PETR HÁJEK

A New Small Emendation of Gödel's Ontological Proof

Keywords: Ontological proof, Gödel, modal logic, comprehension, positive properties.

1. Introduction

Gödel's ontological proof of necessary existence of a godlike being was finally published in the third volume of Gödel's collected works [7]; but it became known in 1970 when Gödel showed the proof to Dana Scott and Scott presented it (in fact a variant of it) at a seminar at Princeton. Detailed history is found in Adams' introductory remarks to the ontological proof in [7]. The proof uses modal logic and its analysis is an exciting exercise in systems of formal modal logic. Needles to say, formal modal logic has found several uses e.g. in the theory of consistency proofs (provability logic) as well as in foundations of computer science and artificial intelligence (dynamic logic, logics of belief, modal reasoning under uncertainty).

From the literature analyzing and modifying Gödel's proof let me first cite Sobel [12] observing that (with enough comprehension) the system suffers by collapse of modalities (proving $\varphi \to \Box\varphi$ for each φ); Anderson's emendation [1] having no collapse, my two papers [8, 9] in which I offer another analysis of the collapse and show that (under an appropriate modal predicate logic) Anderson's axiom system is redundant; further [2] where the authors point out that this is not the case in a weaker modal logic.[1] The original proof by Gödel as well as the emendations can be understood in a formalism having variables for objects, variables for (unary) properties of objects, a binary application predicate (object x has the property X) and a unary predicate (a property X is positive). Gödel (also in Scott's and Anderson's modification) has 5 axioms; below the reader will encounter Anderson's variant which will be our starting point. What makes the proof interesting is the fact that it depends on several choices: (1) the underlying modal propositional logic (S5 or less?), (2) the asssumed comprehension

[1] I also mention the paper [6], presently in final stages of preparation.

Presented by **Heinrich Wansing**; *Submitted* April 3, 2001; *Revised* July 12, 2001; *Accepted* August 27, 2001

Studia Logica 71: 149–164, 2002.

(which formulas define properties?), (3) the underlying behaviour of quantifiers with respect to modalities. To illustrate the last point the question is if we think of Kripke models with a set W of possible worlds, a fixed domain M of objects (not depending on W) and properties interpreted as mappings of $W \times M$ into $\{0,1\}$ (plus an interpretation of positiveness) *or* if we assume also the domain to be dependent on possible worlds, each world bearing its own domain. The latter semantics is reducible to the former by introducing a special property E describing for each possible worlds the subset of objects of the general domain existing in the given world. The details of these semantics and Gödel's proof are discussed in depth in the forthcoming monograph [5] by Fitting where the author also explains Anderson's emendation and mentions mine.

The emendations have concerned mainly Gödel's first axiom (saying that a property is positive iff its negation is not positive). Gödel's second axiom can be read as saying that if a property X is positive and Y is its necessary consequence then Y is also positive. It has not been questioned by any of the above authors but I find it difficult to motivate: If $\mathtt{H}$ stands for godlike and D for "devillike" (a predicate not occuring in Gödel) then assuming $\mathtt{H}$ to be positive we would have to assume that the disjunction of $\mathtt{H}$ and D (being godlike or being devillike) is also positive, which appears to be counterintuitive. (This example also occurs in [4].) In the present paper we show how to weaken the axiom to get rid of this unpleasant feature, preserving the desired consequence.

Acknowledgement. Support of the project LN00A056 (ITI) of the Ministery of Education of the Czech Republic is acknowledged. Thanks are due to the anonymous referee for calling my attention to many misprints and similar errors in the first version of the paper.

2. Preliminaries

In this section we describe in detail our underlying logic and summarize known results related to Gödels ontological proof.

DEFINITION 1 (Language; formulas). We shall keep the logic first-order, but two-sorted; we shall have a sort of individuals and a sort of properties. (Cf. my [9]) Variables for individuals are lower-case x, y, z, $\ldots$; variables for properties are upper case X, Y, Z, $\ldots$. Our language will have one unary predicate $\mathtt{H}$ applied to individuals (godlike), one unary predicate $\mathtt{P}$ applied to properties (positive) and one binary predicate $\in$ of application ($x \in X$

will be abbreviated as $X(x)$). Furthermore we shall have two predicates $=_1$, $=_2$ (later understood as equality of individuals, equality of properties). New predicates, constants and functions may be introduced.

Formulas are built from atomic formulas of the form $x \in X$, $x =_1 y$, $X =_2 Y$, $\mathtt{H}(x)$, $\mathtt{P}(X)$ using connectives, quantifiers $\forall$, $\exists$ and modalities $\Box$, $\Diamond$ ($\exists$ and $\Diamond$ being defined from $\forall$ and $\Box$ in the obvious way). Free and bound occurrences of a variable in a formula are defined as usual; note that $\Box$ does not bind any variable.

DEFINITION 2 (The modal predicate logic S5). Logical axioms are:

- axioms of propositional logic
- usual axioms for the universal quantifier, i.e.

$$(\forall x)\varphi \to \varphi(x/t)$$

  where t is a term of the sort of individuals free for x in φ (substitution; presently the only terms are variables, later we shall have more terms),

$$(\forall X)\varphi \to \varphi(X/T)$$

  (similarly for T a term of the sort of properties free for X in T,

$$(\forall x)(\alpha \to \varphi) \to (\alpha \to (\forall x)\varphi)$$

  where x is not free in φ, similarly for X instead of x.
- axioms of the modal logic S5, i.e.

$$\begin{gathered}\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)\\ \Box\varphi \to \Box\Box\varphi,\\ \Box\varphi \to \varphi,\\ \Diamond\varphi \to \Box\Diamond\varphi.\end{gathered}$$

Deductions rules are modus ponens, generalization (from φ infer $(\forall x)\varphi$, $(\forall X)\varphi$) and necessitation (from φ infer $\Box\varphi$). A *theory* is a set of formulas. A *proof* in a theory Q is a sequence of formulas, in which each member is a logical axiom, an element of Q or it follows from some preceding formulas by one of deduction rules.

