Reflections on Gödel's Ontological Argument

Christopher G. Small[†]
University of Waterloo

[†] Mailing address: Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2G 3G1, Canada. E-mail:cgsmall@uwaterloo.ca.

1. Introduction. Kurt Gödel is best known to mathematicians and logicians for his celebrated incompleteness theorems which demonstrate the limitations of recursive axiomatizations in providing a complete account of mathematical truth for standard arithmetic. Physicists also know his famous relativistic model for a rotating universe in which time-like lines close back on themselves so that the distant past and the distant future are one and the same.[1] What is less well known is the fact that Gödel sketched a revised version of Anselm's traditional ontological argument for the existence of God. Gödel's contributions to the ontological argument have been slower to percolate through the literature than many of his other works. This is undoubtedly due to the fact that his work on the ontological argument was not published directly. There is evidence that he was unsure of the reception of such overtly theological work if it were published.[2] However, in 1970, Gödel showed his argument to Dana Scott, apparently to ensure that his ideas on the subject would not be lost. It is through the notes in Gödel's own hand and the notes of Dana Scott[3] that we have a primary source for this material.

What were the philosophical issues that motivated Gödel to revisit the ontological argument for God? Like his friend and colleague Albert Einstein, Gödel was philosophically inclined. Unlike Einstein, the major philosophical influence on Gödel was Gottfried Leibniz rather than Spinoza, according to Hao Wang.[4] Gödel described[5] his own philosophy as "rationalistic, idealistic, optimistic and theological," words that could well apply to Leibniz himself. His philosophy also involved "a monadology with a central monad [God]. It is like the monadology of Leibniz in its general structure."[6]. Leibniz himself devoted considerable attention to the traditional arguments for God's existence, and particularly to the ontological argument. As we shall see, Leibniz and Gödel concurred in diagnosing the main problems with Anselm's original argument. Another aspect of Gödel's work that can be traced back to Leibniz is in proof theory. In a brief paper[7] published in 1933, Gödel modified the necessity operator of C. I. Lewis' modal logic to show how modal logic could be used to discuss issues of provability and consistency in logical systems. Modal logic, and its interpretation in terms of Kripke's

---

[1] Malament (1995) contains a basic description of Gödel's relativistic cosmology. The philosophical motivation for the model and its inspiration from Kant's philosophy of time can be found in Stein (1995).

[2] Adams (1995).

[3] Sobel (1987).

[4] For much of the information we have about Gödel, we are indebted to Hao Wang who interviewed him. The fruits of those discussions can be found in Wang (1987) and Wang (1996).

[5] As reported by Wang (1996), p. 8.

[6] Many of the basic ideas of Gödel's argument can also be found in Leibniz' *Monadology*. Leibniz' entire philosophical system was built around an atomism of substances, attributes and concepts. Leibniz called the atoms of substance *monads*. Although simple in substance, monads differ from each other in attributes. God is necessary substance rather than accidental (*Monadology* §38); God has no element of privation (§41); and God is a simple substance (monad) of compatible attributes having no privation (§2-5, §45, §47). The idea that privation of simple attributes prevents the consistent combination of attributes is common to both Leibniz' metaphysics and to Gödel's ontological argument. See Blumenfeld (1995) for a discussion of Leibniz' ideas on the ontological argument. One of my intentions here is to delineate the Leibnizian assumptions within Gödel's ontological argument.

[7] See Gödel (1933).

possible world semantics, is fundamental both to the modern versions of the ontological argument and to the foundational issues in mathematics that Gödel pioneered. But Kripke's possible world semantics in turn has its roots in the philosophical work of Leibniz who speculated that God has actualized the best of all possible worlds. Other than Leibniz, the other main philosophical influences on Gödel were Immanuel Kant and Edmund Husserl.[8] In particular, it was the speculation of Kant on the nature of time which led Gödel to propose his rotating universe in which the ordinary rules of cause and effect appear to break down. It was quickly recognized that the rotating universe could not be physically realistic. However, Gödel's intention was to show that Kant's ideas on time were compatible with contemporary general relativity theory.

