CHAPTER 14
Gödel's Ontological Proof and Its Variants

Petr Hájek

In the early 1970s, we learned that Gödel had produced a proof of the existence of God after he showed it to Dana Scott, who discussed it in a seminar at Princeton. Notes began to circulate, and the first public analysis of the proof was performed by Sobel (1987). Only after Gödel's death, in the third volume of his collected works (Gödel, 1995), was the proof finally published. It is just one page, preceded by an extensive and very informative introduction by Adams (Gödel, 1995). The volume also contains notes from Gödel's *Nachlass*, dated 1940, containing his first drafts of the proof. Presently, there exist several papers on the topic, including two extremely interesting monographs (Sobel, 2004; Fitting, 2002). A very important variant of Gödel's system, resulting from the first criticisms made of it, was generated by Anderson (1990). Anderson's variant will play an important role in this chapter.

Gödel is famous for his completeness and incompleteness theorems as well as for his work with set theory, so his ontological proof has never received the same attention. The proof belongs to the family of *ontological* arguments, that is to say, arguments that try to establish the existence of God by relying only on pure logic. Such arguments were presented by Anselm (1033–1109), Descartes (1598–1650), Leibniz (1646–1716), and others, and Gödel is known to have studied particularly Leibniz's works (the two books mentioned earlier are recommended for information on the old ontological proofs and their relation to Gödel's proof). Here let us only mention that even if the old proofs were clearly not formal proofs – in the sense of formal logic – and did not use modalities, they show a striking similarity in their form: they (try to) argue that God exists (actually, really, necessarily) if God is possible (consistent, present in our mind), and then to prove that God is indeed possible. Details and weaknesses of these proofs are analyzed particularly in Sobel's book. Hartshorne (1962) seems to be the first to formalize the modal substance of these proofs (see later); Gödel's proof has this structure too.

This chapter concentrates on the formal (mathematical) aspects of Gödelian ontological proofs; philosophical aspects are completely ignored. However, at the end, we shall comment on the (possible) relevance of ontological proofs to religious faith. Our plan is as follows: in Section 14.1, we specify the formal logic used (first-order modal S5); Section 14.2 discusses Gödel's original version (as presented by Scott), and we analyze the first criticisms of it (made by Sobel and Magari); Section 14.3 focuses on Anderson's variant and its variants; in Section 14.4, we analyze criticism formulated by Oppy; Section 14.5 contains "Miscellanea," for example, the use of different modal logics, the question of proving the devil's existence, a recent variant of Gödel's original proof by Koons, and some other matters; in Section 14.6, we ask what we have learned about Gödel; finally, in Section 14.7, we consider the importance of ontological proofs for religious belief. Section 14.8 concludes.

Let us close this introduction, for the reader's pleasure, by quoting at least a part of the famous Anselm's Proslogion II as an example of an ontological proof (Hopkins and Richardson, 1974):

> Therefore, Lord, Giver of understanding to faith, grant me to understand – to the degree You deem best – that you exist, as we believe, and that you are what we believe You to be. Indeed, we believe You to be something than which nothing greater can be thought… even the Fool is convinced that something than which nothing greater can be thought exists at least in his understanding; for when he hears of this being, he understands [what he hears], and whatever is understood is in the understanding. But surely that than which a greater cannot be thought cannot be only in the understanding. For if it were only in the understanding, it could be thought to exist also in reality – which is greater [than existing only in the understanding]. Hence, without doubt, something than which a greater cannot be thought exists both in the understanding and in reality.

14.1 The Logic Used

We shall use the best-known modal logic S5 (for other possibilities, see remarks in Section 14.5). The reader is assumed to be familiar with classical propositional and first-order predicate logic as well as with (at least) propositional S5 (for detailed treatment of first-order S5, see Fitting and Mendelsohn, 1990). As usual, we use $\Box$ for the modality of necessity and $\Diamond$ for possibility. Let us recall the propositional axioms of S5. Besides the axioms of classical propositional logic, these are as follows

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$$

$$\Box\varphi \equiv \Box\Box\varphi$$

$$\Box\varphi \equiv \Diamond\Box\varphi$$

$$\Box\varphi \to \varphi.$$

Possibility is defined by

$$\Diamond\varphi \equiv \neg\Box\neg\varphi.$$

Deduction rules are modus ponens and necessitation: from φ infer $\Box\varphi$. Semantics come into play when considering Kripke models consisting of a nonempty set W of possible worlds and an evaluation w of each propositional variable in each possible world by a truth value 0 or 1. Connectives are evaluated in each possible world by the usual truth tables; $\Box\varphi$ is true in the model if and only if φ is true in each possible world. Propositional S5 is complete with respect to these models; a formula is provable if and only if it is an S5-tautology (which is true in all Kripke models). (In particular, recall that $\Diamond\Box\varphi = \Box\varphi$ is an S5-tautlogy.)

Let us illustrate the use of propositional S5 by presenting Hartshorne's proof. Let q stand for "God exists." Assuming that $\Diamond q$, and $q \to \Box q$ (Anselm's principle), then

$$\Box(q \to \Box q),$$

$$\Diamond q \to \Diamond\Box q,$$

$$\Diamond\Box q \to \Box q,$$

$$\Diamond q \to \Box q,$$

$$\Box q.$$

As stated earlier, this is, in some sense, the common structure of ontological proofs discussed by us. The question is how (or from what) to prove the two assumptions. Let's turn to the first-order S5. There are variables for an object's alias beings (lowercase letters) and for properties (uppercase latters), one constant G for the property being godlike, and one predicate P of properties (being a positive property). This can be understood either as a second-order logic or as a first-order, two-sorted logic with one binary predicate *Appl* with the formula $Appl(Y, x)$, abbreviated as $Y(x)$. (The reader may choose which version he or she likes.) We shall also use the extensional equality predicate = for beings (thus we assume $x = y \to \Box(x = y)$).