FACT. *The following formulas are provable in our logic:* $\Box(\forall x)\varphi \equiv (\forall x)\Box\varphi$, $\Box(\forall X)\varphi \equiv (\forall X)\Box\varphi$ *(Barcan formulas)* $(\exists x)\Box\varphi \to \Box(\exists x)\varphi$. *(Well known)*

DEFINITION 3. A *fixed universe Kripke model* for our language is a structure

$$\boldsymbol{K} = (W, M_1, M_2, r_{\mathtt{H}}, r_{\mathsf{P}}, r_{\in}, r_{=_1}, r_{=_2})$$

where $W \neq \emptyset$ is a set of possible worlds, $M_1 \neq \emptyset$ a set of individuals, M_2 a set of properties and the r's are interpretations of the respective predicates, i.e.

$$r_{\mathtt{H}}\colon M_1 \times W \to \{0,1\}, \quad r_{\mathsf{P}}\colon M_2 \times W \to \{0,1\},$$
$$r_{\in}\colon M_1 \times M_2 \times W \to \{0,1\},$$
$$r_{=_1}\colon M_1 \times M_1 \times W \to \{0,1\}, r_{=_2}\colon M_2 \times M_2 \times W \to \{0,1\}.$$

A $\boldsymbol{K}$-*evaluation of variables* is a mapping e assigning to each variable x an element $e(x) \in M_1$ and to each X an $e(X) \in M_2$.

Satisfaction is defined as usual:

$$\boldsymbol{K}, w \Vdash \mathtt{H}(x)\,[e] \quad \text{iff} \quad r_{\mathtt{H}}(e(x), w) = 1,$$

analogously for $\mathsf{P}(X)$;

$$\boldsymbol{K}, w \Vdash x \in X\,[e] \quad \text{iff} \quad r_{\in}(e(x), e(X), w) = 1,$$

analogously for $x = y$, $X = Y$.

The induction step for connectives is evident;

$$\boldsymbol{K}, w \Vdash (\forall x)\varphi\,[e] \quad \text{iff} \quad \boldsymbol{K}, w \Vdash \varphi[e'] \quad \text{for each } e' \text{ coinciding with } e \text{ for all variables except possibly } x;$$

analogously for $(\forall X)\varphi$.

$$\boldsymbol{K}, w \Vdash \Box\varphi\,[e] \quad \text{iff} \quad \text{for all } w',\ \boldsymbol{K}, w' \Vdash \varphi\,[e].$$

A formula φ is *true* in $\boldsymbol{K}$ if $\boldsymbol{K}, w \Vdash \varphi\,[e]$ for all w, e. $\boldsymbol{K}$ is a *model* of Q if all axioms $x \in Q$ are true in $\boldsymbol{K}$.

We have the following classical completeness theorem:[2]

THEOREM 1. *Let Q be a theory and φ a formula. $Q \vdash \varphi$ (Q proves φ) iff φ is true in all models of Q.*

Remark 1. Our logic will have further axioms (thus it will be a theory in the predicate logic S5). The axioms concern equality, extensionality and comprehension.

[2] See e.g. [10].

DEFINITION 4. The *equality axioms* for $=_1$, $=_2$ are

- reflexivity, symmetry and transitivity, i.e.
  $x =_1 x, \quad x =_1 y \to y =_1 x, \quad (x =_1 y \,\&\, y =_1 z) \to x =_1 z$
  $X =_2 X, \quad X =_2 Y \to Y =_2 X$, etc.
- congruence for $\mathtt{H}$, $\in$:
  $x = y \to \mathtt{H}(x) \equiv \mathtt{H}(y)$
  $(x = y \,\&\, X = Y) \to (X(x) \equiv Y(y))$
- weak congruence for P :
  $\Box(X = Y) \to (\mathsf{P}(X) \equiv \mathsf{P}(Y))$.

The last axiom expresses the fact that positiveness is understood as a *global* attribute of a property; it is admitted that in a particular possible world a positive property may equal to a non-positive (e.g. both may be empty in the given world).

DEFINITION 5. The *extensionality axiom* is the following:

$$X = Y \equiv (\forall x)(X(x) \equiv Y(x)).$$

It corresponds to a "set-theoretical" understanding of properties; assuming it makes the system simpler. One could eliminate it for the prize of unnecessary complications.

FACT. *Assume all the axioms postulated up to now and let* $\boldsymbol{K} = (W, M_1, M_2, r_{\mathtt{H}}, r_{\mathsf{P}}, e_{\in}, r_{=_1}, r_{=_2})$ *be a model. By an obvious factorization we may assume, without loss of generality, that necessary equality is interpreted as identity, i.e., if* $\boldsymbol{K} \Vdash \Box(x =_1 y)\,[e]$ *then* $e(x) = e(y)$*; and if* $\boldsymbol{K} \Vdash \Box(X =_2 Y)\,[e]$ *then* $e(X) = e(Y)$. *Thus a property* $p \in M_2$ *is uniquely given by the function* $f_p : M_1 \times W \to \{0,1\}$ *defined as* $f_p(m,w) = r_{\in}(m,p,w)$. *The interpretation* $r_{=_2}$ *of* $=_2$ *is superfluous since uniquely determined by extensionality. If* $r_{=_1}$ *happens to be the identity on* M_1 *(which is often possible), then* $\boldsymbol{K}$ *is uniquely determined by* $(W, M_1, Prop, r_{\mathtt{H}}, r_{\mathsf{P}})$*, where Prop is a set of mappings* $f\colon M_1 \times W \to \{0,1\}$ *(the set of all* f_p*'s),* $r_{\mathtt{H}}\colon M_1 \times W \to \{0,1\}$ *and* $r_{\mathsf{P}}\colon Prop \times W \to \{0,1\}$*. Clearly,* $r_{\in}$ *is recovered as follows: for* $u \in M_1$*,* $f \in Prop$*,* $r_{\in}(u,f,w) = f(u,w)$*.*

DEFINITION 6. The following is the schema of *full comprehension* ($\mathrm{C_{full}}$): for each formula $\varphi(x)$ not containing the variable Y,

$$(\exists Y)\Box(\forall x)(Y(x) \equiv \varphi(x))$$

is the comprehension axiom for φ (stating that the formula $\varphi(x)$ defines a property). Note that φ may have free variables other than x, but different from Y.