The ontological argument that we shall consider in the following pages will be a modified version of Gödel's original argument. Sobel[9] has argued quite forcefully that Gödel's argument leads to unacceptable consequences such as a modal collapse. Anderson[10] proposed an emended form of the argument to avoid the problems found by Sobel. However, many of these emendations may be unnecessary if the concept of attribution in Gödel's argument is treated with sufficient care. Within the following pages, I would like to explain this ontological argument, and provide some context that will help us assess its strengths and weaknesses. Readers who are not familiar with modal ontological arguments will find enough here, I hope, to provide an introduction. On the other hand, I also hope to provide some food for thought for those who are already familiar with both Hartshorne's and Gödel's versions of the ontological argument. In the next section, I will briefly review some tools from modal logic which are pivotal, both for the modern modal version of Anselm's argument due to Hartshorne,[11] and for Gödel's version.

2. Some Tools from Modal Logic. Gödel's ontological argument uses the modal logic S_5 which we shall describe briefly in this section. Let φ and ψ be propositions. We shall denote the negation of φ by $\neg\varphi$. The weak disjunction of φ and ψ — the statement "φ or ψ" — shall be denoted by $\varphi \vee \psi$, and the conjunction — the statement "φ and ψ" — by $\varphi \wedge \psi$. The formula $\varphi \rightarrow \psi$ shall denote the statement "φ implies ψ," while $\varphi \leftrightarrow \psi$ shall be interpreted as "φ is equivalent to ψ." Collectively, these operators, when applied to a collection of propositional variables such as φ and ψ with the usual truth table assignments, make up standard propositional logic.

The weakness of propositional logic in formulating counterfactual arguments is one of the main reasons for modal propositional logic. As is well known, the statement $\varphi \rightarrow \psi$, that φ implies ψ, is formally equivalent to the statement $(\neg\varphi) \vee \psi$. Thus a false statement can be said to imply any statement at all, regardless of its truth value. So in propositional

---

[8] For the influence of Kant, see Stein (1995). The influence of Husserl is described in Wang (1996), p. 164-172.

[9] See Sobel (1987). The final word on these objections is not yet in. However, much depends on how narrowly or widely the concept of an attribute is defined.

[10] See Anderson (1990). While Anderson's emendations address Sobel's objections, they also lead to difficult philosophical problems. I will discuss some of these in later notes.

[11] See Hartshorne (1962).

logic, the statement "If the Rome had not fallen, then computers would be using Roman numerals today" is in a certain sense true if truth values are assigned naively, because the antecedent is false.

In ordinary discourse, we say that φ implies ψ if it is not possible for φ to hold true without ψ being true as well. To capture this idea, modal logic introduces two new operators to propositional logic. These modal operators, denoted by $\Box$ and $\Diamond$, capture the ideas of necessity (as opposed to contingent or accidental) truth and possible truth respectively. For example, if φ is the statement that there exists a prime number between n and $2n$ for all positive integers n, and ψ is the statement that a tree grows in Brooklyn, then $\Box\varphi$ is true because the statement is provable, and therefore necessarily true. However $\Box\psi$ is false. That a tree grows in Brooklyn is a contingent or accidental truth, best formalized as $\psi \wedge (\neg\Box\psi)$. Formally, the statement $\Diamond\varphi$ can be defined as $\neg\Box\neg\varphi$, namely that it not necessarily true that φ be false. It is an immediate consequence of this definition that we can write $\Box\varphi$ equivalently as $\neg\Diamond\neg\varphi$.

In order to capture the full sense of necessity and possibility in ordinary discourse, we must interpret the necessity operator $\Box$ in a variety of ways that are appropriate to the nature of the discourse. Alvin Plantinga has noted that we should distinguish *natural necessity* on one hand and *logical necessity* on the other.[12] For example, the statement "Voltaire once swam the Atlantic" is possible in the sense of strict logical necessity, but impossible in the natural sense. Between natural and logical necessity undoubtedly lie many other types of necessity. For example, we can say that physical laws such as the conservation of angular momentum have a status that is less than logical necessity, because we can imagine a world where these laws do not hold. However, the law of conservation of angular momentum seems to be more universally true beyond the proposition that Voltaire did not manage to swim the Atlantic. (This is not to say that it is more credible.) A possible world in which Voltaire might have swum the Atlantic would still presumably have a law of angular momentum as an empirical truth. As modal ontological arguments attempt to prove that the existence of God is a necessary truth, beyond being a contingency, it is clear that the appropriate interpretation for the modal operators for the ontological argument is closer to logical necessity than natural necessity. However, as Gödel and other mathematical Platonists have argued, logical necessity need not be equated with provability. Nor should we presume that mathematical and logical truth encompass all necessary truths. There may well be many others. Plato felt that necessary truths could be found in aesthetics and ethics, also.