We shall discuss two variants of the logic: the first has Kripke models with fixed domain $(W, M, Prop, Appl, P_K, G_K)$, where W is a nonempty set of possible worlds, M a nonempty set of objects (beings), *Prop* a nonempty set of properties, $G_k \in Prop$, $P_K \subseteq Prop \times W$, and $Appl \subseteq Prop \times M \times W$. This has the following meaning: $(X, w) \in P_K$ means that X is a positive property in the world w and that $(X, a, w) \in Appl$ means that the being a has the property X in the world w. In particular, $(G_K, a, w) \in Appl$ means that a is a godlike object in the world w. Satisfaction is defined in the usual way, in particular, $\Box\varphi$ (φ being a sentence) is true in the model if and only if φ is true in each possible world.

Note that with these models, the predicate modal logic S5, with its usual axioms (just axioms, rules of classical predicate logic, of propositional modal S5, Barcan formula $\Box(\forall x)\varphi \equiv (\forall x)\Box\varphi$, and the rule of necessitation (from φ infer $\Box\varphi$)) is complete (see Fitting and Mendelsohn, 1990).

The second set of semantics has, in addition to the preceding, among the predicates a constant predicate E of (actual) existence; that $E(u)$ is true in a world, w means that the object u actually exists in this world. An axiom demands that in each world, at least one object actually exists. The formula $(\exists^E u)\varphi(u)$ stands for $(\exists u)(E(u) \wedge \varphi(u))$; similarly, $(\forall^E u)\varphi(u)$ stands for $(\forall u)(E(u) \to \varphi(u))$ (see, e.g., Fitting, 2002).

Furthermore, we assume the axiom schema with full comprehension. For each formula $\varphi(u,\ldots)$ (the dots standing for possible other free variables for beings or properties), and each variable Y not occurring in φ, the axiom says

$$(\forall\ldots)(\exists Y)\Box(\forall u)(Y(u) \equiv \varphi(u,\ldots)$$

(satisfying that φ is a property). Full comprehension demands restriction to Kripke models such that properties defined by formulas exist in the model. (One can also work with weaker systems not containing the full comprehension; see comments later.)

14.2 Gödel's Proof

We reproduce Gödel's original system (as presented by Scott) and his proof and prove some additional theorems in the system. Recall that $P(X)$ says that X is a positive property, and $G(a)$ says that a is a godlike being. Furthermore, $X \subseteq Y$ (defined later) says that X entails Y, X**Ess**a that the property X is an essence of a (Gödel's name), and $NE(a)$ that a necessarily exists. $\neg X$ denotes a property Y satisfying $(\forall a)(Y(a) \equiv \neg X(a))$. The axioms follow:

(G1) $P(X) \equiv \neg P(\neg X)$

(def) $X \subseteq Y \equiv (\forall u)(X(u) \to Y(u))$

(G2) $(P(X) \wedge \Box(X \subseteq Y)) \to P(Y)$

(def) $G(a) \equiv (\forall Y)(P(Y) \to Y(a))$ (Godlike)

(G3) $P(G)$

(G4) $P(Y) \to \Box P(Y)$

(def) $X\mathbf{Ess}a \equiv X(a) \wedge (\forall Y)(Y(a) \to \Box(X \subseteq Y))$

(def) $NE(a) \equiv (\forall X)(X\mathbf{Ess}a \to \Box(\exists a)X(a))$

(G5) $P(NE)$

In words, a property is positive if and only if its negation is not positive; a property necessarily entailed by a positive property is itself positive;

a positive property is necessarily positive; and both godlikeness and necessary existence are positive. This theory (over the predicate modal logic S5) will be denoted by $\mathcal{G}O$.

Theorem 14.2.1 $\mathcal{G}O$ *proves* $P(X) \rightarrow \Diamond(\exists a)X(a)$.

PROOF The following chain of implications is provable:
$\Box(\forall u)\neg X(u) \rightarrow \Box(\forall u)(X(u) \rightarrow \neg X(u)) \rightarrow (P(X) \rightarrow \neg P(X)) \rightarrow \neg P(X)$. Thus we have proved $\neg\Diamond(\exists a)X(a) \rightarrow \neg P(X)$, which, in turn, proves the theorem.

Remark $\mathcal{G}O$ proves $G(u) \equiv (\forall Y)(P(Y) \equiv Y(u))$.

Indeed, $\leftarrow$ is trivial; for $\rightarrow$, observe $(G(u) \wedge Y(u) \wedge \neg P(Y)) \rightarrow ((G(u) \wedge P\neg Y)) \rightarrow \neg Y(u)$, a contradiction.

Theorem 14.2.2 $\mathcal{G}O$ *proves* $G(u) \rightarrow G\mathbf{Ess}u$.

PROOF The following are provable:
$(G(u) \wedge Y(u)) \rightarrow P(Y) \rightarrow \Box P(Y)$; $P(Y) \rightarrow (\forall x)(G(x) \rightarrow Y(x))$, thus
$\Box P(Y) \rightarrow \Box(G \subseteq Y)$, hence $(G(u) \wedge Y(u)) \rightarrow \Box(G \subseteq Y)$.

Theorem 14.2.3 $\mathcal{G}O$ *proves* $\Box(\exists x)G(x)$

PROOF $\mathcal{G}O$ proves $\Diamond(\exists x)G(x)$ by Theorem 14.2.1, $G(x) \rightarrow \mathbf{NE}(x)$ because $P(\mathbf{NE})$, thus $G(x) \rightarrow \Box(\exists x)G(x)$ because $G\mathbf{Ess}x$. Take q to be $(\exists x)G(x)$ and apply Hartshorne's proof.

This completes Gödel's proof of the necessary existence of a godlike being. We present some further consequences of the axioms.

Theorem 14.2.4 $\mathcal{G}O$ *proves* $(G(x) \wedge Y(x)) \rightarrow \Box Y(x)$.