FACT. *If a theory contains the full comprehension schema then, in particular, properties can be combined using Boolean connectives, e.g. we can prove*

$$(\exists Y)\Box(\forall x)(Y(x) \equiv \neg X(x)),$$
$$(\exists Z)\Box(\forall x)(Z(x) \equiv U(x) \,\&\, V(x)), \quad etc.$$

DEFINITION 7. The logic $\mathcal{L}_{\mathrm{ont}}$ in which we analyze the ontological proofs has the axioms formulated till now, i.e. predicate modal S5 logic with equality axioms as above, extensionality for properties and full comprehension. We define complement of a property and intersection of two properties as follows:

$$Y = \neg X \equiv \Box(\forall x)(Y(x) \equiv \neg X(x))$$
$$Z = X \,\&\, Y \equiv \Box(\forall x)(Z(x) \equiv X(x) \,\&\, Y(x)).$$

(Evidently, this extends the logic conservatively.) Let us also define

$$X \subseteq Y \equiv (\forall x)(X(x) \to Y(x)).$$

(Hence $X = Y$ is equivalent to $(X \subseteq Y) \,\&\, (Y \subseteq X)$.)

Remark 2. It follows from comprehension that there is a property X such that $\Box(\forall x)(X(x) \equiv \mathtt{H}(x))$ or more pedantically, $\Box(\forall x)(x \in X \equiv \mathtt{H}(x))$. $\mathtt{H}$ is a predicate; but without any danger of misunderstanding we may denote the corresponding property also by H. (Similarly for other possibly introduced predicates on individuals.)

* * *

Let us now survey some results on Gödel's system.

The original axioms by Gödel (in Scott's version) read as follows ($\mathtt{G}$ for godlike):

$$\mathsf{P}(X) \equiv \neg\mathsf{P}(-X)$$
$$(\mathsf{P}(X) \,\&\, \Box(X \subseteq Y)) \to \mathsf{P}(Y)$$
$$\mathtt{G}(x) \equiv (\forall Y)(\mathsf{P}(Y) \to Y(x))$$
$$\mathsf{P}(\mathtt{G})$$
$$\mathsf{P}(X) \to \Box\mathsf{P}(X)$$
$$\mathtt{Ess}_Y(x) \equiv Y(x) \,\&\, (\forall Z)(Z(x) \to \Box(Y \subseteq Z))$$
$$\mathtt{NE}(x) \equiv (\forall Y)(\mathtt{Ess}_Y(x) \to \Box(\exists u)Y(u))$$
$$\mathsf{P}(\mathtt{NE}).$$

From these Gödel proves $\Diamond(\exists x)\mathtt{G}(x)$ and $(\exists x)\mathtt{G}(x) \to \Box(\exists x)\mathtt{G}(x)$ (the second being sometimes called Anselm's principle); the rest is obvious by modal logic — the system proves $\Box(\exists x)\mathtt{G}(x)$. We may call his system $\mathcal{GO}$ (Gödel-ontological).

As mentioned above, Sobel [12] proves that under full comprehension Gödel's system proves $\varphi \equiv \Box\varphi$ for each φ, hence it has collapse of modalities. To save it one has either to weaken comprehension or to modify the axioms. The former was done in my [8, 9], where I proposed a "cautious" comprehension, i.e. comprehension for $\mathtt{G}$, $\mathtt{NE}$ and the following schema:

$$(\forall x)(\mathtt{G}(x) \to (\Box\varphi(x) \vee \Box\neg\varphi(x)) \;\to\; (\exists Y)\Box(\forall x)(Y(x) \equiv \varphi(x)).$$

Admittedly, this is not too elegant; I showed by constructing a model, that modalities do not collapse.

Anderson [1] presents a modification of Gödel's system $\mathcal{GO}$ and shows that the modified system proves the necessary existence of a godlike being and is free of the collapse. His system will be our starting point (now $\mathtt{H}$ stands for godlike).

(A1) $\quad \mathtt{P}(X) \to \neg\mathtt{P}(\neg X)$

(A2) $\quad (\mathtt{P}(X) \,\&\, \Box(X \subseteq Y)) \to \mathtt{P}(Y)$ (as Gödel)

(def) $\quad \mathtt{H}(x) \equiv (\forall Y)(\Box Y(x) \equiv \mathtt{P}(Y))$

(A3) $\quad \mathtt{P}(\mathtt{H})$

(A4) $\quad \mathtt{P}(Y) \to \Box\mathtt{P}(Y)$ (as Gödel)

(def) $\quad \mathtt{Ess}_Y(x) \equiv (\forall Y)(\Box Y(u) \equiv \Box(X \subseteq Y))$

(def) $\quad \mathtt{NE}(x) \equiv (\forall Y)(\mathtt{Ess}_Y(x) \to \Box(\exists u)Y(u))$

(A5) $\quad \mathtt{P}(\mathtt{NE})$.

Definition 8. $\mathcal{AO}$ is the system of axioms (A1)–(A5) and definitions of $\mathtt{H}$, $\mathtt{Ess}$, $\mathtt{NE}$ just presented (Anderson ontological). $\mathcal{AO}_0$ is the subsystem of axioms (A1)–(A3) plus the definition of $\mathtt{H}$.

In [9, 8] I showed that assuming our logic as above, (A4)–(A5) are redundant since (A1)–(A3) prove both $\Box(\exists x)\mathtt{H}(x)$ and also (A4) and (A5) (using the definitions of $\mathtt{Ess}$, $\mathtt{NE}$). In other words, $\mathcal{AO}$ and $\mathcal{AO}_0$ are equivalent in the logic $\mathcal{L}_{\mathrm{ont}}$. But in [2] the authors stress that Anderson's proof works in a finer logic in which all the axioms are needed. We shall discuss this in Section 4; the next section is devoted to the analysis and emendation of $\mathcal{AO}_0$ in our present logic.

3. The emendation

Gödel's explanation for (A2) is:

> It is possible to interpret positive as perfective... The chief axiom runs then (essentially): A property is perfective iff it implies no negation of a perfective. [7, p. 434]

And in the footnote (same page) says:

> it need not be assumed always either φ or $\neg\varphi$ is positive.