It is doubtful whether we can restrict ourselves merely to the two modal concepts of natural and logical necessity (or possibility). Language and the world abound in various restrictions on possibility. Nor can science ignore modal logic and relegate it to the realms of metaphysical speculation. For example, in order to understand causal relationships — to ask how one event can be a cause of another — it would appear that we have to work with counterfactuals in science. If I say that a certain drug causes a certain response in a

[12] See Plantinga (1974) p. 2.

subject, then I would appear to be saying that if the drug is administered then the response will follow, whereas if the drug is not administered the response will not follow. One of these two statements is a counterfactual argumentation.

What are the axioms or postulates of modal logic? In addition to the usual postulates of propositional logic, which can be verified by a truth table, modal logic also requires axioms of its own. We have already noted that

(M1) $\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$ for all φ (equivalently, $\Box\varphi \leftrightarrow \neg\Diamond\neg\varphi$ for all φ)

is an axiom. Another axiom that captures the some of the idea of necessity is that

(M2) $\Box\varphi \rightarrow \varphi$ for all φ (equivalently, $\varphi \rightarrow \Diamond\varphi$ for all φ.)

A commonly used axiom is

(M3) $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

which is called *modal modus ponens*. Also useful is the necessitation postulate, that

(M4) $\Box\varphi$ is true whenever φ is provable

i.e., whenever φ is itself an axiom or a theorem. These postulates represent the common core of modal logic. Two additional postulates are often added to these to give it additional strength. The first of these is

(M5) $\Box\varphi \rightarrow \Box\Box\varphi$ for all φ

and the second is that

(M6) $\Diamond\varphi \rightarrow \Box\Diamond\varphi$ for all φ .

Together, these two postulates state that the modal status of a proposition is a necessary truth. The principle that the modal status of a proposition is a necessary truth is called *Becker's postulate*. The system with the axioms of propositional logic and (M1) - (M5) is called the modal logic S_4. The axioms (M1) - (M6) together with propositional logic make up the modal logic S_5.

The modal logic S_5 is a suitable domain in which to explain the modern form of Anselm’s argument, such as the one by Charles Hartshorne, for instance. However, for Gödel’s argument, a modalised predicate logic is required. Nor will first order predicate logic be sufficient for the purpose, as we will need to have a second order logic with predicate variables as well. We shall let symbols such as x or y denote variables. A formula such as Fx shall denote that a variable x has attribute or predicate F. We shall write the statement that there exists an individual with property F by $(\exists x)Fx$, and the statement that all individuals have property F by $(\forall x)Fx$. More generally, if $f(x)$ is a formula involving the variable x, then $(\exists x)f(x)$ and $(\forall x)f(x)$ shall denote the corresponding existential and universal statements, respectively. We shall also need to quantify over predicates. An example of such quantification is a formula such as $(\exists F)(Fx \wedge Fy)$ which asserts that x and y have some attribute in common. (Henceforth, we treat terms such as *predicate* and *attribute* as synonyms. There is a small difference in the fact that a predicate is a syntactic object within a logical system, whereas an attribute is the semantic interpretation of the predicate. Where one word is chosen over another, it is with this distinction in mind.) This second order predicate logic can be extended to the a logic appropriate for Gödel's argument by introducing necessity and possibility operators which satisfy the axioms of S_5. So we will allow sentences of the form

$$\Box(\forall x)(\forall y)(\exists F)(Fx \wedge Fy)$$

which asserts that it is necessarily the case that any two individuals have some attribute in common. Beyond the second order, we shall need to consider various collections of predicates, which can be regarded as attributes of predicates.[13]

3. Hartshorne's Ontological Argument. In this section, we shall consider Hartshorne's version[14] of Anselm's ontological argument. It would be fair to say that the ontological argument has had a rough reception over the centuries. Unlike the cosmological and teleological arguments, which have had at least a hearing from sceptics, the ontological argument has often been dismissed as a piece of sophistry. It is true that at first glance one does feel that one has been "taken for a ride" by the argument. Nevertheless, the argument deserves a better reception than it has often had. Hartshorne's modal version of the argument provides a more solid foundation for a discussion of its strengths and weaknesses than does Descartes' version, which has been given much attention, and wittily summarized by C. S. Lewis as

