Computer-Formalizations of the Ontological Argument Using Intensional Higher-Order Modal Logic

David Fuenmayor[1] and Christoph Benzmüller[2,1]

[1]Freie Universität Berlin, Germany
[2]University of Luxembourg, Luxembourg

May 3, 2017

Abstract

A computer-formalization in Isabelle/HOL of several variants of Gödel's ontological argument is presented (as discussed in M. Fitting's textbook *Types, Tableaus and Gödel's God*). Fitting's work introduces an intensional higher-order modal logic (by drawing on Montague/Gallin approach), which we shallowly embed here in classical higher-order logic (Isabelle/HOL). We then utilize the embedded logic for the formalization of the ontological argument. In particular, Fitting's and Anderson's variants are verified and their claims confirmed. These variants aim to avoid the modal collapse, which has been criticized as an undesirable side-effect of Kurt Gödel's (and Dana Scott's) versions of the ontological argument.

Keywords: Automated Theorem Proving. Computational Metaphysics. Isabelle. Modal Logic. Intensional Logic. Ontological Argument

1 Introduction

We present a shallow semantical embedding of an *intensional* higher-order modal logic (IHOML) in Isabelle/HOL which has been introduced Fitting in his textbook *Types, Tableaus and Gödel's God* [12] in order to formalize his emendation of Gödel's ontological argument (for the existence of God). IHOML is a modification of the intentional logic originally developed by Montague and later expanded by Gallin [14] by building upon Church's type theory and Kripke's possible-world semantics. Our approach has been inspired by previous work on the semantical embedding of multimodal logics with quantification [6], which we expand here to allow for actualist quantification, intensional terms and their related operations.

We subsequently present a study on Computational Metaphysics: a computer-formalization and verification of Gödel's [15] (resp. Dana Scott's [18]) modern variant of the ontological argument, followed by Fitting's emendation thereof. A third variant (by Anderson [2]) is also discussed. The motivation is to avoid the *modal collapse* [19, 20], which has been criticized as an undesirable side-effect of the axioms of Gödel (resp. Scott). The modal collapse essentially states that there are no contingent truths and that everything is determined. Several authors (e.g. [2, 1, 16, 10]) have proposed emendations of the argument with the aim of maintaining the essential result (the necessary existence of God) while at the same time avoiding the modal collapse. Related work has formalized several of these variants on the computer and verified or falsified them. For example, Gödel's axioms [15] have been shown inconsistent [8, 9] while Scott's version has been verified [5]. Further experiments, contributing amongst others to the clarification of a related debate between Hájek and Anderson, are presented and discussed in [6]. The enabling technique in all of these experiments has been shallow semantical embeddings of (extensional) higher-order modal logics in classical higher-order logic (see [6, 3] and the references therein).

Fitting's emendation also intends to avoid the modal collapse. However, in contrast to the above variants, Fitting's solution is based on the use of an intensional as opposed to an extensional higher-order modal logic. For our work this imposed the additional challenge to provide a shallow embedding of this more advanced logic. The experiments presented below confirm that Fitting's argument as presented in his textbook [12] is valid and that it avoids the modal collapse as intended. The work presented here originates from the *Computational Metaphysics* lecture course held at FU Berlin in Summer 2016 [7].

2 Embedding of Intensional Higher-Order Modal Logic

2.1 Type Declarations

Since IHOML and Isabelle/HOL are both typed languages, we introduce a type-mapping between them. We follow as closely as possible the syntax given by Fitting (see p. 86). According to this syntax, if τ is an extensional type, $\uparrow\tau$ is the corresponding intensional type. For instance, a set of (red) objects has the extensional type $\langle\mathbf{0}\rangle$, whereas the concept 'red' has intensional type $\uparrow\langle\mathbf{0}\rangle$.

typedecl i — type for possible worlds

type-synonym io = (i⇒$bool$) — formulas with world-dependent truth-value
typedecl e (**0**) — individual objects

Aliases for common complex types (predicates and relations):

type-synonym ie=(i⇒**0**) (↑**0**) — individual concepts map worlds to objects
type-synonym se=(**0**⇒$bool$) (⟨**0**⟩) — (extensional) sets
type-synonym ise=(**0**⇒io) (↑⟨**0**⟩) — intensional (predicate) concepts
type-synonym $sise$=(↑⟨**0**⟩⇒$bool$) (⟨↑⟨**0**⟩⟩) — sets of concepts
type-synonym $isise$=(↑⟨**0**⟩⇒io) (↑⟨↑⟨**0**⟩⟩) — 2nd-order intensional concepts **type-synonym** see=(**0**⇒**0**⇒$bool$) (⟨**0**,**0**⟩) — (extensional) relations
type-synonym $isee$=(**0**⇒**0**⇒io) (↑⟨**0**,**0**⟩) — intensional relational concepts
type-synonym $isisee$=(↑⟨**0**⟩⇒**0**⇒io) (↑⟨↑⟨**0**⟩,**0**⟩) — 2nd-order intensional relation

2.2 Logical Constants as Truth-Sets

We embed each modal operator as the set of worlds satisfying the corresponding HOL formula.

abbreviation $mnot$ :: io⇒io (¬-[52]53)
  where $\neg\varphi \equiv \lambda w.\ \neg(\varphi\ w)$
abbreviation $mand$ :: io⇒io⇒io (**infixr**∧51)
  where $\varphi\wedge\psi \equiv \lambda w.\ (\varphi\ w)\wedge(\psi\ w)$
abbreviation mor :: io⇒io⇒io (**infixr**∨50)
  where $\varphi\vee\psi \equiv \lambda w.\ (\varphi\ w)\vee(\psi\ w)$
abbreviation $mimp$ :: io⇒io⇒io (**infixr**→49)
  where $\varphi\rightarrow\psi \equiv \lambda w.\ (\varphi\ w)\longrightarrow(\psi\ w)$

Following can be seen as modelling *possibilist quantification*:

abbreviation $mforall$::($'t$⇒io)⇒io (∀) **where** $\forall\,\Phi \equiv \lambda w.\forall\,x.\ (\Phi\ x\ w)$
abbreviation $mexists$::($'t$⇒io)⇒io (∃) **where** $\exists\,\Phi \equiv \lambda w.\exists\,x.\ (\Phi\ x\ w)$

The *existsAt* predicate is used to embed actualist quantifiers by restricting the domain of quantification at every possible world. This standard technique has been referred to as *existence relativization* ([13], p. 106), highlighting the fact that this predicate can be seen as a kind of meta-logical 'existence predicate' telling us which individuals *actually* exist at a given world. This meta-logical concept does not appear in our object language.

consts $ExistsAt$::↑⟨**0**⟩ (**infix** $existsAt$ 70)

abbreviation $mforallAct$::↑⟨↑⟨**0**⟩⟩ ($\forall^E$) — actualist variants use superscript!
  where $\forall^E\Phi \equiv \lambda w.\forall\,x.\ (x\ existsAt\ w)\longrightarrow(\Phi\ x\ w)$
abbreviation $mexistsAct$::↑⟨↑⟨**0**⟩⟩ ($\exists^E$)
  where $\exists^E\Phi \equiv \lambda w.\exists\,x.\ (x\ existsAt\ w) \wedge (\Phi\ x\ w)$

aRel is the frame's *accessibility relation* (aliased r) used to embed the modal operators □ and ◇.

consts $aRel$::i⇒i⇒$bool$ (**infixr** r 70)
abbreviation $mbox$ :: io⇒io (□-[52]53) **where** $\Box\varphi \equiv \lambda w.\forall\,v.\ (w\ r\ v)\longrightarrow(\varphi\ v)$
abbreviation $mdia$ :: io⇒io (◇-[52]53) **where** $\Diamond\varphi \equiv \lambda w.\exists\,v.\ (w\ r\ v)\wedge(\varphi\ v)$

2.3 *Extension-of* Operator

According to Fitting's semantics ([12], pp. 92-4) $\downarrow$ is an unary operator applying only to intensional terms. A term of the form $\downarrow\alpha$ designates the extension of the intensional object designated by α, at some *given* world. For instance, suppose we take possible worlds as persons, we can therefore think of the concept 'red' as a function that maps each person to the set of objects that person classifies as red (its extension). We can further state, the intensional term r of type $\uparrow\langle\mathbf{0}\rangle$ designates the concept 'red'. As can be seen, intensional terms in IHOML designate functions on possible worlds and they always do it *rigidly*. We will sometimes refer to an intensional object explicitly as 'rigid', implying that its (rigidly) designated function has the same extension in all possible worlds. (The notion of rigidity was introduced by Kripke in [17], where he discusses its interesting philosophical ramifications at some length.)