PROOF Let I_x be the property of being x, that is to say, satisfying $I_x(z) \equiv z = x$. We have to prove that $(G(x) \wedge Y(x)) \rightarrow \Box Y(x)$. We already have $\Box(\exists z)G(z)$ (using all the axioms of $\mathcal{G}O$). Now observe the following provable implications:

$$
\begin{aligned}
&(G(x) \wedge Y(x)) \rightarrow (P(Y) \wedge \Box P(Y)),\\
&\Box(G(x) \rightarrow (P(Y) \rightarrow (Y(x) \wedge I_x(x)))),\\
&\Box P(Y) \rightarrow \Box(G(x) \rightarrow (Y(x) \wedge I_x(x))),\\
&\Box P(Y) \rightarrow \Box(\exists z)(G(z) \wedge Y(z) \wedge I_x(z)),\\
&\Box P(Y) \rightarrow \Box(G(x) \wedge Y(x)),\\
&(G(x) \wedge Y(x)) \rightarrow \Box Y(x).
\end{aligned}
$$

Theorem 14.2.5 $\mathcal{G}O$ *proves* $G(u) \equiv (\forall Y)(P(Y) \equiv \Box Y(u))$.

This follows immediately from the preceding remark and theorem. This equivalence says that u is godlike if and only if the properties that u necessarily has are exactly all positive properties. It will play an important role in Anderson's variant of the ontological proof. (It appears that provability of this theorem in $\mathcal{G}O$ has not previously been observed.)

The next theorem is from Sobel (1987) and is called the "theorem on the collapse of modalities in $\mathcal{G}O$."

Theorem 14.2.6 *For each formula* φ, $\mathcal{G}O$ *proves* $(\varphi \rightarrow \Box\varphi)$.

PROOF The following chain of implications is provable:
$\varphi \rightarrow [G(x) \rightarrow (G(x) \wedge \varphi)] \rightarrow [G(x) \rightarrow \Box(G(x) \wedge \varphi)] \rightarrow [G(x) \rightarrow \Box\varphi]$.
Thus $G(x) \rightarrow (\varphi \rightarrow \Box\varphi)$, hence $(\exists x)G(x) \rightarrow (\varphi \rightarrow \Box\varphi)$, and therefore $\varphi \rightarrow \Box\varphi$.

Consequently, in $\mathcal{G}O$, for each formula φ, the formulas φ, $\Box\varphi$, and $\Diamond\varphi$ are mutually equivalent, and modalities become superfluous. This appears to be a significant disadvantage of $\mathcal{G}O$ and led Anderson to his emendation of the system (Anderson, 1990; he says there that, unfortunately, too much follows from these (Gödel's) axioms). But surprisingly, in a very recent paper, Sobel (2006) argues that it is possible that Gödel knew that his axioms lead to the collapse of modalities and accepted this for some philosophical reasons (see Adams in Gödel, 1990, 400). Nevertheless, our main attention will be paid to variants of the system not having a collapse of modalities.

Among the first critics of Gödel's proof, let us mention Magari (1988). He says, *"Teofili hanno spesso fornito ingegnosi argomenti … esistono anche teofobi* (*io lo sono di tutto cuore*)*"* (he declares himself to be a "theofob") and *"non è più facile ammettere gli assiomi che ammettere direttamente il teorema"* (it is not easier to accept the axioms than to accept the conclusion). This second quotation is interesting with respect to the possible religious significance of Gödel's proof. Magari claims that the sentence $\vdash \Box(\exists x)G(x)$ is provable using only the first three axioms (G1)–(G3) of Gödel, but this is false. (My counterexample in Hájek (1996) satisfies even Gödel's (G1)–(G4).) Finally, let us mention Fitting's observation (in his monograph) that *in* $\mathcal{G}O$, the axiom $P(G)$ (godlikeness is positive) can be equivalently replaced by $\Diamond(\exists x)G(x)$.

14.3 Anderson's Variant and Its Variants

Anderson's variant (mentioned at the end of the preceding section) consists first in weakening Gödel's axiom (G1) to $P(X) \rightarrow \neg P(\neg X)$. Thus, at most, one of the properties, X or $\neg X$, may be positive, but not both (but possibly neither of them). The second difference in Anderson's variant is made by changing the definitions of god-likeness and essence, keeping the axioms (G2)–(G5) and the definition of necessary existence as they are but using the new notions of godlikeness and necessary existence. Call these axioms now (A2)–(A5) (Anderson's version). The definition of godlikeness is now (def) $G(u) \equiv (\forall Y)(P(Y) \equiv \Box Y(u))$.

(Recall that we have shown that this formula is provable in $\mathcal{G}O$.) We postpone any discussion of Anderson's notion of essence and necessary existence because we shall show that in Anderson's system (which we may call $\mathcal{A}O$), they are redundant. (This was first shown in Hájek (1996)). For the ontological proof in the logic with fixed domains, we only need the following three axioms (the three first axioms in Anderson's theory; $X \supseteq Y$ stands for $(\forall u)(X(u) \to Y(u))$):

(Al) $P(X) \to \neg P(\neg X)$, where $(\neg X)(u) \equiv \neg X(u)$,

(A2) $(P(X) \wedge \Box(X \subseteq Y)) \to P(Y)$,

(A3) $P(G)$, where $G(u) = (\forall Y)(P(Y) = \Box Y(u))$.

In words, the negation of a positive property is not positive; a property necessarily implied by a positive property is positive; and godlikeness is positive (u is godlike if properties it necessarily has are exactly all positive properties). This theory will be denoted (in accordance with Hájek, 2002b) by $\mathcal{A}O_0$.

Theorem 14.3.1 $\mathcal{A}O_0$ *proves* $\Box(\exists x)G(x)$, $(\exists x)\Box G(x)$ *and also uniqueness:*

$(\forall x, y)((G(x) \wedge G(y)) \to x = y)$.