> D is for Descartes who said "God couldn't be
> So complete if he weren't. So he is. Q.E.D."[15]

To judge the ontological argument, it is best to return to Anselm's own words. In the discussion which follows, I shall only consider those parts of Anselm's argument which are relevant to the construction of Hartshorne's proof.[16] Within his introductory passages

---

[13] Gödel's original proof was developed within second order modal logic, despite the fact that it contains higher-order objects such as the positivity operator. After careful consideration I have chosen to present his proof using a third order predicate algebra built on second order logic. I have made this choice because the argument behind the positivity of being God-like (section 8) is that the conjunction of positive predicates is consistent and positive. Stating this clearly requires the kind of predicate algebra presented in the paper.

[14] See Hartshorne (1962).

[15] From *The Oxford Magazine*, November 30, 1933.

[16] It was Hartshorne's insight that Descartes' ontological argument did not bring out the full flavor of Anselm's reasoning. Moreover, the more problematic parts of Anselm's reasoning, so incisively criticized

of the *Proslogion*, Anselm[17] asks

> By what signs, by what forms, shall I seek you [i.e., God]? (27)

I take this question to be an inquiry into the essence of God. Anselm answers his own question with his famous definition

> We believe you [i.e., God] are that thing than which nothing greater can be thought. (160)

The expression "can be thought" is apparently not meant in the psychological sense, but as a limitation on possibility: God is that for which it is impossible that there be a greater being in any respect. This establishes the ontological argument in the domain of modal logic. It also points the way to Anselm's basic idea, namely that it is not possible for God to possess any attribute which would impose a limitation on greatness or perfection. Anselm continues by arguing

> But when the fool[18] hears me use the phrase "something than which nothing greater can be thought" he understands what he hears ... . (163)
> So the fool has to agree that the concept of something than which nothing greater can be thought exists in his understanding. (173)

By the "understanding" Anselm is apparently referring to the modal realm of possibility.[19] That which exists in the understanding is that which is possible. So Anselm is asserting that even the atheist would agree that it is possible that God exists. Even though the atheist asserts that God's existence is false, the atheist surely would agree that this is a contingent falsehood. Like a painter who imagines a picture before painting it, the atheist can conceive of a world in which God exists even if that world is not the true world. This leads us to the first axiom, namely

(H1) $\Diamond g$

---

by Kant, are actually unnecessarily in the modal logic S_5.

[17] The translation of the *Proslogion* that I shall use is that of Benedicta Ward, which can be found in the popular Penguin Classics edition *The Prayers and Meditations of Saint Anselm with the Proslogion*, published in 1973. The numbering of verses follows that edition.

[18] The word "fool" is not to be understood in its modern sense. Anselm means an atheist by this term. This usage derives from the Bible, in which the Hebrew word translated as "fool" denotes a morally deficient person, rather than someone who is stupid in the intellectual sense.

[19] At least, this is Hartshorne's conclusion. It has been challenged by John Hick, who has argued that the interpretation of Anselm's argument through modern modal logic has some inconsistencies. Hick's argument is that Hartshorne misreads Anselm by confusing ontological necessity – that a statement is true from the perspective of eternity – with logical necessity. It is rather difficult to get into Anselm's mind on this point. However, we do not have to disagree with Hick completely on this issue in order to appreciate Hartshorne's arguments on its own merits beyond Anselm's intentions. However, the validity of Hartshorne's argument may well depend upon the appropriate interpretation of necessity. See Hick (1968).

where g is the statement that God exists. When we consider Gödel's ontological argument, we shall see that this statement will not be taken as axiomatic. Indeed, Gödel's contribution to the ontological argument is partly a deeper analysis of this statement, which is axiomatic here. We have to say that Anselm does provide an argument for the possibility of God's existence. However, his argument is not convincing, because he does not distinguish between propositions which are *conceivable*, and propositions which are *possible*. The distinction is very important. For example, I can conceive of an even integer greater than two which cannot be written as the sum of two prime numbers. At present, we do not know whether such integers exist.[20] Similarly, the fact that I can conceive of God's possible existence does not *ipso facto* imply that God might possibly exist.