This leads to the following suggestion of mine: *replace* the first two axioms by the following axiom:

(A12) $\quad (\mathtt{P}(X) \,\&\, \Box(X \subseteq Y)) \to \neg\mathtt{P}(\neg Y).$

(Recall that $X \subseteq Y$ stands for $(\forall u)(X(u) \to Y(u)$.) We shall show that (A12) + (A3) proves $\Box(\exists x)\mathtt{H}(x)$, for an appropriate definition of H. Evidently then we can prove the following:

LEMMA 1. (A12) *proves* $\mathtt{P}(X) \to \neg\mathtt{P}(\neg X)$ *(which is Anderson's* (A1)*) and* $\mathtt{P}(X) \to \Diamond(\exists u)X(u)$.

PROOF. Observe $\Box(X \subseteq X)$ and $\Box(\forall x)\neg X(x) \to \Box(X \subseteq \neg X)$ hence $((\mathtt{P}(X) \,\&\, \Box(\forall x)\neg X(x)) \to \neg\mathtt{P}(\neg\neg X)$, thus $(\mathtt{P}(X) \,\&\, \Box(\forall x)\neg X(x)) \to \neg\mathtt{P}(X)$, contradiction. ■

DEFINITION 9. Let H be defined as follows:

$$\mathtt{H}(u) \equiv (\forall X)(\Box X(u) \equiv (\exists Y)(\Box(Y \subseteq X) \,\&\, \mathtt{P}(Y))).$$

(A3) is $\mathtt{P}(\mathtt{H})$, as before.

$\mathcal{AO}'_0$ is the theory with axioms (A12), (A3) and the last definition of H (fragment of $\mathcal{AO}$ modified).

Evidently, $\mathtt{H}(u)$ implies that u necessarily has all positive properties (since $\mathtt{P}(Y) \,\&\, \Box(Y \subseteq Y)$ implies $\Box Y(u)$ by the definition of H; moreover, if $\Box X(u)$) then X is necessarily implied by a positive property $((\exists Y)(\mathtt{P}(Y) \,\&\, \Box(Y \subseteq X))$.

LEMMA 2. $\mathcal{AO}'_0$ *proves* $\mathtt{H}(u) \to \Box\mathtt{H}(u)$.

PROOF. Assume $\mathtt{H}(u)$; by (A3), H is positive. Since $\mathtt{H}(u)$ implies $\Box Y(u)$ for each positive y, we get $\Box\mathtt{H}(u)$ as claimed. ■

THEOREM 2. *In* $\mathcal{L}_{\mathrm{ont}}$, $\mathcal{AO}'_0$ *proves* $\Box(\exists u)\mathtt{H}(u)$.

PROOF. Conclude $(\exists u)\mathtt{H}(u) \to (\exists u)\Box\mathtt{H}(u)$, $\Diamond(\exists u)\mathtt{H}(u) \to \Diamond(\exists u)\Box\mathtt{H}(u) \to (\exists u)\Diamond\Box\mathtt{H}(u) \to (\exists u)\Box\mathtt{H}(u)$. In our logic $(\exists u)\Box\mathtt{H}(u) \to \Box(\exists u)\mathtt{H}(u)$ is a valid implication, which gives the desired conclusion

$$\Box(\exists u)\mathtt{H}(u).$$

■

Example. Let us present some models of the system. Let $M \neq \emptyset$, $W \neq \emptyset$ and let $Prop = 2^{M\times W}$ (the set of all mappings $f\colon M \times W \to \{0,1\}$). Let $h \in M$ and let $r_{\mathtt{H}}(u,v) = 1$ iff $u = h$ (for all w). Note $r_{\mathtt{H}} \in Prop$. Let $r_{\mathsf{P}}\colon Prop \times W \to \{0,1\}$ be such that $r_{\mathsf{P}}(r_{\mathtt{H}}, w) = 1$ for all w, and for any $f \in Prop$ and any w, $r_{\mathsf{P}}(f,w) = 1$ implies that $f(h,w') = 1$ for all w'. Then $\boldsymbol{K} = (W, M, Prop, r_{\mathtt{H}}, r_{\mathsf{P}})$ is a model of (A12), (A3); h is the godlike object of the model. Particular choices for r_{P}:

(i) $r_{\mathtt{H}}$ (the property of being godlike) is the only positive property, for each w

(ii) positive are exactly all properties that h necessarily has, for all w;

(iii) any choice including (i) and contained in (ii), i.e. for each w, $r_{\mathtt{H}}$ is positive and h has all positive properties. For example, for w_1, only $r_{\mathtt{H}}$ is positive, for w_2 all properties sub (ii) all positive. (Thus $\mathcal{AO}'_0$ does *not* prove $\mathsf{P}(X) \to \Box\mathsf{P}(X)$; but see the theorem below.)

Remark 3. I find the axiom (A12) much more natural than Anderson's (A1), (A2). On the other hand, the systems are closely related. Recall that an interpretation of a theory T_1 in a theory T_2 is *faithful* if for each sentence φ of T_1, T_1 proves φ iff T_2 proves φ^*. (Cf. Gödel's interpretation of the Zermelo-Fraenkel set theory $ZF + V = L$ with the axiom of constructibility in ZF itself)

THEOREM 3. *Anderson's* $\mathcal{AO}_0$ *has a faithful interpretation in* $\mathcal{AO}'_0$ *(both with full comprehension).*

PROOF. The interpretation is identical except for P define $\mathsf{P}^{\#}(X) \equiv (\exists Y)(\mathsf{P}(Y) \& \Box(Y \subseteq X))$ (property is #-positive if there is a positive property Y necessarily implying X). Let g be such that $\mathtt{H}(g)$; if $\mathsf{P}^{\#}(X)$ then $\Box X(g)$, thus $\neg\mathsf{P}^{\#}(\neg X)$ — this is (Anderson's) (A1) for $\mathsf{P}^{\#}$. (A2) is evident from the definition of $\mathsf{P}^{\#}$. Further we obviously have $\mathtt{H}(u) \equiv (\forall X)(\mathsf{P}^{\#}(X) \equiv \Box X(u))$ and $\mathsf{P}^{\#}(\mathtt{H})$. Thus # interprets Anderson in our system. On the other hand, in Anderson we get $\mathsf{P}(X) \equiv \mathsf{P}^{\#}(X)$, thus # as an interpretation of Anderson in Anderson is identity. Thus # is a faithful interpretation of our system in Anderson. ■

Remark 4. In [9, 8] I have proved that the original system by Gödel with cautious comprehension has a faithful interpretation in Anderson's $\mathcal{AO}'_0$ with full comprehension (in our logic). Thus Gödels original system with cautions comprehension has a faithful interpretation in our $\mathcal{AO}'_0$.