Terms of the form $\downarrow\alpha$ are called *relativized* (extensional) terms; they are always derived from intensional terms and their type is *extensional* (in the color example $\downarrow r$ would be of type $\langle\mathbf{0}\rangle$). Relativized terms may vary their denotation from world to world of a model, because the extension of an intensional term can change from world to world, i.e. they are non-rigid.

To recap: an intensional term denotes the same function in all worlds (i.e. it's rigid), whereas a relativized term denotes a (possibly) different extension (an object or a set) at every world (i.e. it's non-rigid). To find out the denotation of a relativized term, a world must be given. Relativized terms are the *only* non-rigid terms.

For our Isabelle/HOL embedding, we had to follow a slightly different approach; we model $\downarrow$ as a predicate applying to formulas of the form $\Phi(\downarrow\alpha_1,\ldots\alpha_n)$ (for our treatment we only need to consider cases involving one or two arguments, the first one being a relativized term). For instance, the formula $Q(\downarrow a_1)^w$ (evaluated at world w) is modelled as $\downharpoonleft(Q,a_1)^w$ (or $(Q \downharpoonleft a_1)^w$ using infix notation), which gets further translated into $Q(a_1(w))^w$.

(a) Predicate φ takes as argument a relativized term derived from an (intensional) individual of type $\uparrow\mathbf{0}$:

abbreviation *extIndivArg*::$\uparrow\langle\mathbf{0}\rangle\Rightarrow\uparrow\mathbf{0}\Rightarrow io$ (**infix** $\downharpoonleft$ *60*)
  where $\varphi \downharpoonleft c \equiv \lambda w.\ \varphi\ (c\ w)\ w$

(b) A variant of (a) for terms derived from predicates (types of form $\uparrow\langle t\rangle$):

abbreviation *extPredArg*::$(('t\Rightarrow bool)\Rightarrow io)\Rightarrow('t\Rightarrow io)\Rightarrow io$ (**infix** $\downarrow$ *60*)
  where $\varphi \downarrow P \equiv \lambda w.\ \varphi\ (\lambda x.\ P\ x\ w)\ w$

2.4 Equality

abbreviation *meq* :: $'t\Rightarrow 't\Rightarrow io$ (**infix**$\approx$*60*) — normal equality (for all types)

where $x \approx y \equiv \lambda w.\ x = y$
abbreviation *meqC* :: $\uparrow\langle\uparrow\mathbf{0},\uparrow\mathbf{0}\rangle$ (**infixr**$\approx^C$ *52*) — eq. for individual concepts
where $x \approx^C y \equiv \lambda w.\ \forall v.\ (x\ v) = (y\ v)$
abbreviation *meqL* :: $\uparrow\langle\mathbf{0},\mathbf{0}\rangle$ (**infixr**$\approx^L$ *52*) — Leibniz eq. for individuals
where $x \approx^L y \equiv \forall \varphi.\ \varphi(x) \rightarrow \varphi(y)$

2.5 Verifying the Embedding

The above definitions introduce modal logic K with possibilist and actualist quantifiers, as evidenced by the following tests:

abbreviation *valid*::*io*$\Rightarrow$*bool* ($\lfloor$-$\rfloor$) **where** $\lfloor\psi\rfloor \equiv \forall w.(\psi\ w)$ — modal validity

Verifying K principle and the *necessitation* rule:

lemma *K*: $\lfloor(\Box(\varphi \rightarrow \psi)) \rightarrow (\Box\varphi \rightarrow \Box\psi)\rfloor$ **by** *simp* — K schema
lemma *NEC*: $\lfloor\varphi\rfloor \Longrightarrow \lfloor\Box\varphi\rfloor$ **by** *simp* — necessitation

Local consequence implies global consequence (but not the other way round!):

lemma *localImpGlobalCons*: $\lfloor\varphi \rightarrow \xi\rfloor \Longrightarrow \lfloor\varphi\rfloor \longrightarrow \lfloor\xi\rfloor$ **by** *simp*
lemma $\lfloor\varphi\rfloor \longrightarrow \lfloor\xi\rfloor \Longrightarrow \lfloor\varphi \rightarrow \xi\rfloor$ **nitpick oops** — countersatisfiable

(Converse-)Barcan formulas are satisfied for possibilist, but not for actualist, quantifiers:

lemma $\lfloor(\forall x.\Box(\varphi\ x)) \rightarrow \Box(\forall x.(\varphi\ x))\rfloor$ **by** *simp*
lemma $\lfloor\Box(\forall x.(\varphi\ x)) \rightarrow (\forall x.\Box(\varphi\ x))\rfloor$ **by** *simp*
lemma $\lfloor(\forall^E x.\Box(\varphi\ x)) \rightarrow \Box(\forall^E x.(\varphi\ x))\rfloor$ **nitpick oops** — countersatisfiable
lemma $\lfloor\Box(\forall^E x.(\varphi\ x)) \rightarrow (\forall^E x.\Box(\varphi\ x))\rfloor$ **nitpick oops** — countersatisfiable

We have made use of (counter-)model finder *Nitpick* [11] for the first time. For all the conjectured lemmas above, *Nitpick* has found a countermodel, i.e. a model satisfying all the axioms which falsifies the given formula. This means, the formulas are not valid.

$\beta\eta$-redex is valid for non-relativized (intensional or extensional) terms:

lemma $\lfloor((\lambda\alpha.\ \varphi\ \alpha)\ \ (\tau::\uparrow\mathbf{0})) \leftrightarrow (\varphi\ \ \tau)\rfloor$ **by** *simp*
lemma $\lfloor((\lambda\alpha.\ \varphi\ \alpha)\ \ (\tau::\mathbf{0})) \leftrightarrow (\varphi\ \ \tau)\rfloor$ **by** *simp*
lemma $\lfloor((\lambda\alpha.\ \Box\varphi\ \alpha)\ (\tau::\uparrow\mathbf{0})) \leftrightarrow (\Box\varphi\ \tau)\rfloor$ **by** *simp*
lemma $\lfloor((\lambda\alpha.\ \Box\varphi\ \alpha)\ (\tau::\mathbf{0})) \leftrightarrow (\Box\varphi\ \tau)\rfloor$ **by** *simp*

$\beta\eta$-redex is valid for relativized terms as long as no modal operators occur inside the predicate abstract:

lemma $\lfloor((\lambda\alpha.\ \varphi\ \alpha)\ \downharpoonleft(\tau::\uparrow\mathbf{0})) \leftrightarrow (\varphi\ \downharpoonleft\tau)\rfloor$ **by** *simp*
lemma $\lfloor((\lambda\alpha.\ \Box\varphi\ \alpha)\ \downharpoonleft(\tau::\uparrow\mathbf{0})) \leftrightarrow (\Box\varphi\ \downharpoonleft\tau)\rfloor$ **nitpick oops** — countersatisfiable
lemma $\lfloor((\lambda\alpha.\ \Diamond\varphi\ \alpha)\ \downharpoonleft(\tau::\uparrow\mathbf{0})) \leftrightarrow (\Diamond\varphi\ \downharpoonleft\tau)\rfloor$ **nitpick oops** — countersatisfiable