PROOF $\mathcal{A}O$ proves $P(X) \to \Diamond(\exists u)X(u)$ because assuming X is necessarily empty gives $\Box(X \subseteq \neg X)$; thus $P(X)$ gives $P(\neg X)$, which contradicts (Al). This proves $\Diamond(\exists x)G(x)$. Furthermore, $\mathcal{A}O_0$ proves $G(u) \to \Box G(u)$ using (A3) and the definition of G. From this follows $\Diamond(\exists u)G(u) \to \Box(\exists u)G(u)$ and hence $\Box(\exists u)G(u)$. Thus one gets $(\exists x)G(x)$, $(\exists x)\Box G(x)$. (Indeed, because $G(x) \to \Box G(x)$, we can conclude that $\Box(\exists x)G(x) \to \Box(\exists x)\Box G(x) \to (\exists x)\Box G(x)$.) To prove uniqueness, observe that for each x, there exists a property I_x such that $\Box(\forall y)(I_x(y) \equiv y = x)$ (being x). Clearly $\Box I_x(x)$ for all x; if $G(x)$, then $P(I_x)$; thus if $G(y)$, then $\Box I_x(y)$, hence $\Box(y = x)$.

Theorem 14.3.2 $\mathcal{A}O_0$ *proves* $P(Y) \to \Box P(Y)$ *and* $P(Y) \to P(\Box Y)$. *(The former formula is Anderson's axiom (A4).)*

PROOF First,

$G(g) \to (P(Y) \equiv \Box Y(g))$,
$\Box G(g) \to (\Box P(Y) \equiv \Box Y(g))$,
$\Box G(g) \to (\Box P(Y) \equiv \Box Y(g) \equiv P(Y))$,
$P(Y) \equiv \Box P(Y)$.

We prove the second formula. Indeed, $\Box(\Box Y \to Y)$ is a theorem, thus $P(\Box Y)$ implies $P(Y)$. Conversely, assuming $G(x)$, we get

$P(Y) \to \Box Y(x) \to \Box(\Box Y(x)) \to P(\Box Y)$.

Following Anderson, let us define (individual) essence and necessary existence as follows:

$$X\mathbf{Ess}u \equiv (\forall Y)(\Box Y(u) \equiv \Box(X \subseteq Y))$$

$$\mathbf{NE}(x) \equiv (\forall Y)(Y\mathbf{Ess}x \to \Box(\exists u)Y(u)).$$

Theorem 14.3.3 $\mathcal{A}O_0$ *proves* $(\forall x)I_x$ **Ess**x, $(\forall x)$**NE**(x) *and hence* P(**NE**). *(The last formula is Anderson's axiom (A5).)*

PROOF The first claim is evident from definition, the second from the provability of $\Box I_x(x)$ and from the provability of the formula $(X\mathbf{Ess}u \wedge Y\mathbf{Ess}u) \to (X \subseteq Y \wedge Y \subseteq X)$. The last formula follows from the fact that any property that all objects have is positive.

Summarizing, $\mathcal{A}O_0$ proves all the axioms (A1)–(A5) of Anderson's theory $\mathcal{A}O$.

Now let us turn to the second set of models, which has among the predicates, in addition to the preceding, a constant predicate E of (actual) existence. Recall that $E(u)$ is true in a world w means that the object u actually exists in this world and that the formula $(\exists^E u)\varphi(u)$ stands for $(\exists u)(E(u) \wedge \varphi(u)$; similarly, $(\forall^E u)\varphi(u)$ stands for $(\forall u)(E(u) \to \varphi(u))$ (see, e.g., Fitting, 2002).

Over this logic, we keep the axioms (Al), (A2), and (A3) and only change the definition of godlikeness:

$$G(x) \equiv (E(x) \wedge (\forall Y)(P(Y) \equiv \Box Y(u))).$$

The theory will be called $\mathcal{A}OE_0$.

Theorem 14.3.4 $\mathcal{A}O\,\mathrm{E}_0$ *proves* $(\exists x)\Box G(x)$, *and hence* $\Box(\exists^E x)G(x)$ *(necessary actual existence of a godlike being) and also uniqueness of the godlike being.*

PROOF Clearly the theory proves $G(x) \to \Box G(x)$ and $\Diamond(\exists x)G(x)$, that is, $(\exists x)\Diamond G(x)$. From $G(x) \to \Box G(x)$, we get $\Diamond G(x) \to \Box G(x)$ and $(\exists x)\Diamond G(x) \to (\exists x)\Box G(x)$; thus we get $(\exists x)\Box G(x)$. This gives $\Box(\exists x)G(x)$ and hence $\Box(\exists^E x)G(x)$ (because $G(x)$ implies $E(x)$). Uniqueness uses the property I_x as earlier.

Define X**Ess**x as earlier and prove that I_x is the individual essence of x for each x, but change the definition of necessary existence to

$$\mathbf{NE}(x) \equiv (\forall Y)(Y\mathbf{Ess}x \to \Box(\exists^E y)Y(y)).$$

Theorem 14.3.5 $\mathcal{A}OE_0$ *proves* **NE**$(x) = \Box E(x)$ *and also* P(**NE**).

PROOF We know that each essence of x is necessarily equivalent to I_x; thus **NE**$(x) \equiv \Box(\exists^E u)I_x(u) = \Box E(x)$. Furthermore, for each x, $G(x) \to \Box G(x) \to \Box E(x)$, and thus $\Box(G \subseteq \Box E)$ and hence $P(\Box E)$.

In the rest of this section, we give an alternative presentation of the theory $\mathcal{A}OE_0$ in the logic with variable domains. (This part may be omitted by a hurrying reader.) We now will call the theory $\mathcal{A}OE_0$, $\mathcal{A}OE_{01}$ and its predicate of godlikeness G_1. Now define $X \subseteq^E Y$ to be $(\forall^E x)$

$(X(x) \to Y(x))$, equivalently $(\forall x)((E(x) \wedge X(x)) \to Y(x))$ (see Fitting, 2002). Take $\mathcal{A}OE_0$ and change the axiom (A2) to (A2E), saying $(P(X) \wedge \Box(X \subseteq^E Y)) \to P(Y)$. Keep the original definition of godlikeness; that is, let $G_2(x) \equiv (\forall Y)(P(Y) \equiv \Box Y(x))$; the axiom (A3) becomes $P(G_2)$. Call this theory $\mathcal{A}OE_{02}$.