4. A version with variable domains

In his analysis Anderson [1] uses a weaker (more general) logic (introduced in [3]) admitting that the set of (actually existing) individuals varies from one possible world to another. This possibility is studied in depth in [5]. Wanting to keep the formalism as simple as possible we use a technique of Fitting and understand the variable domain logic as the extension of our logic $\mathcal{L}_{\mathrm{ont}}$ by a new predicate E of actual existence:

DEFINITION 10. The logic $\mathcal{L}_{\mathrm{ont,E}}$ is the extension of $\mathcal{L}_{\mathrm{ont}}$ by a unary predicate E applied to individuals (reading $\mathtt{E}(x)$ "x actually exists") and by the axiom of non-emptiness

$$\Box(\exists x)\mathtt{E}(x).$$

For simplicity we also accept the axiom $x = y \to \Box(x = y)$ (enabling absolute interpretation of equality between individuals; this is eliminable).

Following Fitting we introduce quantifiers $\forall^{\mathrm{E}}$, $\exists^{\mathrm{E}}$ as follows:

$$\begin{array}{lll} (\forall^{\mathrm{E}}x)\varphi & \text{stands for} & (\forall x)(\mathtt{E}(x) \to \varphi), \\ (\exists^{\mathrm{E}}x)\varphi & \text{stands for} & (\exists x)(\mathtt{E}(x)\ \&\ \varphi). \end{array}$$

We also write $X \subseteq^{\mathrm{E}} Y$ for $(\forall^{\mathrm{E}}u)(X(u) \to Y(u))$.

DEFINITION 11. In $\mathcal{L}_{\mathrm{ont,E}}$, Anderson's system $\mathcal{AOE}$ (Anderson ontological with the existence predicate) is as follows:

(A1)	$\mathtt{P}(X) \to \neg\mathtt{P}(\neg X)$	(as above)
(A2)$^{\mathrm{E}}$	$(\mathtt{P}(X)\ \&\ \Box(X \subseteq^{\mathrm{E}} Y)) \to \mathtt{P}(Y)$	
(def)	$\mathtt{H}(u) \equiv (\forall Y)(\Box Y(u) \equiv \mathtt{P}(Y))$	(as above)
(A3)	$\mathtt{P}(\mathtt{H})$	(as above)
(A4)	$\mathtt{P}(X) \to \Box\mathtt{P}(X)$	(as above)
(def)	$\mathtt{Ess}_X(u) \equiv (\forall Y)(\Box Y(u) \equiv \Box(X \subseteq^{\mathrm{E}} Y))$	
(def)	$\mathtt{NE}(u) \equiv (\forall Y)(\mathtt{Ess}_Y(u) \to \Box(\exists^{\mathrm{E}} X)Y(x))$	
(A5)	$\mathtt{P}(\mathtt{NE})$	(as above, of course for the new NE).

Our modification $\mathcal{AOE}'$ consists in replacing (A1), (A2) by the following axiom

$(\mathrm{A12})^{\mathrm{E}}$ $\quad (\mathtt{P}(X) \,\&\, \Box(X \subseteq^{\mathrm{E}} Y)) \to \neg\mathtt{P}(\neg Y)$

and changing the definition of H to be

$$\mathtt{H}(u) \equiv (\forall Y)(\Box Y(u) \equiv (\exists Z)(\mathtt{P}(Z) \,\&\, \Box(Z \subseteq^{\mathrm{E}} Y)).$$

Our aim is to prove

THEOREM 4. *$\mathcal{AOE}'$ proves* $\Box(\exists^{\mathrm{E}} u)\mathtt{H}(u)$.

We check that a natural modification of the proof by Gödel and Anderson works.

LEMMA 3. (1) $(\mathrm{A12})^{\mathrm{E}}$ *implies* $\mathtt{P}(X) \to \neg\mathtt{P}(\neg X)$.
(2) $(\mathrm{A12})^{\mathrm{E}}$ *implies* $\mathtt{P}(X) \to \Diamond(\exists^{\mathrm{E}} u)X(u)$
(3) $(\mathrm{A12})^{\mathrm{E}}$, (A3) *implies* $\mathtt{H}(u) \to \Box\mathtt{H}(u)$.
(4) $(\mathrm{A12})^{\mathrm{E}}$, (A3), (A4) *implies* $\mathtt{P}(Z) \to \Box(H \subseteq^{\mathrm{E}} Z)$.

PROOF. (1) is evident.

(2) Assume $\neg\Diamond(\exists^{\mathrm{E}} u)X(u)$, i.e. $\Box(\forall^{\mathrm{E}} u)\neg X(u)$. Then $\Box(X \subseteq^{\mathrm{E}} \neg X)$, thus $\mathtt{P}(X) \to \neg\mathtt{P}(\neg\neg X)$, hence $\mathtt{P}(X) \to \neg\mathtt{P}(X)$, which gives $\neg\mathtt{P}(X)$.

(3) Since $\Box(\mathtt{H} \subseteq^{\mathrm{E}} \mathtt{H})$ and $\mathtt{P}(\mathtt{H})$ we get $\mathtt{H}(u) \to \Box\mathtt{H}(u)$ from the definition of H.

(4) We have $\mathtt{P}(Z) \to (\forall^{\mathrm{E}} u)(\mathtt{H}(u) \to \Box Z(u))$; using (S5) delete the box and apply necessitation, getting $\Box\mathtt{P}(Z) \to \Box(\mathtt{H} \subseteq^{\mathrm{E}} Z)$. Then (A4) gives the result. ■

LEMMA 4. $(\mathrm{A12})^{\mathrm{E}}$, (A3), (A4) *implies* $\mathtt{H}(u) \to \mathtt{Ess}_{\mathtt{H}}(u)$.

