
26  have L2: "⌊¬(∃X.((𝒫 X) ∧ ¬(∃ᴱ X)))⌋" by (metis L1 A1')
27  have T1': "⌊∀X.((𝒫 X) → (∃ᴱ X))⌋" by (metis L2)
28  have T3': "⌊∃ᴱ 𝒢⌋" by (metis T1' T2)
29  have L3: "⌊◇(∃ᴱ 𝒢)⌋" by (metis T3' T') (*not needed*)
30  thus ?thesis using T3' by simp qed
31 end
```

Figure 9: Simplified variant with simple entailment in logic **T**.

```
1  theory SimpleVariantHF imports HOML MFilter BaseDefs
2  begin
3  (*Definition: Set of supersets of X, we call this ℋℱ X*)
4  abbreviation HF::"γ⇒(γ⇒σ)"  where "HF X ≡ λY.(X⊑Y)"
5
6  (*Postulate: ℋℱ 𝒢 is a filter; i.e., Hauptfilter of 𝒢*)
7  axiomatization where F1: "⌊Filter (HF 𝒢)⌋"
8
9  (*Necessary existence of a Godlike entity*)
10 theorem T6: "⌊□(∃ᴱ 𝒢)⌋" using F1 by auto (*Proof found*)
11
12 theorem T6again: "⌊□(∃ᴱ 𝒢)⌋"
13 proof -
14  have T3': "⌊∃ᴱ 𝒢⌋" using F1 by auto
15  have T6:  "⌊□(∃ᴱ 𝒢)⌋" using T3' by blast
16  thus ?thesis by simp qed
17
18 (*Possible existence of Godlike entity not implied*)
19 lemma T3: "⌊◇(∃ᴱ 𝒢)⌋" nitpick oops (*Countermodel*)
20
21 (*Axiom T enforces possible existence of Godlike entity*)
22 axiomatization
23 lemma T3: assumes T: "⌊∀φ.((□φ) → φ)⌋"
24           shows "⌊◇(∃ᴱ 𝒢)⌋"     using F1 T by auto
25
26 lemma True nitpick[satisfy] oops (*Consistency*)
27
28 (*Modal collapse: not implied anymore*)
29 lemma MC: "⌊∀Φ.(Φ → □Φ)⌋" nitpick oops (*Countermodel*)
30 lemma MT: "⌊∀x y.(((𝒢 x) ∧ (𝒢 y)) → (x=y))⌋"
31          nitpick oops (*Countermodel*)
32 end
```

Figure 10: Hauptfiltervariant.

Simple Entailment in Logic T The above discussion motivates a further alternative of the simplified modal ontological argument; see theory file SimpleVariantSEinT in Fig. 9. This argument is assuming modal logic **T** (which comes with axiom T as discussed above), and, as before, it postulates axioms A1' A2" and T2 (lines 5–7):

A1' Self-difference is not a positive property.

A2" A property entailed by a positive property is positive.

T2 Being Godlike is a positive property.

One possible proof argument for T6 is as before; see lines 15–19. However, there is also a much simpler proof, see lines 23–30, which we explain in more detail (this simple proof is applicable to previous variants as well; it is also key to proving T5 from A1', A2'/A2" and T2/A3 in previous variants, including the simplified variant in Fig. 6):[9]

L1 The existence of a non-exemplified positive property implies that self-difference is a positive property—*This follows from axiom A2".*

L2 There is no non-exemplified positive property—*From L1 and axiom A1'.*

T1' Positive properties are exemplified—*Equivalent to L2.*

T3' A Godlike entity exists—*From T1' and axiom T2.*

L3 A Godlike entity possibly exists—*From T3' and T' (contrapositive of axiom T); note that L3 is not needed to obtain T6 in the next step; generally, axiom T (resp. its contrapositive T') is only needed for deriving T3/L3.*

[9]The existence of such an unintended derivation was already hinted at by Fuenmayor and Benzmüller (2017, Footn. 11).