Modal collapse is countersatisfiable:

lemma $\lfloor\varphi \rightarrow \Box\varphi\rfloor$ **nitpick oops** — countersatisfiable

2.6 Useful Definitions for Axiomatization of Further Logics

The best known normal logics (*K4, K5, KB, K45, KB5, D, D4, D5, D45, ...*) can be obtained by combinations of the following axioms:

abbreviation M **where** $M \equiv \forall\, \varphi.\ \Box\varphi \rightarrow \varphi$
abbreviation B **where** $B \equiv \forall\, \varphi.\ \varphi \rightarrow \Box\Diamond\varphi$
abbreviation D **where** $D \equiv \forall\, \varphi.\ \Box\varphi \rightarrow \Diamond\varphi$
abbreviation IV **where** $IV \equiv \forall\, \varphi.\ \Box\varphi \rightarrow \Box\Box\varphi$
abbreviation V **where** $V \equiv \forall\, \varphi.\ \Diamond\varphi \rightarrow \Box\Diamond\varphi$

Instead of postulating (combinations of) the above axioms we instead make use of the well-known *Sahlqvist correspondence*, which links axioms to constraints on a model's accessibility relation (e.g. reflexive, symmetric, etc). We show that reflexivity, symmetry, seriality, transitivity and euclideanness imply axioms M, B, D, IV, V respectively.

lemma *reflexive aRel* $\Longrightarrow \lfloor M \rfloor$ **by** *blast* — aka T
lemma *symmetric aRel* $\Longrightarrow \lfloor B \rfloor$ **by** *blast*
lemma *serial aRel* $\Longrightarrow \lfloor D \rfloor$ **by** *blast*
lemma *transitive aRel* $\Longrightarrow \lfloor IV \rfloor$ **by** *blast*
lemma *euclidean aRel* $\Longrightarrow \lfloor V \rfloor$ **by** *blast*
lemma *preorder aRel* $\Longrightarrow \lfloor M \rfloor \wedge \lfloor IV \rfloor$ **by** *blast* — S4: reflexive + transitive
lemma *equivalence aRel* $\Longrightarrow \lfloor M \rfloor \wedge \lfloor V \rfloor$ **by** *blast* — S5: preorder + symmetric

lemma *reflexive aRel* $\wedge$ *euclidean aRel* $\Longrightarrow \lfloor M \rfloor \wedge \lfloor V \rfloor$ **by** *blast* — S5

Using these definitions, we can derive axioms for the most common modal logics (see also [4]). Thereby we are free to use either the semantic constraints or the related *Sahlqvist* axioms. Here we provide both versions. In what follows we use the semantic constraints (for improved performance).

2.7 Textbook Examples

In this section we provide further evidence that our embedded logic works as intended by proving the examples discussed in Fitting's textbook [12]. We were able to confirm that all results agree with his claims.

Example 7.13, p. 96:

lemma $\lfloor(\lambda X.\ \Diamond\exists\, X)\ \ (P{::}\uparrow\langle\mathbf{0}\rangle) \rightarrow \Diamond((\lambda X.\ \exists\, X)\ \ P)\rfloor$ **by** *simp*
lemma $\lfloor(\lambda X.\ \Diamond\exists\, X) \downarrow(P{::}\uparrow\langle\mathbf{0}\rangle) \rightarrow \Diamond((\lambda X.\ \exists\, X) \downarrow P)\rfloor$
nitpick[*card 't=1*, *card i=2*] **oops** — nitpick finds same counterexample as book

Example 7.14, p. 98:

lemma $\lfloor(\lambda X.\ \Diamond\exists\, X) \downarrow(P{::}\uparrow\langle\mathbf{0}\rangle) \rightarrow (\lambda X.\ \exists\, X) \downarrow P\rfloor$ **by** *simp*
lemma $\lfloor(\lambda X.\ \Diamond\exists\, X)\ \ (P{::}\uparrow\langle\mathbf{0}\rangle) \rightarrow (\lambda X.\ \exists\, X)\ \ P\rfloor$
nitpick[*card 't=1*, *card i=2*] **oops** — countersatisfiable

Example 7.15, p. 99:

lemma $\lfloor\Box(P\ \ (c{::}\uparrow\mathbf{0})) \to (\exists\, x{::}\uparrow\mathbf{0}.\ \Box(P\ x))\rfloor$ **by** *auto*

Example 7.16, p. 100:

lemma $\lfloor\Box(P\ \downharpoonleft(c{::}\uparrow\mathbf{0})) \to (\exists\, x{::}\mathbf{0}.\ \Box(P\ x))\rfloor$
nitpick[*card 't=2*, *card i=2*] **oops** — counterexample with two worlds found

Example 7.17, p. 101:

lemma $\lfloor\forall\, Z{::}\uparrow\mathbf{0}.\ (\lambda x{::}\mathbf{0}.\ \ \Box((\lambda y{::}\mathbf{0}.\ \ x \approx y)\ \downharpoonleft Z))\ \downharpoonleft Z\rfloor$
nitpick[*card 't=2*, *card i=2*] **oops** — countersatisfiable
lemma $\lfloor\forall\, z{::}\mathbf{0}.\ \ (\lambda x{::}\mathbf{0}.\ \ \Box((\lambda y{::}\mathbf{0}.\ \ x \approx y)\ \ z))\ z\rfloor$ **by** *simp*
lemma $\lfloor\forall\, Z{::}\uparrow\mathbf{0}.\ (\lambda X{::}\uparrow\mathbf{0}.\ \Box((\lambda\, Y{::}\uparrow\mathbf{0}.\ X \approx\ Y)\ \ Z))\ Z\rfloor$ **by** *simp*

Example 9.1, p.116 (using normal-, Leibniz- and concept-equality)

lemma $\lfloor((\lambda X.\ \Box(X\ \downharpoonleft(p{::}\uparrow\mathbf{0})))\ \downarrow(\lambda x.\ \Diamond(\lambda z.\ z \approx x)\ \downharpoonleft p))\rfloor$ **by** *auto*
lemma $\lfloor((\lambda X.\ \Box(X\ \downharpoonleft(p{::}\uparrow\mathbf{0})))\ \downarrow(\lambda x.\ \Diamond(\lambda z.\ z \approx^{L} x)\ \downharpoonleft p))\rfloor$ **by** *auto*
lemma $\lfloor((\lambda X.\ \Box(X\ \ (p{::}\uparrow\mathbf{0})))\ \downarrow(\lambda x.\ \Diamond(\lambda z.\ z \approx^{C} x)\ \ p))\rfloor$ **by** *simp*

2.8 Stability Conditions and Rigid Designation

As said before, intensional terms are trivially rigid. The following predicate tests whether an intensional predicate is 'rigid' in the sense of denoting a world-independent function.

abbreviation *rigidPred*::('t⇒io)⇒io **where**
rigidPred $\tau \equiv (\lambda\beta.\ \Box((\lambda z.\ \beta \approx z)\ \downarrow\tau))\ \downarrow\tau$

Following definitions are called 'stability conditions' by Fitting ([12], p. 124).

abbreviation *stabilityA*::('t⇒io)⇒io **where** *stabilityA* $\tau \equiv \forall\,\alpha.\ (\tau\ \alpha) \to \Box(\tau\ \alpha)$
abbreviation *stabilityB*::('t⇒io)⇒io **where** *stabilityB* $\tau \equiv \forall\,\alpha.\ \Diamond(\tau\ \alpha) \to (\tau\ \alpha)$

We prove them equivalent in *S5* logic (using *Sahlqvist correspondence*).