PROOF. Assume $\mathtt{H}(u)$ and $\Box Y(u)$; then for some Z, $\mathtt{P}(Z) \,\&\, \Box(Z \subseteq^{\mathrm{E}} Y)$. By the preceding lemma, $\Box(H \subseteq^{\mathrm{E}} Z)$ and hence $\Box(\mathtt{H} \subseteq^{\mathrm{E}} Y)$. Consequently, if $\mathtt{H}(u)$ and $\Box(\mathtt{H} \subseteq^{\mathrm{E}} Y)$ then (since $\mathtt{P}(\mathtt{H})$), $\Box Y(u)$ by the definition of H. ■

PROOF OF THEOREM 4. First assume $\mathtt{H}(u)$; then $\mathtt{Ess}_{\mathtt{H}}(u)$ and $\mathtt{NE}(u)$ (since NE is positive, we have $\Box\mathtt{NE}(u)$). Thus, by the definition of NE, we get $\mathtt{H}(u) \to \Box(\exists^{\mathrm{E}} x)\mathtt{H}(x)$, and therefore by generalizing by $(\forall^{\mathrm{E}} u)$ and moving the quantifier $(\forall^{\mathrm{E}} u)$ we get

$$(\exists^{\mathrm{E}} u)\mathtt{H}(u) \to \Box(\exists^{\mathrm{E}} x)\mathtt{H}(x),$$
$$\Diamond(\exists^{\mathrm{E}} u)\mathtt{H}(u) \to \Diamond\Box(\exists^{\mathrm{E}} x)\mathtt{H}(x),$$
$$\Diamond(\exists^{\mathrm{E}} u)\mathtt{H}(u) \to \Box(\exists^{\mathrm{E}} x)\mathtt{H}(x),$$

which gives $\Box(\exists^{\mathrm{E}}x)\mathtt{H}(x)$. This completes the proof; our emendation (use of $(\mathrm{A12})^{\mathrm{E}}$) works. ■

5. Yet another emendation

Continuing the investigation of the last system, let us prove the following:

LEMMA 5. *$\mathcal{AOE}$ proves $\mathtt{P}(H\ \&\ E)$, i.e. being an actually existing godlike individual is positive.*

PROOF. Note that Fitting [5] has this fact too. It is enough to prove $\mathtt{H}(x) \to \mathtt{E}(x)$, hence $\Box(\forall x)(\mathtt{H}(x) \equiv (\mathtt{H}(x)\,\&\,\mathtt{E}(x)))$, (thus $\Box(\mathtt{H} = (\mathtt{H}\,\&\,\mathtt{E}))$; then $\mathtt{P}(\mathtt{H})$ implies $\mathtt{P}(\mathtt{H}\,\&\,\mathtt{E})$. Assume $\mathtt{H}(x)$ and let $(\forall y)\Box(Y(y) \equiv y = x)$ (by comprehension). Since $\Box(x = x)$, we get $\Box Y(x)$ and by the above $\mathtt{H}(x)$ implies $\mathtt{Ess}_{\mathtt{H}}(x)$, thus $\Box(\mathtt{H} \subseteq^{\mathrm{E}} Y)$, which gives $\mathtt{H}(x) \to \Box(\forall u)((\mathtt{H}(u)\&\mathtt{E}(u)) \to u = x)$. But we know that $\Box(\exists^{\mathrm{E}}u)\mathtt{H}(u)$, hence $\mathtt{H}(x) \to \Box(\exists u)(\mathtt{E}(u)\ \&\ \mathtt{H}(u)\ \&\ u = x)$, which gives $\mathtt{H}(x) \to (\exists u)(\mathtt{E}(u)\ \&\ u = x)$, i.e. $\mathtt{H}(x) \to \mathtt{E}(x)$. ■

This leads us to another emendation: instead of introducing Ess, NE, (A4) and (A5), just modify (A3) to the following:

DEFINITION 12. $(\mathrm{A3})^{\mathrm{E}}$ is the axiom $\mathtt{P}(\mathtt{H}\,\&\,\mathtt{E})$. $\mathcal{AOE}'_0$ (modified fragment of Anderson ontological with existence predicate) is the theory $(\mathrm{A12})^{\mathrm{E}}$, $(\mathrm{A3})^{\mathrm{E}}$ plus the definition of H as in $\mathcal{AOE}$.

THEOREM 5. *$\mathcal{AOE}'_0$ proves $\Box(\exists^{\mathrm{E}}u)\mathtt{H}(u)$.*

PROOF. Note the provability of $\Diamond(\exists^{\mathrm{E}}u)\mathtt{H}(u)$. Now $\mathtt{H}(u)$ implies $\Box(\mathtt{H}(u)\ \&\ \mathtt{E}(u))$ (by $(\mathrm{A3})^{\mathrm{E}}$ and the definition of H), hence $\Box(\exists x)(\mathtt{E}(x)\,\&\,\mathtt{H}(x))$, which is $\Box(\exists^{\mathrm{E}}x)\mathtt{H}(x)$. ■

Our final aim is to find a natural subtheory $\mathcal{AOE}''$ of $\mathcal{AOE}'$ dealing with necessary existence and existence and equivalent to $\mathcal{AOE}'_0$ (i.e. $(\mathrm{A12})^{\mathrm{E}}$, $(\mathrm{A3})^{\mathrm{E}}$ plus the corresponding definition of H). Here you have it:

DEFINITION 13. $\mathcal{AOE}''$ in the following modification of $\mathcal{AOE}'$:

- axioms $(\mathrm{A12})^{\mathrm{E}}$, (A3),
- definition of H as in $\mathcal{AOE}'$,
- definition $\mathtt{P}^{\#}(Y) \equiv (\exists Z)(\mathtt{P}(Z)\,\&\,\Box(Z \subseteq Y))$
  (thus the definition of H can be written as

$$\mathtt{H}(u) \equiv (\forall Y)(\Box Y(u) \equiv \mathtt{P}^{\#}(Y))),$$

and

$(\mathrm{A4})^{\#}$ $\qquad \mathtt{P}(Y) \to \Box\mathtt{P}^{\#}(Y)$

$(\mathrm{A5})^{\#}$ $\qquad \mathtt{P}^{\#}(\mathtt{NE})$.