T6 A Godlike entity necessarily exists—*From T3' by necessitation.*

Hauptfiltervariant Another drastically simplified variant of the modal ontological argument is related to the observations discussed earlier at the end of Sec. 4. There it has been shown that the set HF $\mathcal{G}$, consisting of all supersets of $\mathcal{G}$, is a modal filter (so that HF $\mathcal{G}$ is the Hauptfilter of $\mathcal{G}$); this then directly implies the necessary existence of a Godlike entity. A new variant based on this observation is presented in theory file SimpleVariantHF in Fig. 10. Here the set of all supersets HF $\mathcal{G}$ of property $\mathcal{G}$ (being Godlike) is postulated to be a modal filter (axiom F1 in line 7).[10] The existence and necessary existence of a Godlike entity then directly follows from F1, see lines 10–16. And as already observed before, the possible existence of a Godlike entity is independent, but can be enforced by postulating modal axiom T (lines 19–24). Moreover, the theory is consistent (line 26) and neither modal collapse nor monotheism are implied (lines 29–31).

8 Related Work

Fitting (2002) suggested to carefully distinguish between intensions and extensions of positive properties in the context of Gödel's modal ontological argument, and, in order to do so within a single framework, he introduced a sufficiently expressive HOML enhanced with means for the explicit representation of intensional terms and their extensions; cf. also the intensional operations used by Fuenmayor

[10]In fact, the only essential requirement that is enforced here is that $\emptyset$ is not in HF $\mathcal{G}$, hence $\mathcal{G}$ cannot be identical to $\emptyset$.

and Benzmüller (2017; 2020) in their formal study of the works of Fitting and Anderson (1990; 1996).

The application of computational methods to philosophical problems was initially limited to first-order theorem provers. Fitelson and Zalta (2007) used PROVER9 to find a proof of the theorems about situation and world theory in (Zalta 1993) and they found an error in a theorem about Plato's Forms that was left as an exercise in (Pelletier and Zalta 2000). Oppenheimer and Zalta (2011) discovered, using Prover9, that one of the three premises used in their reconstruction of Anselm's ontological argument (Oppenheimer and Zalta 1991) was sufficient to derive the conclusion. The first-order conversion techniques that were developed and applied in these works are outlined in some detail in related work by Alama, Oppenheimer and Zalta (2015).

More recent related work makes use of higher-order proof assistants. Besides the already mentioned own prior work with colleagues, this includes Rushby's (2018) study on Anselm's ontological argument in the PVS system and Blumson's (2017) related study in Isabelle/HOL. Related is also includes Odifreddi's (2000) discussion on ultrafilters, dictators and God, which I was pointed to by a reviewer.

The development of Gödel's ontological argument has recently been addressed by Kanckos and Lethen (2019). They discovered previously unknown variants of the argument in Gödel's Nachlass, whose relation to the presented simplified variants should be further investigated in future work.[11]

9 Discussion

The simplifications of Gödel's theory as presented are far-reaching. In fact, one may ask *"Is this really what Gödel had in mind?"*, or are there some technical issues, such as the alternative proof in lines 22–30 from Fig. 9, that have been ignored so far? Moreover, is the definition of property entailment ($\sqsubseteq$) really adequate in the context of the modal ontological argument, or shouldn't this definition be replaced by concept containment, so that self-difference, resp. the empty property, would no longer $\sqsubseteq$-entail any other property?[12]

Moreover, assuming that the simplified theory SimpleVariant from Fig. 6 is indeed still in line with what Gödel had in mind, why not presenting the definitions and axioms using an alternative wording, for example, as follows (where we replace "positive property" by "rational/consistent property" and "Godlike entity" by "maximally-rational entity"):

$\mathcal{G}$ An entity x is maximally-rational ($\mathcal{G}$) iff it has all rational/consistent properties.

A1' Self-difference is not a rational/consistent property.

A2' A property entailed or necessarily entailed by a rational/consistent property is rational/consistent.

T2 Maximal-rationality is a rational/consistent property.

It follows: A maximally-rational entity necessarily exists.

It would still be possible, but not mandatory, to understand a maximally-rational being also as a Godlike being.

Independent of this discussion we expect the Isabelle/HOL theory files we contributed to be useful for teaching quantified modal logics in classroom, as previously demonstrated in our awarded lecture course on "Computational Metaphysics" (Wisniewski et al. 2016). The developed corpus of example problems is furthermore suited as a benchmark for other ambitious knowledge representation and reasoning projects in the KR community: Can the alternative approaches encode such metaphysical arguments as well? How about their proof automation capabilities and how about model finding? Description logics or argumentation theory, for example, due to their limited expressivity, appear unsuited to support such ambitious applications, while the techniques presented here demonstrably scale for the encoding and assessment of other ambitious theories in metaphysics (Kirchner, Benzmüller, and Zalta 2020) and mathematics (Benzmüller and Scott 2020; Tiemens et al. 2020). Moreover, our problem set constitutes an interesting benchmark for other HOL automated theorem provers and should therefore be converted into TPTP THF representation (Sutcliffe and Benzmüller 2010) and be used in theorem prover competitions.