Theorem 14.3.6 *$\mathcal{A}OE_{02}$ proves $(\exists x)\Box(G_2(x) \wedge E(x))$ and hence $\Box(\exists^E x)\, G_2(x)$. It also proves $(\exists! x)G_2(x)$.*

PROOF *Claim 1:* The theory proves $P(X) \to \Diamond(\exists^E u)X(u)$. Indeed, $(\Box(\forall^E u)\neg X(u)) \to (\Box(X \subseteq^E \neg X)) \to (P(X) \to P(\neg X)) \to \neg P(X)$. Consequently, the theory proves $\Diamond(\exists^E x)G_2(x)$ and hence $(\exists x)\Diamond(E(x) \wedge G_2(x))$.

Claim 2: The theory proves $\Diamond(E(x) \wedge G_2(x)) \to \Box(E(x) \wedge G_2(x))$. Indeed, $\vdash \Box(G_2 \subseteq^E (G_2 \wedge E))$, which gives $\vdash P(G_2 \wedge E)$ and hence $\vdash (G_2(u) \wedge E(u)) \to \Box(G_2(u) \wedge E(u))$. From this, we get $\vdash \Diamond(G_2(u) \wedge E(u)) \to \Box(G_2(u) \wedge E(u))$. The rest is evident. Uniqueness is proved as usual.

Theorem 14.3.7 *$\mathcal{A}OE_{01}$ proves AOE_{02}, and $G_1(x) \equiv G_2(x)$.*

PROOF $X \subseteq^E (X \wedge E)$, thus $\vdash P(X) \to P(X \wedge E)$ and $(P(X) \wedge \Box(X \subseteq^E Y)) \to P(Y)$. We see that *$AOE_{01}$ proves* (A2E). Clearly $\vdash \Box(G_1(x) \to G_2(x))$, thus $\vdash P(G_2)$. Furthermore, $\vdash (\exists! x)G_1(x)$ and, also, $\vdash (\exists! x)G_2(x)$ (because all axioms of $\mathcal{A}OE_{02}$ are provable). This gives $\vdash \Box(\forall x)((G_1(x) \wedge G_2(y)) \to x = y)$.

Theorem 14.3.8 *$\mathcal{A}OE_{02}$ proves $\mathcal{A}OE_{01}$ and $G_1(x) \equiv G_2(x)$.*

PROOF Over $\mathcal{A}OE_{02}$, $\vdash (P(X) \wedge \Box(X \subseteq Y)) \to (P(X) \wedge \Box(X \subseteq^E Y)) \to P(Y)$, thus $\vdash$(A2). Moreover, $\vdash P(G_2 \wedge E)$, thus $P(G_1)$. The rest is as in the previous proof.

14.4 Oppy's Criticism

First, let us quickly mention a criticism from Oppy (1996). The idea was to define a god*-like being to be a being necessarily satisfying many but not all positive properties (and no nonpositive properties) and prove its necessary existence. Without going into details, let us prove a theorem showing that this is not possible.

Theorem 14.4.1 *Let $\mathcal{A}O_0$ be the theory over logic with fixed domain, as earlier. Let $\mathcal{A}O'_0$ be an Oppy-style extension of AO_0 with a new predicate G^* for god*-like such that T' proves the following:*

(B1) $G^*(x) \to (\forall Y)(\Box Y(x) \to P(Y))$,
(B2) $G^*(x) \to \Box G^*(x)$,
(B3) $(\exists x)G^*(x)$,
(B4) $(G^*(x) \wedge G^*(y)) \to x = y$.

Then $\mathcal{A}O'_0$ proves $G^*(x) \to G(x)$. In wards, assume that $\mathcal{A}O'_0$ proves that god-likeness of x implies that each property that x necessarily has is positive; god*-likeness implies necessary god*-likeness; and there is exactly one god*-like being. Then $\mathcal{A}O'_0$ proves that the unique god*-like being is the same as the unique godlike being.*

PROOF First, observe that (B2) and (B1) imply that G^* is positive. Second, introduce a constant g^* for the unique god*-like being (and g for the unique godlike being). Owing to uniqueness, for each property Y, $Y(g^*)$ is equivalent to $(\forall x)(G^*(x) \to Y(x))$. Let Y be an arbitrary positive property; we claim $\Box Y(g^*)$ (proving this we shall be ready). Indeed,

$\Diamond\neg Y(g^*) \to (\forall x)(G^*(x) \to \Diamond\neg Y(x))$ (see earlier), hence
$\Diamond\neg Y(g^*) \to \Box(\forall x)(G^*(x) \to \Diamond\neg Y(x))$ (by necessitation, $\Diamond$ being equivalent to $\Box\Diamond$),

thus $\Diamond\neg Y(g^*) \to P(\Diamond\neg Y)$ (because $\Diamond\neg Y$ necessarily follows from G^*, as we have just proved). But $\Diamond\neg Y$ is the same as $\neg\Box Y$, and $\Box Y$ is positive, as proved earlier. Thus we get a contradiction with the axiom (Al). We have proved $\Box Y(g^*)$. Thus g^* satisfies the formula defining the godlike being, and hence $g^* = g$. This completes the proof.