lemma *equivalence aRel* $\Longrightarrow \lfloor$*stabilityA* $(\tau{::}\uparrow\langle\mathbf{0}\rangle)\rfloor \longrightarrow \lfloor$*stabilityB* $\tau\rfloor$ **by** *blast*
lemma *equivalence aRel* $\Longrightarrow \lfloor$*stabilityB* $(\tau{::}\uparrow\langle\mathbf{0}\rangle)\rfloor \longrightarrow \lfloor$*stabilityA* $\tau\rfloor$ **by** *blast*

A term is rigid if and only if it satisfies the stability conditions.

theorem $\lfloor$*rigidPred* $(\tau{::}\uparrow\langle\mathbf{0}\rangle)\rfloor \longleftrightarrow \lfloor($*stabilityA* $\tau \wedge$ *stabilityB* $\tau)\rfloor$ **by** *meson*
theorem $\lfloor$*rigidPred* $(\tau{::}\uparrow\langle\uparrow\mathbf{0}\rangle)\rfloor \longleftrightarrow \lfloor($*stabilityA* $\tau \wedge$ *stabilityB* $\tau)\rfloor$ **by** *meson*

2.9 *De Re* and *De Dicto*

De re is equivalent to *de dicto* for non-relativized (i.e. rigid) terms:

lemma $\lfloor\forall\,\alpha.\ ((\lambda\beta.\ \Box(\alpha\ \beta))\ (\tau{::}\langle\mathbf{0}\rangle))\ \ \leftrightarrow \Box((\lambda\beta.\ (\alpha\ \beta))\ \tau)\rfloor$ **by** *simp*
lemma $\lfloor\forall\,\alpha.\ ((\lambda\beta.\ \Box(\alpha\ \beta))\ (\tau{::}\uparrow\langle\mathbf{0}\rangle)) \leftrightarrow \Box((\lambda\beta.\ (\alpha\ \beta))\ \tau)\rfloor$ **by** *simp*

De re is not equivalent to *de dicto* for relativized terms:

lemma $\lfloor\forall\,\alpha.\ ((\lambda\beta.\ \Box(\alpha\ \beta))\ \downarrow(\tau{::}\uparrow\langle\mathbf{0}\rangle)) \leftrightarrow \Box((\lambda\beta.\ (\alpha\ \beta))\ \downarrow\tau)\rfloor$
nitpick[*card 't=1*, *card i=2*] **oops** — countersatisfiable

3 Gödel's Ontological Argument

"Gödel's particular version of the argument is a direct descendent of that of Leibniz, which in turn derives from one of Descartes. These arguments all have a two-part structure: prove God's existence is necessary, if possible; and prove God's existence is possible." [12], p. 138.

3.1 Part I - God's Existence is Possible

For this first part, while Leibniz provides some kind of proof for the compatibility of all perfections, Gödel goes on to prove an analogous result: *(T1) 'Every positive property is possibly instantiated'*, which together with *(T2) 'God is a positive property'* directly implies the conclusion. In order to prove *T1*, Gödel assumes *(A2) 'Any property entailed by a positive property is itself positive'*. The definition of property entailment introduced by Gödel can be criticized on the grounds that it lacks some notion of relevance and is therefore exposed to the paradoxes of material implication. In particular, when we assert that property A does not entail property B, we implicitly assume that A is possibly instantiated. Conversely, an impossible property (like being a round square) entails any property (like being a triangle). It is precisely by virtue of these paradoxes that Gödel manages to prove *T1*.

TODO: comment about the original Leibnizian notion of *concept containment*? cite Zalta and who else?

abbreviation *Entailment*::$\uparrow\langle\uparrow\langle\mathbf{0}\rangle,\uparrow\langle\mathbf{0}\rangle\rangle$ (**infix** $\Rrightarrow$ *60*) **where**
$X \Rrightarrow Y \equiv \Box(\forall^E z.\ X\ z \rightarrow Y\ z)$

lemma $\lfloor(\lambda x\ w.\ x \neq x) \Rrightarrow \chi\rfloor$ **by** *simp*
lemma $\lfloor\neg(\varphi \Rrightarrow \chi) \rightarrow \Diamond\exists^E\ \varphi\rfloor$ **by** *auto*

consts *Positiveness*::$\uparrow\langle\uparrow\langle\mathbf{0}\rangle\rangle$ ($\mathcal{P}$) — positiveness applies to intensional predicates
abbreviation *Existence*::$\uparrow\langle\mathbf{0}\rangle$ ($E!$) — object-language existence predicate
where $E!\ x \equiv \lambda w.\ (\exists^E y.\ y{\approx}x)\ w$
abbreviation *appliesToPositiveProps*::$\uparrow\langle\uparrow\langle\uparrow\langle\mathbf{0}\rangle\rangle\rangle$ (*pos*) **where**
$pos\ Z \equiv\ \forall X.\ Z\ X \rightarrow \mathcal{P}\ X$
abbreviation *intersectionOf*::$\uparrow\langle\uparrow\langle\mathbf{0}\rangle,\uparrow\langle\uparrow\langle\mathbf{0}\rangle\rangle\rangle$ (*intersec*) **where**
$intersec\ X\ Z \equiv\ \Box(\forall x.(X\ x \leftrightarrow (\forall Y.\ (Z\ Y) \rightarrow (Y\ x))))$ — possibilist quantifier

axiomatization where
A1a:$\lfloor\forall X.\ \mathcal{P}\ (\rightharpoondown X) \rightarrow \neg(\mathcal{P}\ X)\ \rfloor$ **and** — axiom 11.3A
A1b:$\lfloor\forall X.\ \neg(\mathcal{P}\ X) \rightarrow \mathcal{P}\ (\rightharpoondown X)\rfloor$ **and** — axiom 11.3B
A2: $\lfloor\forall X\ Y.\ (\mathcal{P}\ X \wedge (X \Rrightarrow Y)) \rightarrow \mathcal{P}\ Y\rfloor$ **and** — axiom 11.5
A3: $\lfloor\forall Z\ X.\ (pos\ Z \wedge intersec\ X\ Z) \rightarrow \mathcal{P}\ X\rfloor$ — axiom 11.10

lemma *True* **nitpick**[*satisfy*] **oops** — model found: axioms are consistent

lemma $\lfloor D \rfloor$ **using** *A1a A1b A2* **by** *blast* — D axiom is implicitely assumed

Positive properties are possibly instantiated.

theorem *T1*: $\lfloor\forall X.\ \mathcal{P}\ X \rightarrow \Diamond\exists^E\ X\rfloor$ **using** *A1a A2* **by** *blast*

Being Godlike is defined as having all (and only) positive properties.

abbreviation *God*::$\uparrow\langle\mathbf{0}\rangle$ (G) **where** $G \equiv (\lambda x.\ \forall Y.\ \mathcal{P}\ Y \rightarrow Y\ x)$
abbreviation *God-star*::$\uparrow\langle\mathbf{0}\rangle$ ($G*$) **where** $G* \equiv (\lambda x.\ \forall Y.\ \mathcal{P}\ Y \leftrightarrow Y\ x)$

Both are equivalent. We can use either one for improved performance.

lemma *GodDefsAreEquivalent*: $\lfloor\forall x.\ G\ x \leftrightarrow G*\ x\rfloor$ **using** *A1b* **by** *force*

Being Godlike is itself a positive property. Note that this theorem can be axiomatized directly, as noted by Dana Scott (see [12], p. 152).