10 Conclusion

Gödel's modal ontological argument stands in prominent tradition of western philosophy. It has its roots in the Proslogion of Anselm of Canterbury and it has been picked up in Descartes' Fifth Meditation and in the works of Leibniz, which in turn inspired and informed the work of Gödel.

In this paper we have linked Gödel's theory to a suitably adapted mathematical theory (modal filter and modal ultrafilter), and subsequently we have developed simplified modal ontological arguments which avoid some of Gödel's axioms and consequences, including modal collapse, that have led to criticism in the past. At the same time the offered simplifications are very far reaching, eventually too far. Anyhow, the insights that were presented in this paper appear relevant to inform ongoing studies of the modal ontological argument in theoretical philosophy and theology.

We have in this paper applied modern symbolic AI techniques to arrive at deep, explainable and verifiable models of the metaphysical concepts we are interested in. In particular, we have illustrated how state of the art theorem proving systems, in combination with latest knowledge representation and reasoning technology, can fruitfully be employed to explore and contribute deep new knowledge to other disciplines.

[11]An email discussion (March, 2020) with Tim Lethen revealed the following: In particular version No. 2 of Gödel's argument as presented in (Kanckos and Lethen 2019) appears related, though not equivalent, to our simplified versions. Version No. 2 –which we have meanwhile formalized and verified in Isabelle/HOL – avoids the notions of essence and necessary existence and associated definitions/axioms, just as our simplified versions do. Moreover, their findings also suggest that instead of axiom A1' we may just postulate that a non-positive property exists (and experiments confirm this claim; A1' is then implied). However, I prefer axiom A1'.

[12]See also (Alama, Oppenheimer, and Zalta 2015) and (Fuenmayor and Benzmüller 2017, Footn. 11).

Acknowledgements

I am grateful to friends and colleagues who have supported this line of research in the past and/or directly contributed with comments and suggestions, including (in alphabetical order) C. Brown, M. Droste, D. Fuenmayor, T. Gleißner, D. Kirchner, S. Kovač, T. Lethen, X. Parent, R. Rojas, D. Scott, A. Steen, K. Świętorzecka, L. van der Torre, E. Weydert, D. Streit, B. Woltzenlogel Paleo, E. Zalta. Moreover, I am grateful to the reviewers of this paper for useful comments and for the pointer to P. Odifreddi's paper.

References

Alama, J.; Oppenheimer, P. E.; and Zalta, E. N. 2015. Automating leibniz's theory of concepts. In *Automated Deduction – CADE-25*, volume 9195 of *LNCS*, 73–97. Springer.

Anderson, C. A., and Gettings, M. 1996. Gödel's ontological proof revisited. In *Gödel'96: Logical Foundations of Mathematics, Computer Science, and Physics: Lecture Notes in Logic 6*. Springer. 167–172.

Anderson, C. A. 1990. Some emendations of Gödel's ontological proof. *Faith and Philosophy* 7(3):291–303.

Andrews, P. B. 2002. *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof*, volume 27 of *Applied Logic Series*. Springer.

Benzmüller, C., and Andrews, P. 2019. Church's type theory. In Zalta, E. N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2019 edition.

Benzmüller, C., and Fuenmayor, D. 2020. Computer-supported analysis of positive properties, ultrafilters and modal collapse in variants of Gödel's ontological argument. *Bulletin of the Section of Logic* 49(2):127–148.

Benzmüller, C., and Paulson, L. 2013. Quantified multimodal logics in simple type theory. *Logica Universalis* 7(1):7–20.

Benzmüller, C., and Scott, D. S. 2020. Automating free logic in HOL, with an experimental application in category theory. *Journal of Automated Reasoning* 64(1):53–72.

Benzmüller, C., and Woltzenlogel Paleo, B. 2014. Automating Gödel's ontological proof of God's existence with higher-order automated theorem provers. In *ECAI 2014*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, 93–98. IOS Press.

Benzmüller, C., and Woltzenlogel Paleo, B. 2016. The inconsistency in Gödel's ontological argument: A success story for AI in metaphysics. In Kambhampati, S., ed., *IJCAI 2016*, volume 1-3, 936–942. AAAI Press.