THEOREM 6. *In the logic* $\mathcal{L}_{\mathrm{ont},\mathtt{E}}$, *the theories* $\mathcal{AOE}'_0$ *(extended by the definition of* $\mathtt{P}^{\#}$ $\mathtt{Ess}$, $\mathtt{NE}$*) and* $\mathcal{AOE}''$ *are equivalent, i.e.*

(1) $\mathcal{AOE}'_0$ *proves* $(\mathrm{A4})^{\#}$, $(\mathrm{A5})^{\#}$,

(2) $\mathcal{AOE}''$ *proves* $(\mathrm{A3})^{\mathtt{E}}$.

PROOF. We work in the theory $\mathcal{AOE}'_0$ with axioms $(\mathrm{A12})^{\mathtt{E}}$, $(\mathrm{A3})^{\mathtt{E}}$; recall that the theory proves $\Box(\exists^{\mathtt{E}}x)\mathtt{H}(x), \mathtt{H}(u) \to \Box\mathtt{H}(u)$. For simplicity, introduce a constant $\mathbf{g}$ and postulate $\mathtt{H}(\mathbf{g})\ \&\ \mathtt{E}(\mathbf{g})$. Also recall the definition $\mathtt{Ess}_X(u) \equiv (\forall Z)(\Box Z(u) \equiv \Box(Y \subseteq^{\mathtt{E}} Z)$. The theorem is a consequence of the following lemmas. ∎

LEMMA 6. $\mathtt{P}^{\#}(Y) \to \Box\mathtt{P}^{\#}(Y)$ $\qquad$ *(which is* $(\mathrm{A4})^{\#}$*).*

PROOF. $\mathtt{H}(u) \to (\mathtt{P}^{\#}(Z) \equiv \Box Z(u))$, hence $\Box\mathtt{H}(u) \to \Box\mathtt{P}^{\#}(Z) \equiv \Box Z(u)))$ thus $\mathtt{H}(u) \to \Box\mathtt{H}(u)$, $(\exists u)\mathtt{H}(u) \to (\mathtt{P}^{\#}(Z) \equiv \Box\mathtt{P}^{\#}(Z))$ thus $\mathtt{P}^{\#}(Z) \equiv \Box\mathtt{P}^{\#}(Z)$. ∎

LEMMA 7. $\mathtt{Ess}_{\mathtt{H}}(\mathbf{g})$

PROOF. Let $\Box Y(\mathbf{g})$, thus $\mathtt{P}^{\#}(Y)$ and for some Z, $\mathtt{P}(Z)\ \&\ \Box(Z \subseteq^{\mathtt{E}} Y)$. From the definition of $\mathtt{H}$ we get

$$\mathtt{H}(u) \to (\mathtt{P}(Z) \to \Box Z(u)) \to (\mathtt{P}(Z) \to Z(u)),$$
$$\mathtt{P}(Z) \to (\forall u)(\mathtt{H}(u) \to Z(u)) \to (\forall^{\mathtt{E}}u)(\mathtt{H}(u) \to Z(u)),$$
$$\mathtt{P}^{\#}(Y) \to (\forall^{\mathtt{E}}u)(\mathtt{H}(u) \to Y(u)),$$
$$\Box\mathtt{P}^{\#}(Y) \to \Box(\mathtt{H} \subseteq^{\mathtt{E}} Y)$$

and an application of $(\mathrm{A4})^{\#}$ gives

$$\mathtt{P}^{\#}(Y) \to \Box(\mathtt{H} \subseteq^{\mathtt{E}} Y).$$

Thus $(\forall Y)(\Box Y(y) \to \Box(\mathtt{H} \subseteq^{\mathtt{E}} Y))$; conversely, $\mathtt{H}(g)\ \&\ \mathtt{E}(g)$ and $\Box(\mathtt{H} \subseteq^{\mathtt{E}} Y)$ immediately gives $\Box Y(g)$. This proves $\mathtt{Ess}_{\mathtt{H}}(g)$. ∎

LEMMA 8. $\mathtt{Ess}_Y(u)\ \&\ \mathtt{Ess}_Z(u) \to \Box(Y \subseteq^{\mathtt{E}} Z)$ $\quad$ *(and* $\Box(Z \subseteq^{\mathtt{E}} Y)$*).*

PROOF. $\mathtt{Ess}_Z(u) \to (\Box Z(u) \equiv \Box(Z \subseteq^{\mathtt{E}} Z))) \to \Box Z(u)$
$\mathtt{Ess}_Y(u) \to (\Box Z(u) \equiv \Box(Y \subseteq^{\mathtt{E}} Z))) \to \Box(Y \subseteq^{\mathtt{E}} Z)$. ∎

COROLLARY 1. $(\forall Y)(\mathtt{Ess}_Y(\mathtt{g}) \to \Box(\mathtt{H} \subseteq^{\mathtt{E}} Y))$.

LEMMA 9. $\mathtt{P}^{\#}(\mathtt{NE})$

PROOF. Lemma 8 and $\Box(\exists^{\mathtt{E}}u)\mathtt{H}(u)$ gives $(\forall Y)(\mathtt{Ess}_Y(\mathtt{g}) \to \Box(\exists^{\mathtt{E}}u)Y(u))$, then $\mathtt{NE}(\mathtt{g})$. Apply necessitation: $\Box\mathtt{NE}(\mathtt{g})$, hence $\mathtt{H}(\mathtt{g})$ gives $\mathtt{P}^{\#}(\mathtt{NE})$.

This completes the proof of the first claim of our theorem. To prove the second, i.e. to show the provability of $(\mathrm{A3})^{\mathtt{E}}$, in $\mathcal{AOE}''$ with the axioms $(\mathrm{A12})^{\mathtt{E}}$, (A3), $(\mathrm{A4})^{\#}$, $(\mathrm{A5})^{\#}$ just check the proof of the lemma at the beginning of this section; this needs also checking the proofs from Section 4 with $(\mathrm{A4})^{\#}$, $(\mathrm{A5})^{\#}$ instead of (A4), (A5) (but it is routine). ∎

Remark 5. In [9] I mention that the proof in $\mathcal{AO}_0$ works even the belief logic KD45 (S5 with $\Box\varphi \to \varphi$ weakened to $\Box\varphi \to \Diamond\varphi$ and with Barcan formulas added as axioms); one can prove $\Box(\exists x)\mathtt{H}(x)$ (it is believed that there is a godlike being) but not $(\exists x)\mathtt{H}(x)$. The same appears to work for our proof in $\mathcal{AO}'_0$, i.e. from (A12), (A3), and even for the proof of $\Box(\exists^{\mathtt{E}}x)\mathtt{H}(x)$ in $\mathcal{AOE}'_0$, i.e. from $(\mathrm{A12})^{\mathtt{E}}$, $(\mathrm{A3})^{\mathtt{E}}$; but the proof from (A1) $(\mathrm{A2})^{\mathtt{E}}$ (or $(\mathrm{A12}^{\mathtt{E}})$, (A3)–(A5)) does use the schema $\Box\varphi \to \varphi$ and I do not know if it can be eliminated. Thus some room remains for further analysis (including construction of some interesting models in the style of [8]).