theorem *T2*: $\lfloor\mathcal{P}\ G\rfloor$
proof –
{
  fix w
  have *1*: *pos* $\mathcal{P}$ w **by** *simp*
  have *2*: *intersec* G $\mathcal{P}$ w **by** *simp*
  have $\lfloor\forall Z\, X.\ (pos\ Z \wedge intersec\ X\ Z) \rightarrow \mathcal{P}\ X\rfloor$ **by** (*rule A3*)
  hence $(\forall Z\, X.\ (pos\ Z \wedge intersec\ X\ Z) \rightarrow \mathcal{P}\ X)\ w$ **by** (*rule allE*)
  hence $(\forall X.\ ((pos\ \mathcal{P}) \wedge (intersec\ X\ \mathcal{P})) \rightarrow \mathcal{P}\ X)\ w$ **by** (*rule allE*)
  hence $(((pos\ \mathcal{P}) \wedge (intersec\ G\ \mathcal{P})) \rightarrow \mathcal{P}\ G)\ w$ **by** (*rule allE*)
  hence *3*: $((pos\ \mathcal{P} \wedge intersec\ G\ \mathcal{P})\ w) \longrightarrow \mathcal{P}\ G\ w$ **by** *simp*
  hence *4*: $((pos\ \mathcal{P}) \wedge (intersec\ G\ \mathcal{P}))\ w$ **using** *1 2* **by** *simp*
  from *3 4* **have** $\mathcal{P}\ G\ w$ **by** (*rule mp*)
}
thus *?thesis* **by** (*rule allI*)
qed

Conclusion for the first part: Possibly God exists.

theorem *T3*: $\lfloor\Diamond\exists^E\ G\rfloor$ **using** *T1 T2* **by** *simp*

3.2 Part II - God's Existence is Necessary, if Possible

In this part we show that God's necessary existence follows from its possible existence by adding some additional (philosophically controversial) assumptions including an *essentialist* premise and the *S5* axioms. Further derived results like monotheism and absence of free will are also discussed.

axiomatization where *A4a*: $\lfloor\forall X.\ \mathcal{P}\ X \rightarrow \Box(\mathcal{P}\ X)\rfloor$

Following lemma was originally assumed by Gödel as an axiom:

lemma *A4b*: $\lfloor\forall X.\ \neg(\mathcal{P}\ X) \rightarrow \Box\neg(\mathcal{P}\ X)\rfloor$ **using** *A1a A1b A4a* **by** *blast*
lemma *True* **nitpick**[*satisfy*] **oops** — model found: all axioms A1-4 consistent

Axiom *A4a* and its consequence *A4b* together imply that $\mathcal{P}$ satisfies Fitting's 'stability conditions' ([12], p. 124). This means $\mathcal{P}$ designates rigidly. Note that this makes for an *essentialist* assumption which may be considered controversial by some philosophers: every property considered positive in our world (e.g. honesty) is necessarily so.

lemma $\lfloor \mathit{rigidPred}\ \mathcal{P} \rfloor$ **using** *A4a A4b* **by** *blast*

Remark: Essence is defined here (and in Fitting's variant) in the version of Scott; Gödel's original version leads to the inconsistency reported in [8, 9]

abbreviation $\mathit{essenceOf}{::}\uparrow\langle\uparrow\langle\mathbf{0}\rangle,\mathbf{0}\rangle\ (\mathcal{E})$ **where**
$\mathcal{E}\ \ Y\, x \equiv (Y\, x) \wedge (\forall\, Z.\ Z\, x \rightarrow\ Y \Rrightarrow\ Z)$
abbreviation $\mathit{beingIdenticalTo}{::}\mathbf{0}{\Rightarrow}\uparrow\langle\mathbf{0}\rangle\ (\mathit{id})$ **where**
$\mathit{id}\ x\ \equiv (\lambda y.\ y{\approx}x)$ — *id* is here a rigid predicate (following Kripke [17])

Being God-like is an essential property:

theorem *GodIsEssential*: $\lfloor \forall\, x.\ G\ x \rightarrow (\mathcal{E}\ \ G\ x) \rfloor$ **using** *A1b A4a* **by** *metis*

Something can only have *one* essence:

theorem $\lfloor \forall\, X\ \ Y\, z.\ (\mathcal{E}\ \ X\ z \wedge \mathcal{E}\ \ Y\ z) \rightarrow (X \Rrightarrow\ Y) \rfloor$ **by** *meson*

An essential property offers a complete characterization of an individual:

theorem *EssencesCharacterizeCompletely*: $\lfloor \forall\, X\ y.\ \mathcal{E}\ \ X\ y \rightarrow (X \Rrightarrow (\mathit{id}\ y)) \rfloor$
proof (*rule ccontr*) — proof by contradiction not shown here

abbreviation $\mathit{necessaryExistencePredicate}{::}\uparrow\langle\mathbf{0}\rangle\ (\mathit{NE})$
where $\mathit{NE}\ x\ \equiv (\lambda w.\ (\forall\, Y.\ \ \mathcal{E}\ \ Y\, x \rightarrow \Box\exists^{\,E}\ \ Y)\ w)$

axiomatization where $\mathit{A5}$: $\lfloor \mathcal{P}\ \mathit{NE} \rfloor$
lemma *True* **nitpick**[*satisfy*] **oops** — model found: so far all axioms consistent

(Possibilist) existence of God implies its necessary (actualist) existence:

theorem *GodExistenceImpliesNecExistence*: $\lfloor \exists\ \ G \rightarrow\ \ \Box\exists^{\,E}\ G \rfloor$
proof − — direct proof not shown here

We postulate semantic frame conditions for some modal logics. Taken together, reflexivity, transitivity and symmetry make for an equivalence relation and therefore an *S5* logic (via *Sahlqvist correspondence*). They are individually postulated in order to get more detailed information about their relevance in the proofs presented below.

axiomatization where
refl: *reflexive aRel* **and**
tran: *transitive aRel* **and**
symm: *symmetric aRel*

lemma *True* **nitpick**[*satisfy*] **oops** — model found: axioms still consistent

We prove some useful inference rules:

lemma *modal-distr*: $\lfloor\Box(\varphi \to \psi)\rfloor \Longrightarrow \lfloor(\Diamond\varphi \to \Diamond\psi)\rfloor$ **by** *blast*
lemma *modal-trans*: $(\lfloor\varphi \to \psi\rfloor \wedge \lfloor\psi \to \chi\rfloor) \Longrightarrow \lfloor\varphi \to \chi\rfloor$ **by** *simp*

Possible existence of God implies its necessary (actualist) existence:

theorem *T4*: $\lfloor\Diamond\exists\ G\rfloor \longrightarrow \lfloor\Box\exists^E\ G\rfloor$ — only symmetry and transitivity needed
proof –
  have $\lfloor\exists\ G \to \Box\exists^E\ G\rfloor$ **using** *GodExistenceImpliesNecExistence*
    by *simp* — follows from Axioms 11.11, 11.25 and 11.3B
  hence $\lfloor\Box(\exists\ G \to \Box\exists^E\ G)\rfloor$ **using** *NEC* **by** *simp*
  hence *1*: $\lfloor\Diamond\exists\ G \to \Diamond\Box\exists^E\ G\rfloor$ **by** (*rule modal-distr*)
  have *2*: $\lfloor\Diamond\Box\exists^E\ G \to \Box\exists^E\ G\rfloor$ **using** *symm tran* **by** *metis* — frame conditions
  from *1 2* **have** $\lfloor\Diamond\exists\ G \to \Diamond\Box\exists^E\ G\rfloor \wedge \lfloor\Diamond\Box\exists^E\ G \to \Box\exists^E\ G\rfloor$ **by** *simp*
  hence $\lfloor\Diamond\exists\ G \to \Box\exists^E\ G\rfloor$ **by** (*rule modal-trans*)
  thus *?thesis* **by** (*rule localImpGlobalCons*)
qed

Conclusion: Necessary (actualist) existence of God:

lemma *GodNecExists*: $\lfloor\Box\exists^E\ G\rfloor$ **using** *T3 T4* **by** *metis*

By introducing reflexivity to our semantic frame conditions (axiom M/T):

lemma *GodExistenceIsValid*: $\lfloor\exists^E\ G\rfloor$ **using** *GodNecExists refl* **by** *auto*