This is practically the same reasoning as employed by Gettings (1999). Oppy (2000) also suggests another "Gaunilian parody" of the proof of a godlike being. We slightly simplify (owing to the fact that we need fewer axioms than Anderson's original system), and we systematically speak of theories, not only of their models. Let P^* be a new predicate of properties (possibly definable in $\mathcal{A}O_0$, or just extending its language); we postulate the following (still over the logic with fixed domains):

(Al*) $P^*(X) \to \neg P^*(\neg X)$, where $(\neg X)(u) \equiv \neg X(u)$,

(A2*) $(P^*(X) \wedge (X \subseteq Y) \to P^*(Y))$,

(A3*) $P^*(G^*)$, where $G^*(u) \equiv (\forall Y)(P^*(Y) \equiv \Box Y(u))$.

These are the same axioms as earlier, only with P^*, G^* instead of P, G. Clearly the extended system proves $\Box(\exists u)G^*(u)$ (plus uniqueness), but assuming that P is not equivalent to P^*, we clearly get that the unique godlike being is different from the unique god*-like being. Can we assume this? Is it consistent? The answer is yes, and using the logic of a fixed universe, we can define such systems under the assumption that there are at least two objects. Call a predicate P^* on properties definable in AO_0 *Gaunilian* if $\mathcal{A}O_0$ proves (A1*)–(A3*).

Theorem 14.4.2 *Gaunilian predicates are in one-one correspondence with objects: for each Gaunilian predicate P^*, let g^* denote the corresponding god*-like being, and for each object constant a, let $P^a(Y)$ be defined as $\Box Y(a)$. Then P^a is a Gaunilian predicate, and its god^a like being is a. From $a \neq b$, it follows that $G^a(a) \wedge \neg G^b(a)$ (different elements define nonequivalent Gaunilian predicates).*

PROOF For P^a, (A1*) and (A2*) are evident. $G^a(u)$ becomes $(\forall Y)(\Box Y(a) \equiv \Box Y(u))$, which implies $\Box G^a(a)$. Thus $P^a(G^a)$, and we have (A3*). The rest is evident. In particular, if $a \neq b$, then $\Box I^a(a) \wedge I^b(a)$, hence $\neg G^b(a)$.

Thus, in each model, each object determines a system of Gaunilian properties whose god*-like object is this object; systems of Gaunilian properties are in one-one correspondence with objects of the model.

Now let us see what happens when we turn to the logic with variable domains of actually existing objects. Having a Gaunilian predicate P^* with the modified definition of G^* (i.e., $G^*(u) \equiv (E(x) \wedge (\forall Y)(P^*(Y) \equiv \Box Y(u)))$), we prove $\Box(\exists^E x)G^*(x)$, that is to say, that the unique object g^*

satisfying G^* satisfies $\Box E(g^*)$. (The theory may be called $\mathcal{A}OE'_0$.) Then we get the following.

Theorem 14.4.3 *In* $\mathcal{A}OE'_0$*, Gaunilian predicates are in one-one correspondence with necessarily actually-existing objects; one of them is the* (*starting*) *predicate P ofpositivity and the corresponding godlike object.*

PROOF By obvious modification of the preceding proof, we can arrive at this one. In particular, $G^a(u)$ is now equivalent to $E(u) \wedge (\forall Y)(\Box Y(u) \equiv \Box Y(a))$, which gives $\neg E(a) \rightarrow \neg G^a(a)$ and $E(a) \rightarrow [E(a) \wedge (\forall Y)(\Box Y(a) \equiv \Box Y(a))] \rightarrow G^a(a)$.

But we cannot prove existence of objects different from the godlike object, combined with satisfying necessary actual existence from any assumption of (plain) existence of many objects. One easily constructs a model of our theory over the logic with variable domains in which there are as many objects as you want, but the godlike object is the only object necessarily actually existing; that is to say, the model satisfies $(\forall x)(\Box E(x) \rightarrow G(x))$. Let us give a trivial example:

	G	M	E
g	1	0	1
m	0	1	0
g	1	0	1
m	0	1	1

with *G* being godlike; *M* being me; *g* being godlike being; *m* being me; E being actual existence; and on upper world (heaven) and lower world (earth).

14.5 Miscellanea

We shall discuss possible variants of the underlying logic, further variants of the axioms on godlikeness, and the question of proving the existence of the devil. First, let us mention that some authors prefer to present Gödel's system as a system in second-order modal logic, whereas here we prefer to use two-sorted, first-order modal logic. This is unimportant because the second-order logic (and second-order modal logic) with its general Henkin-style (Henkin-Kripke) semantics is equivalent to the corresponding two-sorted, first-order logic (and there is no reason to insist on the standard semantics of second-order logic in analyzing Gödel's ontological proof).

However, what can be discussed are modifications of the underlying modal logic and of the comprehension schema. I argue that the proof is made by using the logic KD45, called the "logic of belief," which is just S5 without the axiom $\Box\varphi \rightarrow \varphi$ (Hájek, 1996). Then $\Diamond\varphi$ can be read "φ is admitted" and $\Box\varphi$ read "φ is believed." This reading seems to be of some interest. In the same paper, a schema of "cautious comprehension" is suggested, and it is shown that the system $(\mathcal{G}O)_{caut}$ (i.e., $\mathcal{G}O$ with only cautious comprehension) is faithfully interpretable in $\mathcal{A}O_0$ (with its full comprehension) (see also Hájek, 2002a).

It is even possible to assume only existence of properties used in the proof and its discussion (as G, I_x, NE) and closedness of properties under negation. One such example is the work of Szatkowski (2005), where the author offers twenty variants of modal logic (second order but with Henkin-Kripke semantics) and shows them being complete with respect to a natural axiomatization. In the appendix, he shows that in all of them, Gödel's ontological proof (of $\Box(\exists x)G(x)$) works.

Kovač (2003) restricts the comprehension schema to prevent the collapse of modalities but keeps it reasonably strong. In particular, it is assumed that the formula $\varphi(x)$ in the comprehension schema does not contain any closed subformula. Axioms on modalities are not uniquely fixed; rather any modal system making the proof work is admitted.