Benzmüller, C. 2019. Universal (meta-)logical reasoning: Recent successes. *Science of Computer Programming* 172:48–62.

Blanchette, J. C., and Nipkow, T. 2010. Nitpick: A counterexample generator for higher-order logic based on a relational model finder. In Kaufmann, M., and Paulson, L. C., eds., *Interactive Theorem Proving — ITP 2010*, volume 6172 of *LNCS*, 131–146. Springer.

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

Blumson, B. 2017. Anselm's God in Isabelle/HOL. *Archive of Formal Proofs*.

Burris, S., and Sankappanavar, H. 1981. *A Course in Universal Algebra*. Springer.

Fitelson, B., and Zalta, E. N. 2007. Steps toward a computational metaphysics. *Journal Philosophical Logic* 36(2):227–247.

Fitting, M. 2002. *Types, Tableaus, and Gödel's God*. Kluwer.

Fuenmayor, D., and Benzmüller, C. 2017. Automating emendations of the ontological argument in intensional higher-order modal logic. In *KI 2017: Advances in Artificial Intelligence*, volume 10505 of *LNAI*, 114–127. Springer.

Garson, J. 2018. Modal logic. In Zalta, E. N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition.

Gödel, K. 1970. Appendix A. Notes in Kurt Gödel's Hand. In Sobel (2004). 144–145.

Kanckos, A., and Lethen, T. 2019. The development of Gödel's ontological proof. *The Review of Symbolic Logic*.

Kanckos, A., and Woltzenlogel Paleo, B. 2017. Variants of Gödel's ontological proof in a natural deduction calculus. *Studia Logica* 105:553–586.

Kirchner, D.; Benzmüller, C.; and Zalta, E. N. 2019. Computer science and metaphysics: A cross-fertilization. *Open Philosophy* 2:230–251.

Kirchner, D.; Benzmüller, C.; and Zalta, E. N. 2020. Mechanizing principia logico-metaphysica in functional type theory. *Review of Symbolic Logic* 13(1):206–218.

Kovač, S. 2012. Modal collapse in Gödel's ontological proof. In Szatkowski, M., ed., *Ontological Proofs Today*. Ontos Verlag. 50–323.

Nipkow, T.; Paulson, L. C.; and Wenzel, M. 2002. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer.

Odifreddi, P. 2000. Ultrafilters, dictators, and gods. In Calude, C., and Paun, G., eds., *Finite Versus Infinite. Discrete Mathematics and Theoretical Computer Science*. Springer, London. 239–245.

Oppenheimer, P. E., and Zalta, E. N. 1991. On the logic of the ontological argument. *Philosophical Perspectives* 5:509–529.

Oppenheimer, P. E., and Zalta, E. N. 2011. A computationally-discovered simplification of the ontological argument. *Australasian Journal of Philosophy* 89(2):333–349.

Pelletier, F. J., and Zalta, E. N. 2000. How to Say Goodbye to the Third Man. *Noûs* 34(2):165–202.

Rushby, J. 2018. A mechanically assisted examination of begging the question in Anselm's ontological argument. *J. Applied Logics* 5(7):1473–1496.

Scott, D. S. 1972. Appendix B: Notes in Dana Scott's Hand. In Sobel (2004). 145–146.

Sobel, J. H. 1987. Gödel's ontological proof. In Tomson, J. J., ed., *On Being and Saying. Essays for Richard Cartwright*. MIT Press. 241–261.

Sobel, J. H. 2004. *Logic and Theism: Arguments for and Against Beliefs in God*. Cambridge University Press.

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

Tiemens, L.; Scott, D. S.; Benzmüller, C.; and Benda, M. 2020. Computer-supported exploration of a categorical axiomatization of modeloids. In *Relational and Algebraic Methods in Computer Science – 18th International Conference, RAMiCS 2020*, volume 12062 of *LNCS*, 302–317. Springer.

Wisniewski, M.; Steen, A.; and Benzmüller, C. 2016. Einsatz von Theorembeweisern in der Lehre. In Schwill, A., and Lucke, U., eds., *Hochschuldidaktik der Informatik: 7. Fachtagung des GI-Fachbereichs Informatik und Ausbildung/Didaktik der Informatik*, Commentarii informaticae didacticae (CID), 81–92. Universitätsverlag Potsdam.

Zalta, E. N. 1993. Twenty-Five Basic Theorems in Situation and World Theory. *Journal of Philosophical Logic* 22(4):385–428.