Monotheism for non-normal models (with Leibniz equality) follows directly from God having all and only positive properties:

theorem *Monotheism-LeibnizEq*: $\lfloor\forall x.\ G\ x \to (\forall y.\ G\ y \to (x\ \approx^L\ y))\rfloor$
  using *GodDefsAreEquivalent* **by** *simp*

Monotheism for normal models is trickier, since we need to consider previous results ([12], p. 162):

theorem *Monotheism-normalModel*: $\lfloor\exists x.\forall y.\ G\ y \leftrightarrow x \approx y\rfloor$
proof – — direct proof not shown here

One of the objections to Gödel's argument is that his axioms imply that positive properties are necessarily instantiated. We can prove this true:

lemma *PosPropertiesNecExist*: $\lfloor\forall Y.\ \mathcal{P}\ Y \to \Box\exists^E\ Y\rfloor$ **using** *GodNecExists A4a*
  by *meson* — follows from corollary 11.28 and axiom A4a

Fitting [12] also discusses the objection raised by Sobel [20], who argues that Gödel's axiom system is too strong: it implies that whatever is the case is so necessarily, i.e. the modal system collapses ($\varphi \longrightarrow \Box\varphi$). The *modal collapse* has been philosophically interpreted as implying the absence of free will. In the context of our S5 axioms, the *modal collapse* becomes valid ([12], pp. 163-4):

lemma *useful*: $(\forall x.\ \varphi\ x \longrightarrow \psi) \Longrightarrow ((\exists x.\ \varphi\ x) \longrightarrow \psi)$ **by** *simp* — useful inf. rule

theorem *ModalCollapse*: $\lfloor \forall \Phi.(\Phi \rightarrow (\Box\ \Phi)) \rfloor$
proof –
  {
  fix w
   {
    fix Q
    have $(\forall x.\ G\ x \rightarrow (\mathcal{E}\ G\ x))\ w$ **using** *GodIsEssential* **by** (*rule allE*)
    hence $\forall x.\ G\ x\ w \longrightarrow \mathcal{E}\ G\ x\ w$ **by** *simp*
    hence $\forall x.\ G\ x\ w \longrightarrow (\forall Z.\ Z\ x \rightarrow \Box(\forall^E z.\ G\ z \rightarrow Z\ z))\ w$ **by** *force*
    hence $\forall x.\ G\ x\ w \longrightarrow ((\lambda y.\ Q)\ x \rightarrow \Box(\forall^E z.\ G\ z \rightarrow (\lambda y.\ Q)\ z))\ w$ **by** *force*
    hence $\forall x.\ G\ x\ w \longrightarrow (Q \rightarrow \Box(\forall^E z.\ G\ z \rightarrow Q))\ w$ **by** *simp*
    hence *1*: $(\exists x.\ G\ x\ w) \longrightarrow ((Q \rightarrow \Box(\forall^E z.\ G\ z \rightarrow Q))\ w)$ **by** (*rule useful*)
    have $\exists x.\ G\ x\ w$ **using** *GodExistenceIsValid* **by** *auto*
    from *1 this* **have** $(Q \rightarrow \Box(\forall^E z.\ G\ z \rightarrow Q))\ w$ **by** (*rule mp*)
    hence $(Q \rightarrow \Box((\exists^E z.\ G\ z) \rightarrow Q))\ w$ **using** *useful* **by** *blast*
    hence $(Q \rightarrow (\Box(\exists^E z.\ G\ z) \rightarrow \Box Q))\ w$ **by** *simp*
    hence $(Q \rightarrow \Box Q)\ w$ **using** *GodNecExists* **by** *simp*
   }
  hence $(\forall \Phi.\ \Phi \rightarrow \Box\ \Phi)\ w$ **by** (*rule allI*)
  }
  thus *?thesis* **by** (*rule allI*)
qed

4 Fitting's Variant

In this section we consider Fitting's solution to the objections raised in his discussion of Gödel's Argument pp. 164-9, especially the problem of *modal collapse*, which has been metaphysically interpreted as implying a rejection of free will. Since we are generally commited to the existence of free will (in a pre-theoretical sense), such a result is philosophically unappealing and rather seen as a problem in the argument's formalization.

Remark: The '$(\!|$-$|\!)$' parentheses are used to convert an extensional object into its 'rigid' intensional counterpart (e.g. $(\!|\varphi|\!) \equiv \lambda w.\ \varphi$).

abbreviation *Entailment*::$\uparrow\langle\langle\mathbf{0}\rangle,\langle\mathbf{0}\rangle\rangle$ (**infix**$\Rrightarrow$ *60*)
  where $X \Rrightarrow Y \equiv \Box(\forall^E z.\ (\!|X\ z|\!) \rightarrow (\!|Y\ z|\!))$
consts *Positiveness*::$\uparrow\langle\langle\mathbf{0}\rangle\rangle$ ($\mathcal{P}$)
abbreviation *Existence*::$\uparrow\langle\mathbf{0}\rangle$ ($E!$) **where** $E!\ x \equiv \lambda w.\ (\exists^E y.\ y \approx x)\ w$
abbreviation *God*::$\uparrow\langle\mathbf{0}\rangle$ (G) **where** $G \equiv (\lambda x.\ \forall Y.\ \mathcal{P}\ Y \rightarrow (\!|Y\ x|\!))$

4.1 Part I - God's Existence is Possible

axiomatization where
  A1a: $\lfloor \forall X.\ \mathcal{P}\ (\rightharpoondown X) \rightarrow \neg(\mathcal{P}\ X)\ \rfloor$ **and** — axiom 11.3A
  A1b: $\lfloor \forall X.\ \neg(\mathcal{P}\ X) \rightarrow \mathcal{P}\ (\rightharpoondown X) \rfloor$ **and** — axiom 11.3B
  A2: $\lfloor \forall X\ Y.\ (\mathcal{P}\ X \wedge (X \Rrightarrow Y)) \rightarrow \mathcal{P}\ Y \rfloor$ **and** — axiom 11.5
  T2: $\lfloor \mathcal{P}\ \downarrow G \rfloor$ — proposition 11.16 (modified)
lemma *True* **nitpick**[*satisfy*] **oops** — model found: axioms are consistent

T1 Positive properties are possibly instantiated

theorem *T1*: $\lfloor \forall X::\langle \mathbf{0} \rangle.\ \mathcal{P}\ X \rightarrow \Diamond(\exists^E z.\ (\!| X\ z |\!)) \rfloor$ **using** *A1a A2* **by** *blast*

T3 (God exists possibly) can be formalized in two different ways, using a *de re* or a *de dicto* reading.

theorem *T3-deRe*: $\lfloor (\lambda X.\ \Diamond\exists^E\ X)\ \downarrow G \rfloor$ **using** *T1 T2* **by** *simp*
theorem *T3-deDicto*: $\lfloor \Diamond\exists^E \downarrow G \rfloor$ **nitpick oops** — countersatisfiable: not used