> Something than which nothing greater can be thought so truly exists
> that it is not possible to think of it as not existing. (194)

With this statement, Anselm rules out the accidental existence of God. In modal logic this becomes

(H2) $\quad g \to \Box g.$

Anselm's argument is that if God existed only in a accidental sense, then we could imagine a greater being, namely one which whose existence is necessary. Since God is the greatest being of which we can conceive, it follows that God can exist only necessarily and not accidentally. Perhaps a more modern language would help in thinking about (H2). Anselm seems to be saying, in effect, that if God existed accidentally, and not necessarily, then we could ask embarrassing questions like "If God made the universe, then who made God?" A being invoked to explain the existence of the universe, but whose existence in turn is in need of explanation, is not great enough for Anselm to call God.[21]

While many sceptical people might be willing to accept (H1) and (H2), it turns out that within the modal logic S_5, the acceptance of both statements is inconsistent with atheism. To prove this, we need only note the following. First, from Becker's postulate (M6), we get

(H3) $\quad (\Diamond \neg g) \to (\Box \Diamond \neg g).$

But the law of the excluded middle tells us that $(\Box g) \vee (\neg \Box g)$. This is equivalent to

[20] This is Goldbach's conjecture, a famous unsolved problem in arithmetic.
[21] This is the place where Hartshorne's argument most closely parallels to cosmological argument for God's existence. In the cosmological argument as developed in one version by Leibniz, the *principle of sufficient reason* demands that all accidental facts must have a reason or cause. The universe, being in existence as a contingent fact, must have an explanation, namely God. God becomes the ultimate grounding of the principle of sufficient reason. If God were accidental, then this would contradict the principle of sufficient reason. Therefore God's existence is necessary. Note that Leibniz' version of the cosmological argument implies (H2). However, (H2) does not require the principle of sufficient reason or Leibniz' cosmological argument.

(H4) $(\Box g) \vee (\Diamond \neg g)$.

So using the substitution rule on (H3) and (H4), we can conclude that

(H5) $(\Box g) \vee (\Box \Diamond \neg g)$.

However, taking the contrapositive of (H2), and writing $\neg \Box g$ as $\Diamond \neg g$, we obtain

(H6) $(\Diamond \neg g) \rightarrow (\neg g)$.

To goal here is to clean out the cupboards by showing that accidental types of existence are impossible for God. Since we have derived (H6) by the postulates (H1) and (H2) for the statement g, using the rules of deduction for modal logic, (H6) is a theorem. So we can apply the necessitation postulate (M4) to (H6). This gives us

(H7) $\Box[(\Diamond \neg g) \rightarrow (\neg g)]$.

Now applying modal modus ponens (M3) to (H7), we find that

(H8) $(\Box \Diamond \neg g) \rightarrow (\Box \neg g)$.

Now (H8) allows us to apply the substitution rule to (H5). This yields

(H9) $(\Box g) \vee (\Box \neg g)$.

Statement (H9) has cleaned out the cupboards. Among the four modes of existence — necessary existence, accidental existence, accidental nonexistence and necessary nonexistence — we have ruled out the accidental cases. It remains to bring in (H1) to finish it off. To do this, we write (H1) as

(H10) $\neg \Box \neg g$.

Putting (H9) and (H10) together finally gives us

(H11) $\Box g$.

Thus God exists necessarily. Q.E.D.

But is it a proof? If a proof is something that starts with certain statements as axioms and combines them together with clearly defined rules of deduction, then it is certainly a proof. However, as we normally understand a logical proof, the axioms and rules of deduction have to be *self-evident*. Statements (H1) and (H2) are certainly not self-evident. However, that does not make them false. Objections of various kinds have been proposed. Critical scrutiny has been brought to bear on Becker's postulate, (M5) and

(M6), and Anselm's statements (H2) and (H1).

Some people have had difficulty accepting Becker's postulate. However, when we interpret the postulate in possible world semantics, we find that leads to nothing that is particularly remarkable or unacceptable for ordinary discourse. Imagine that two individuals from different possible worlds were able to compare their understanding of the meaning of the operators $\Box$ and $\Diamond$. Then Becker's postulate is a statement of equivalence in their interpretations of modal possibility and necessity. Suppose, for example, that a Napoleon who won the battle of Waterloo was able to sit down at the table with a Napoleon who lost the battle of Waterloo to discuss how the battle went. If Becker's postulate is satisfied, then they should be able to agree about which contingencies are truly possible. This seems to be most closely in accord with our naive views of modal possibility. That is not to say that it is forced upon us.