I argue for a further weakening of Anderson's (A1) and (A2), motivated by the observation that (A2) would force us, if we had a predicate "devillike," to infer that *being godlike or devillike* is a positive property (Hájek, 2002b). This may be considered counterintuitive; if so, then the suggestion is to replace (A1) and (A2) by the axiom (A12), saying $(P(X) \wedge \Box(X \subseteq Y)) \rightarrow \neg P(\neg Y)$. In words, if *X* is positive, and *X* necessarily entails *Y*, then the negation of *Y* is not positive. Also, this weakened system proves $\Box(\exists x)G(x)$ (and has no collapse of modalities). Furthermore, a version exists with varying domains.

Koons (2005) extensively discusses possible restrictions of the comprehension schema and modifies the original system, $\mathcal{G}O$, by replacing the axiom (G5) of Gödel (positivity of *NE*) by (G6) (he names Gödel's axioms (A1)–(A5), and thus he names his axiom (A6)). Koon's axiom is just $P(X) \rightarrow P(\Box X)$. He shows that the modified system proves $\Box(\exists x)G(x)$.

Both the papers by Kovač and by Koons pay much attention to philosophical aspects; moreover, Sobel (2006) further analyzes these aspects of Kovač' and Koons' papers. These aspects are not discussed here.

Possibilities of changing Gödelian proofs to a proof of existence of the devil were discussed by Cook (2003) and Pfeiffer (n.d.). Trivially, if you replace the predicate *P* by *N* (negative), and *G* by *D* (devillike), for instance, in $\mathcal{A}O$, you prove $\Box(\exists x)D(x)$. But you cannot have both the original $\mathcal{A}O$ and its devillike modification and assume $P(X) \rightarrow \neg N(X)$ (that no property is both positive and negative); this is obviously an inconsistent theory. What you can aruge is the variant with (A12) just described earlier, in addition to its devillike variant, and $P(X) \rightarrow \neg N(X)$, but this theory is rather weak. See the papers just cited for more information.

14.6 What Do We Learn about Gödel?

Gödel is reported to have said to Oskar Morgenstern that he did not want to publish his proof, lest he be thought to actually believe in God, whereas he was only engaged in a logical investigation. But Wang reports that Gödel's wife, Adele, revealed that even if Gödel did not go to church, he was religious and read the Bible in bed every Sunday morning (Wang, 1987, 1996). In the so-called Grandjean questionnaire, Gödel stated that his belief was theistic, not pantheistic. From his letters to his mother, it is clear that he believed in a life after death. Wang formulated Gödel's argument as follows: "Science shows that the world is rationally arranged; but without a next life, the potentialities of each person and the preparations in this life make no sense." Concerning religion, Gödel wrote (Hao Wang's translation; German original in Schimanowich-

Galidescu, 2002), "I believe that there is much more reason in religion, though not in the churches, than one commonly believes, but we were brought up from early youth to a prejudgment against it. We are, of course, far from being able to confirm scientifically the theological world picture, but it might, I believe, already be possible to perceive by pure reason (without appealing to the faith in any religion) that the theological worldview is thoroughly compatible with all known data (including the conditions that prevail on our earth). The famous philosopher and mathematician Leibniz already tried to do this 250 years ago, and this is also what I tried in my previous letters."

I think that these and similar pieces of information on Gödel's opinions and beliefs show that the ontological proof was, for him, not a pure logical game. The proof helps us to better understand what sort of person he was. See also Köhler (2002).

14.7 Meaning for Religion?

Gaunilian examples discussed earlier were built from the assumption that there are some necessarily actually existing objects; they just show that there may be various systems of properties satisfying the axioms of the "ontological" theory (say, $\mathcal{A}O_0$). It should be explicitly stated that this does not destroy the ontological proof, proving, not assuming, necessary actual existence of a godlike being in the sense of the definition of godlikeness. But to make this notion of godlikeness really interesting from the religious (and philosophical) point of view, one should have an elaborated theory of positiveness (and of actual existence). A theologian should analyze the meaning of axioms (A1)–(A3) and the notion of positive properties, especially in the context of the Bible. Muck (1992) seems to be one of first attempts. Among other things, would the axiom (not used in Gödel's proof) "nobody except (possibly) God has necessary existence" be a theologically acceptable assumption? I believe that Gödel's proof would deserve an analysis similar to the one the famous protestant theologian K. Barth (1958, 622) presented for Anselm's proof. But it must be kept in mind that religious belief is not first a matter of accepting some axioms but accepting a kind of life (following an invitation). The religious notion of God is not the same as the philosophical one; recall Pascal's famous "Fire. God of Abraham, of Isaac, of Jacob, not of philosophers and scientists." On the other hand, let us quote another famous theologian (Tillich, 1955): "Is there any possibility of uniting ontology with biblical religion, if ontology could not accept the central assertion of biblical religion that Jesus is the Christ?… To ask the ontological question is a necessary task. *Against* Pascal I say: The God of Abraham, Isaac and Jacob, and the God of the philosophers is the same God. He is a person and the negation of himself as a person."

Finally, the famous Catholic theologian Hans Küng (1978, 626), in his long book, says (shortened):

> If God existed, then the substantiating reality itself would not ultimately be unsubstantiated. God would be the *primal reason* of all reality.
>
> If God existed, then the supporting reality itself would not ultimately be unsupported. God would be the *primal support* of all reality.
>
> If God existed, then evolving reality would not ultimately be without aim. God would be the *primal goal* of all reality.
>
> If God existed, then there would be no suspicion that reality, suspended between being and not being, might ultimately be void. God would be the *being itself* of all reality.
>
> *If God existed*,
>
> then, against all threats of fate and death, I could justifiably affirm the unity and identity of my human existence: for God would be the first ground also of my life;
>
> then, against all threats of emptiness and meaninglessness, I could justifiably affirm the truth and significance of my existence: for God would be the ultimate meaning also of my life;
>
> *If* God is, he is the answer to the radical uncertainty of reality.
>
> *That* God is, can however be accepted neither stringently in virtue of a proof or demonstration of pure reason nor absolutely in virtue of a moral postulate of practical reason, still less solely in virtue of the biblical testimony.
>
> *That* God is, can ultimately be accepted only in a confidence founded on reality itself.