4.2 Part II - God's Existence is Necessary if Possible

axiomatization where
  A4a: $\lfloor \forall X.\ \mathcal{P}\ X \rightarrow \Box(\mathcal{P}\ X) \rfloor$ — axiom 11.11
lemma *A4b*: $\lfloor \forall X.\ \neg(\mathcal{P}\ X) \rightarrow \Box\neg(\mathcal{P}\ X) \rfloor$ **using** *A1a A1b A4a* **by** *blast*

lemma *True* **nitpick**[*satisfy*] **oops** — model found: so far all axioms consistent
lemma $\lfloor$*rigidPred* $\mathcal{P}\rfloor$ **using** *A4a A4b* **by** *blast* — $\mathcal{P}$ designates rigidly

abbreviation *essenceOf*::$\uparrow\langle\langle \mathbf{0}\rangle,\mathbf{0}\rangle$ ($\mathcal{E}$) **where**
  $\mathcal{E}\ Y\ x \equiv (\!| Y\ x |\!) \wedge (\forall Z::\langle \mathbf{0}\rangle.\ (\!| Z\ x |\!) \rightarrow Y \Rrightarrow Z)$
theorem *GodIsEssential*: $\lfloor \forall x.\ G\ x \rightarrow ((\mathcal{E} \downarrow_1 G)\ x) \rfloor$ **using** *A1b* **by** *metis*

abbreviation *necessaryExistencePredicate* :: $\uparrow\langle \mathbf{0}\rangle$ (*NE*) **where**
  $NE\ x \equiv \lambda w.\ (\forall Y.\ \mathcal{E}\ Y\ x \rightarrow \Box(\exists^E z.\ (\!| Y\ z |\!)))\ w$

axiomatization where *A5*: $\lfloor \mathcal{P} \downarrow NE \rfloor$
lemma *True* **nitpick**[*satisfy*] **oops** — model found: so far all axioms consistent

Theorem 11.26 (Informal Proposition 7) - (possibilist) existence of God implies necessary (actualist) existence. This theorem can be formalized in two ways. Both of them are proven valid:

theorem *GodExImpNecEx-v1*: $\lfloor \exists \downarrow G \rightarrow \Box\exists^E \downarrow G \rfloor$ **proof** – — not shown here

theorem *GodExImpNecEx-v2*: $\lfloor \exists \downarrow G \rightarrow ((\lambda X.\ \Box\exists^E\ X) \downarrow G) \rfloor$
  using *A4a GodExImpNecEx-v1* **by** *metis* — can be proven by automated tools

In contrast to Gödel's argument (as presented by Fitting), the following theorems can be proven in logic *K* (the *S5* axioms are no longer needed):

theorem *possExImpNecEx-v1*: $\lfloor \Diamond\exists \downarrow G \rightarrow \Box\exists^E \downarrow G \rfloor$
  using *GodExImpNecEx-v1 T3-deRe* **by** *metis*
theorem *possExImpNecEx-v2*: $\lfloor (\lambda X.\Diamond\exists^E\ X) \downarrow G \rightarrow (\lambda X.\ \Box\exists^E\ X) \downarrow G \rfloor$
  using *GodExImpNecEx-v2* **by** *blast*

lemma *T4-v1*:$\lfloor \Diamond\exists \downarrow G \rfloor \longrightarrow \lfloor \Box\exists^E \downarrow G \rfloor$ **using** *possExImpNecEx-v1* **by** *simp*
lemma *T4-v2*:$\lfloor (\lambda X.\ \Diamond\exists^E\ X) \downarrow G \rfloor \longrightarrow \lfloor (\lambda X.\ \Box\exists^E\ X) \downarrow G \rfloor$ **using** *possExImpNecEx-v2* **by** *simp*

4.3 Conclusion (*De Re* and *De Dicto* Reading)

Version I - Necessary Existence of God (*de dicto*):

lemma *GodNecExists-v1*: $\lfloor\Box\exists^E \downarrow G\rfloor$
  using *GodExImpNecEx-v1 T3-deRe* **by** *fastforce* — corollary 11.28
lemma $\lfloor\Box(\lambda X.\ \exists^E\ X)\ \downarrow G\rfloor$
  using *GodNecExists-v1* **by** *simp* — *de dicto* shown here explicitly

Version II - Necessary Existence of God (*de re*)

lemma *GodNecExists-v2*: $\lfloor(\lambda X.\ \Box\exists^E\ X)\ \downarrow G\rfloor$
  using *T3-deRe T4-v2* **by** *blast*

4.4 Modal Collapse

Modal collapse is countersatisfiable even in *S5*. Note that countermodels with a cardinality of one for the domain of individuals are found by *Nitpick* (the countermodel shown in the book has cardinality of two).

axiomatization where *S5*: *equivalence aRel* — *S5* axioms assumed
lemma $\lfloor\forall\,\Phi.(\Phi \rightarrow (\Box\ \Phi))\rfloor$ **nitpick**[*card 't=1*, *card i=2*] **oops** — countermodel

5 Anderson's Alternative

In this final section, we verify Anderson's emendation of Gödel's argument, as it is presented by Fitting in [12], pp. 169-171).

abbreviation *Entailment*::$\uparrow\langle\uparrow\langle\mathbf{0}\rangle,\uparrow\langle\mathbf{0}\rangle\rangle$ (**infix** $\Rrightarrow$ *60*) **where**
  $X \Rrightarrow Y \equiv \Box(\forall^E z.\ X\ z \rightarrow Y\ z)$
consts *Positiveness*::$\uparrow\langle\uparrow\langle\mathbf{0}\rangle\rangle$ ($\mathcal{P}$)
abbreviation *Existence*::$\uparrow\langle\mathbf{0}\rangle$ ($E!$) **where** $E!\ x \equiv \lambda w.\ (\exists^E y.\ y{\approx}x)\ w$
abbreviation *God*::$\uparrow\langle\mathbf{0}\rangle$ (G^A) **where** $G^A \equiv \lambda x.\ \forall\,Y.\ (\mathcal{P}\ Y) \leftrightarrow \Box(Y\ x)$

5.1 Part I - God's Existence is Possible

axiomatization where
  A1a: $\lfloor\forall\,X.\ \mathcal{P}\ (\rightharpoondown X) \rightarrow \neg(\mathcal{P}\ X)\ \rfloor$ **and** — Axiom 11.3A
  A2: $\lfloor\forall\,X\ Y.\ (\mathcal{P}\ X \wedge (X \Rrightarrow Y)) \rightarrow \mathcal{P}\ Y\rfloor$ **and** — Axiom 11.5
  T2: $\lfloor\mathcal{P}\ G^A\rfloor$ — Proposition 11.16

lemma *True* **nitpick**[*satisfy*] **oops** — model found: axioms are consistent

T1 Positive properties are possibly instantiated

theorem *T1*: $\lfloor\forall\,X.\ \mathcal{P}\ X \rightarrow \Diamond\exists^E\ X\rfloor$ **using** *A1a A2* **by** *blast*

T3 God exists possibly

theorem *T3*: $\lfloor\Diamond\exists^E\ G^A\rfloor$ **using** *T1 T2* **by** *simp*

5.2 Part II - God's Existence is Necessary if Possible

$\mathcal{P}$ now satisfies only one of the stability conditions. But since this variant uses an *S5* logic, the other stability condition is implied (see [12], p. 124). Therefore $\mathcal{P}$ becomes rigid .

axiomatization where $A4a$: $\lfloor \forall X.\ \mathcal{P}\ X \to \Box(\mathcal{P}\ X)\rfloor$

We again postulate our *S5* axioms:

axiomatization where
refl: *reflexive aRel* **and**
tran: *transitive aRel* **and**
symm: *symmetric aRel*
lemma *True* **nitpick**[*satisfy*] **oops** — model found: so far all axioms consistent

lemma $A4b$: $\lfloor \forall X.\ \neg(\mathcal{P}\ X) \to \Box\neg(\mathcal{P}\ X)\rfloor$
using *A4a symm* **by** *auto* — symmetry is needed (corresponding to B axiom)
lemma $\lfloor rigidPred\ \mathcal{P}\rfloor$ **using** *A4a A4b* **by** *blast* — $\mathcal{P}$ is rigid in B

abbreviation $essenceOf{::}\uparrow\langle\uparrow\langle\mathbf{0}\rangle,\mathbf{0}\rangle$ $(\mathcal{E}^A)$ **where** — Essence, Anderson Version
$\mathcal{E}^A\ \ Y\, x \equiv (\forall Z.\ \Box(Z\ x) \leftrightarrow\ Y \Rrightarrow Z)$
abbreviation $necessaryExistencePred{::}\uparrow\langle\mathbf{0}\rangle$ (NE^A)
where $NE^A\ x\ \equiv (\lambda w.\ (\forall\ Y.\ \ \mathcal{E}^A\ \ Y\, x \to \Box\exists^{\,E}\ Y)\ w)$