Summary

Our main results are the following:

1. In the logic $\mathcal{L}_{\mathrm{ont}}$ with fixed domains, the system $\mathcal{AO}'_0$ of axioms

(A12) $\quad (\mathtt{P}(X) \,\&\, \Box(X \subseteq Y)) \to \neg\mathtt{P}(\neg Y)$

(def1) $\quad \mathtt{H}(x) \equiv (\forall Y)(\Box Y(x) \equiv (\exists Z)(\mathtt{P}(Z) \,\&\, \Box(Z \subseteq Y))$

(A3) $\quad \mathtt{P}(\mathtt{H})$

proves $\Box(\exists x)\mathtt{H}(x)$. (Modified fragment of Anderson.)

2. In the logic $\mathcal{L}_{ont,E}$ with the existence predicate E, the system $\mathcal{AOE}'_0$ of axioms

$(\mathrm{A12})^{\mathtt{E}}$ $\quad (\mathtt{P}(X) \,\&\, \Box(X \subseteq^{\mathtt{E}} Y)) \to \neg\mathtt{P}(\neg Y)$

(def 1) $\quad \mathtt{H}(x) \equiv (\forall Y)(\Box Y(x) \equiv (\exists Z)(\mathtt{P}(Z) \,\&\, \Box(Z \subseteq^{\mathtt{E}} Y))$

$(\mathrm{A3})^{\mathtt{E}}$ $\quad \mathtt{P}(H \,\&\, E)$.

proves $\Box(\exists^{\mathtt{E}}x)\mathtt{H}(x)$. (Modified fragment of Anderson with existence.)

3. Alternatively, you may take $(\mathrm{A}12)^{\mathrm{E}}$, (def1), (A3) as above together with Anderson's definitions and axioms (A4), (A5), i.e.

(def2)	$\mathtt{Ess}_X(u) \equiv (\forall Y)(\Box Y(u) \equiv \Box(X \subseteq^{\mathrm{E}} Y)$
(def3)	$\mathtt{NE}(u) \equiv (\forall Y)(\mathtt{Ess}_Y(u) \to \Box(\exists^{\mathrm{E}} x)Y(x))$
(A4)	$\mathtt{P}(X) \to \Box\mathtt{P}(X)$
(A5)	$\mathtt{P}(\mathtt{NE})$

This axiom system $\mathcal{AOE}'$ proves $\Box(\exists^{\mathrm{E}} x)\mathrm{H}(x)$, as well as it proves $(\mathrm{A}3)^{\mathrm{E}}$. (Modified Anderson with existence.)

Our axiom (A12) (or $(\mathrm{A}12)^{\mathrm{E}}$) seems to correspond exactly to Gödel's ideas and is free from the unwanted consequence of the original axiom (A2).

This seems to show that the Axioms $(\mathrm{A}12)^{\mathrm{E}}$, $(\mathrm{A}3)^{\mathrm{E}}$ give a rather pleasant emendation of Gödel's axioms.

6. Conclusion

Admittedly, our emendation is simple and small. Nevertheless, I hope it makes the assumptions more natural and acceptable. Needless to say, the relevance for theology and religion remains very limited, mainly as a sort of mathematical analysis or modelling of old attempts to prove the existence of God. (See [11] for a fascinating discussion of the question of the existence of God.) On the other hand, I hope that the reader will agree that the analysis of Gödel's proof is an unusual exercise in modal logic.

References

[1] ANDERSON, C. A., 'Some emendations of Gödel's ontological proof', *Faith and Philosophy* 7 (1990), 291–303.

[2] ANDERSON, C. A., M. GETTINGS, 'Gödel's ontological proof revisited', in Hájek (ed.), *Gödel'96*, Springer 1996, pp. 167–172.

[3] COCCHIARELLA, N. B., 'A completeness theorem in second order modal logic', *Theoria* 35 (1969), 81–103.

[4] ESSLER, W. K., 'Gödels Beweis', in F. Ricken (ed.) *Klassische Gottesbeweise in der Sicht der gegenwärtigen Logik und Wissenschaftstheorie*, Verlag W. Kohlhammer 1991.

[5] FITTING, M., *Types, Tableaux and Gödel's God.* Draft of a book, August 2000, see `http://comet.lehman.cuny.edu/fitting`

[6] FUHRMANN, A., '"G" für Gödel: Kurt Gödels axiomatische Theologie', in: P. Schroeder-Heister et al. (ed.) *Logik in der Philosophie.* Book in preparation.

[7] GÖDEL, K., *Collected Works III*, ed. by S. Feferman et al., Oxford Univ. Press 1995.

[8] HÁJEK, P., *Der Mathematiker und die Frage der Existenz Gottes*, should appear in Buldt and Schimanovich, ed.: Wahrkeit und Beweisbarkeit (the book was planned to appear in 1995(!) expected to appear in 2001).

[9] HÁJEK, P., 'Magari and others on Gödel's ontological proof', in Ursini et al. (ed.), *Logic and Algebra*, Marcel Dekker 1996, pp. 125–136.

[10] HUGHES, G. E., and M. J. CRESSWELL, *An Introduction to Modal Logic.* Methuen London 1968/1972.

[11] KÜNG, J., *Existiert Gott?*, Piper Verlag, Stuttgart 1978.

[12] SOBEL, J. H., 'Gödel's ontological proof', in J. V. Thomson, (ed.), *On Being and Saying. Essays for Richard Cartwright*, The MIT Press 1987.

PETR HÁJEK
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2
182 07 Prague, Czech Republic
hajek@cs.cas.cz