14.8 Conclusion

Gödel's ontological proof is certainly not of central importance in his scientific achievement. However, "the proof and criticisms of it have inspired interesting work" (Fitting, 2002). In particular, in modal logic it gives a nontrivial exercise in modal axiomatic theories and their double semantics (fixed universe, varying universe). Furthermore, we have learned more about Gödel as a personality. Finally, concerning religion, "Even if one concludes that even this form of an ontological argument is no sufficient proof of God, it is a help for clarifying the notion of a property as used in the context of properties of God" (Muck, 1992; my translation). For all of the other incredibly Gödelian accomplishments, Gödel's ontological argument, while often overlooked, is one of his most fascinating legacies.

Acknowledgments

A version of this chapter was presented during the conference "Horizons of Truth," celebrating the one-hundredth anniversary of the birth of Kurt Gödel. The work of the author was partly supported by grant A100300503 of the Grant Agency of the Academy of Sciences of the Czech Republic and partly by the Institutional Research Plan AV0Z10300504.

References

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

Barth, K. (1958). *Fides Quaerens Intellectum*. Zollikon, Switzerland: Evangelischer.

Buldt, B., Köhler, E., Stöltzner, M., et al., eds. (2002). *Kurt Gödel: Wahrheit und Beweisbarkeit, Band 2: Kompendium zum Werk*. Wien: öbv & hpt.

Cook, R. T. (2003). God, the Devil and Gödel's other proof. In *Logica Yearbook 2003*, ed. L. Behounek, pp. 97–109. Prague: Filosofia.

Fitting, M. (2002). *Types, Tableaux and Gödel's God*. Dordrecht, Netherlands: Kluwer.

Fitting, M., and Mendelsohn, R. L. (1990). *First-Order Modal Logic*. Dordrecht, Netherlands: Kluwer.

Gettings, M. (1999). Gödel's ontological argument: A response to Oppy. *Analysis*, **59**, 309–13.

Gödel, K. (1970). Ontological proof. In *Collected Works*, vol. 3 (1995), pp. 403–4. [Introductory note by R. M. Adams, pp. 388–402. Appendix B: Texts relating to the ontological proof, including Gödel's first version, 1941, pp. 429–37.]

Hájek, P. (1996). Magari and others on Gödel's ontological proof. In *Logic and Algebra*, ed. A. Ursini and P. Agliano, pp. 125–35. New York: Marcel Dekker.

_______ (2002a). *Der Mathematiker und die Frage der Existenz Gottes*. In Buldt et al., 2002, 325–36.

_______ (2002b). A new small emendation of Gödel's ontological proof. *Studia Logica*, **71**, 149–64.

Hartshorne, C. (1962). *The Logic of Perfection and Other Essays in Neoclassical Metaphysics*. LaSalle, IL: Open Court.

Hopkins, J., and Richardson, H., eds. (1974). Proslogion by Anselm of Canterbury. In *Anselm of Cantebury, Vol. 1 (Monologion, Proslogion, Debate with Gaunilo and a Meditation on Human Redemption)*. Toronto: Edwin Mellen Press.

Köhler, E. (2002). *Gödels Platonismus*. In *Kurt Gödel: Wahrheit und Beweisbarkeit, Band 2: Kom-pendium zum Werk*. Ed. B. Buldt et al., pp. 341–86. Viennea: öbv & hpt.

Koons, R. C. (2005). Sobel on Gödel's ontological proof. http://www.scar.utoronto.ca/~sobel/OnL_T/Koons_SobelonGoedel.pdf.

Kovač, S. (2003). Some weakened Gödelian ontological arguments. *Journal of Philosophical Logic*, **32**, 565–88.

Küng, H. (1978). *Existiert Gott?* Munich, Germany: Piper.

Magari, R. (1988). Logica e teofilia. *Notizie di Logica*, 7.

Muck, O. (1992). Religiöser Glaube und Gödels ontologischer Gottesbeweis. *Theologie und Philosophie*, **67**, 263–67.

Oppy, G. (1996). Gödelian ontological arguments. *Analysis*, **56**, 226–30. [See also http://plato.stanford.edu/entries/ontological-arguments, Sect. 6.]

_______ (2000). Response to Gettings. *Analysis*, **60**, 363–67.

Pfeiffer, H. (n.d.). Die Existenz Gottes und des Teufels. Unpublished manuscript.

Schimanowich, W. de P. (2002). *Gödels Briefe an seine Mutter*. In *Kurt Gödel: Wahrheit und Beweis-barkeit, Bandl, Dokumente und historische Analysen*, ed. Köhler et al., pp. 185–208. Vienna: öbv & hpt.

Sobel, J. H. (1987). Gödel's ontological proof. In *On Being and Saying: Essays in Honor of Richard Cartwright*, ed. J. J. Thompson, pp. 24–61. Cambridge, MA: MIT Press.

_______ (2004). *Logic and Theism: Arguments for and against Beliefs in God*. Cambridge: Cambridge University Press.

_______ (2006). On Gödel's ontological proof. In *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*, ed. H. Lagerund et al., pp. 397–421. Uppsala, Sweden: Uppsala Philosophical Studies.

Szatkowski, M. (2005). Semantic analysis of some variants of Anderson-like ontological proofs. *Studia Logica*, **79**, 317–55.

Tillich, P. (1955). *Biblical Religion and the Search for Ultimate Reality*. Chicago: University of Chicago Press.

Wang, H. (1987). *Reflections on Kurt Gödel*. Cambridge, MA: MIT Press.

_______ (1996). *A Logical Journey: From Gödel to Philosophy*. Cambridge, MA: MIT Press.