If g is God-like, the property of being God-like is its essence. As shown before, this theorem's proof could be completely automatized for Gödel's and Fitting's variants. For Anderson's version however, we had to provide Isabelle with some help based on the corresponding natural-language proof given by Anderson (see [2] Theorem 2*, p. 296)

theorem *GodIsEssential*: $\lfloor \forall x.\ G^A\ x \to (\mathcal{E}^A\ \ G^A\ x)\rfloor$ **proof** − — not shown here
axiomatization where $A5$: $\lfloor \mathcal{P}\ NE^A\rfloor$
lemma *True* **nitpick**[*satisfy*] **oops** — model found: so far all axioms consistent

Possibilist existence of God implies necessary actualist existence:

theorem *GodExistenceImpliesNecExistence*: $\lfloor\exists\ \ G^A\ \to\ \ \Box\exists^{\,E}\ G^A\rfloor$ **proof** − — not shown here

Some useful rules:

lemma *modal-distr*: $\lfloor\Box(\varphi \to \psi)\rfloor \Longrightarrow \lfloor(\Diamond\varphi \to \Diamond\psi)\rfloor$ **by** *blast*
lemma *modal-trans*: $(\lfloor\varphi \to \psi\rfloor \wedge \lfloor\psi \to \chi\rfloor) \Longrightarrow \lfloor\varphi \to \chi\rfloor$ **by** *simp*

Anderson's version of Theorem 11.27

theorem $T4$: $\lfloor\Diamond\exists\ \ G^A\rfloor \longrightarrow \lfloor\Box\exists^{\,E}\ G^A\rfloor$
proof −
have $\lfloor\exists\ \ G^A \to \Box\exists^{\,E}\ G^A\rfloor$ **using** *GodExistenceImpliesNecExistence*
by *simp* — follows from Axioms 11.11, 11.25 and 11.3B
hence $\lfloor\Box(\exists\ \ G^A \to \Box\exists^{\,E}\ G^A)\rfloor$ **using** *NEC* **by** *simp*
hence 1: $\lfloor\Diamond\exists\ \ G^A \to \Diamond\Box\exists^{\,E}\ G^A\rfloor$ **by** (*rule modal-distr*)

have *2*: $\lfloor \Diamond\Box\exists^E\ G^A \rightarrow \Box\exists^E\ G^A \rfloor$ **using** *symm tran* **by** *metis*
from *1 2* **have** $\lfloor \Diamond\exists\ \ G^A \rightarrow \Diamond\Box\exists^E\ G^A \rfloor \wedge \lfloor \Diamond\Box\exists^E\ G^A \rightarrow \Box\exists^E\ G^A \rfloor$ **by** *simp*
hence $\lfloor \Diamond\exists\ \ G^A \rightarrow \Box\exists^E\ G^A \rfloor$ **by** (*rule modal-trans*)
thus *?thesis* **by** (*rule localImpGlobalCons*)
qed

Conclusion - Necessary (actualist) existence of God:

lemma *GodNecExists*: $\lfloor \Box\exists^E\ G^A \rfloor$ **using** *T3 T4* **by** *metis*
lemma *MC*: $\lfloor \forall\,\Phi.(\Phi \rightarrow (\Box\ \Phi)) \rfloor$ **nitpick oops** — modal collapse countersatisfiable

6 Conclusion

We presented a shallow semantical embedding in Isabelle/HOL for an intensional higher-order modal logic (a successor of Montague/Gallin intensional logics) as introduced by M. Fitting in his textbook *Types, Tableaus and Gödel's God* [12]. We employed this logic to formalize and verify all results relevant to the subsequent discussion of three different variants of the ontological argument: the first one by Gödel himself (respectively, Scott), the second one by Fitting and the last one by Anderson.

By employing an interactive theorem-prover like Isabelle, we were not only able to verify Fitting's results, but also to guarantee consistency. We could prove even stronger versions of many of the theorems and find better countermodels (i.e. with smaller cardinality) than the ones presented in his book. Another interesting aspect was the possibility to explore the implications of alternative formalizations for definitions and theorems which shed light on interesting philosophical issues concerning entailment, essentialism and free will, which are currently the subject of some follow-up analysis.

The latest developments in *automated theorem proving* allow us to engage in much more experimentation during the formalization and assessment of arguments than ever before. The potential reduction (of several orders of magnitude) in the time needed for proving or disproving theorems (compared to pen-and-paper proofs), results in almost real-time feedback about the suitability of our speculations. The practical benefits of computer-supported argumentation go beyond mere quantitative (easier, faster and more reliable proofs). The advantages are also qualitative, since it fosters a different approach to argumentation: We can now work iteratively (by 'trial-and-error') on an argument by making gradual adjustments to its definitions, axioms and theorems. This allows us to continuously expose and revise the assumptions we indirectly commit ourselves everytime we opt for some particular formalization.

References

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

[2] C. Anderson. Some emendations of Gödel's ontological proof. *Faith and Philosophy*, 7(3), 1990.

[3] C. Benzmüller. Universal reasoning, rational argumentation and human-machine interaction. *arXiv, http://arxiv.org/abs/1703.09620*, 2017.

[4] C. Benzmüller, M. Claus, and N. Sultana. Systematic verification of the modal logic cube in Isabelle/HOL. In C. Kaliszyk and A. Paskevich, editors, *PxTP 2015*, volume 186, pages 27–41, Berlin, Germany, 2015. EPTCS.

[5] C. Benzmüller and B. W. Paleo. Automating Gödel's ontological proof of God's existence with higher-order automated theorem provers. In T. Schaub, G. Friedrich, and B. O'Sullivan, editors, *ECAI 2014*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 93 – 98. IOS Press, 2014.

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

[7] C. Benzmüller, A. Steen, and M. Wisniewski. The computational metaphysics lecture course at Freie Universität Berlin. In S. Krajewski and P. Balcerowicz, editors, *Handbook of the 2nd World Congress on Logic and Religion, Warsaw, Poland*, page 2, 2017.

[8] C. Benzmüller and B. Woltzenlogel Paleo. The inconsistency in Gödels ontological argument: A success story for AI in metaphysics. In *IJCAI 2016*, 2016.

[9] C. Benzmüller and B. Woltzenlogel Paleo. An object-logic explanation for the inconsistency in Gödel's ontological theory (extended abstract). In M. Helmert and F. Wotawa, editors, *KI 2016: Advances in Artificial Intelligence, Proceedings*, LNCS, Berlin, Germany, 2016. Springer.

[10] F. Bjørdal. Understanding Gödel's ontological argument. In T. Childers, editor, *The Logica Yearbook 1998*. Filosofia, 1999.

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

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

[13] M. Fitting and R. Mendelsohn. *First-Order Modal Logic*, volume 277 of *Synthese Library.* Kluwer, 1998.

[14] D. Gallin. *Intensional and Higher-Order Modal Logic.* N.-Holland, 1975.

[15] K. Gödel. *Appx.A: Notes in Kurt Gödel's Hand*, pages 144–145. In [20], 2004.

[16] P. Hájek. A new small emendation of Gödel's ontological proof. *Studia Logica*, 71(2):149–164, 2002.

[17] S. Kripke. *Naming and Necessity.* Harvard University Press, 1980.

[18] D. Scott. *Appx.B: Notes in Dana Scott's Hand*, pages 145–146. In [20], 2004.

[19] J. Sobel. Gödel's ontological proof. In *On Being and Saying. Essays for Richard Cartwright*, pages 241–261. MIT Press, 1987.

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