Kant's criticism of the ontological argument is probably the most famous. Even today, many people consider it to have effectively demolished Anselm's argument. However, it must be said that Kant's criticism is more appropriate for Descartes' version of the ontological argument than Hartshorne's version. Kant's argument was that *existence is not a predicate*. That is, existence is not an attribute of individuals in the same way that being short or red is. It is certainly true that we have to be careful here. If we can arbitrarily add existence as a defining attribute for an individual, there seems to be no limit to what we can prove to exist. Suppose we allow existence to be one of the defining attributes of some being, such as a unicorn, for example. So a unicorn, by definition, is a beast resembling a horse, with a horn on its head, which exists. While many of us will not mind defining a unicorn as a horse-like animal with a horn on its head, we will hesitate to allow existence to be a defining attribute in the same sense. It would seem that allowing existence to be a predicate ensures that unicorns exist. Kant proposed that the ontological argument slips existence in the back door, so to speak, by ensuring that it is a consequence of the perfection of God.

Is existence a predicate? When we predicate something, we usually impose some restriction upon it. To say that unicorns are horse-like is apparently to say that the set of unicorns is a subset of the set of all horse-like things. However, what restriction is imposed by saying that something exists? Presumably there is a class of objects which includes actual and mythical things. To assert that unicorns exist is to assert that unicorns are to be found among the set of actual things. If this is the case, the problem may seem to be solved. However, many problems arise in talking about non-existent things. When we consider Gödel's argument later we shall see that to a certain extent he follows Frege in asserting that existence is not a first-order predicate such as being horse-like but a second-order concept about the exemplification of a predicate or attribute.[22]

---

[22] The brief discussion here cannot do justice to the full range of arguments about existence and predication. To attempt to cover the topic more fully would take us too far afield. Is existence a non-nuclear predicate? Is it a first-order predicate or of higher order? The reader is referred to Knuuttila and Hintikka (1986) for a more complete treatment of the subject.

Moreover, Kant's objection is only directly relevant to Descartes' version of the ontological argument and not Hartshorne's version. Whether or not existence is a predicate, the distinction between necessary existence and accidental existence does seem to make a valid distinction between two ways that things can exist. For example, we can say that a prime number between 100 and 105 exists necessarily, whereas the Statue of Liberty exists accidentally. These two modes of existence seem to describe attributes of the objects in question. A prime number between 100 and 105 is a Platonic object, while the Statue of Liberty — American sentiment notwithstanding — is not. Proposition (H2) makes a claim for the nature of God's ontological status that is similar to saying that numbers are Platonic objects. Suppose, for example we interpret g to be the proposition that there exists an even integer greater than 2 which cannot be written as the sum of two prime numbers. Although I do not know whether g is true when interpreted thus,[23] I can conclude that $g \to \Box g$ is correct. If g is false, then the proposition is vacuously true. If g is true, then the proposition is true because the truths of arithmetic are necessary truths. Hartshorne's argument makes a similar claim for God and there is no obvious reason to deny such an ontological status to God.

There remains only one proposition left to criticise, and that is (H1). From the discussion so far, the main reason we have for the truth of (H1) is that God's existence does seem to be conceivable. However, as was stated above, conceivable existence and possible existence are not the same. It is perfectly reasonable that we find the existence of God conceivable because of a limitation in our understanding of metaphysics. Without (H1), the best that we can prove is that $\Diamond g \to \Box g$, namely that if God possibly exists, then God must exist necessarily. But where do we go from here?

4. The Algebra of Predicates. A primary difficulty that we encounter in ordinary discourse is that we often assume that to *name* something is to *identify* it. This is clearly philosophically unsatisfactory, but it forms the basis for much ordinary discussion. The problem of identity is a vexing one precisely because to name someone is not to identify him. Uniqueness of characteristics or attributes is one of the ways we identify individuals: For a physicist, to inquire into the existence of a hitherto unknown subatomic particle is to inquire into the *exemplification of a given attribute or set of attributes*.[24] So if we define a property F of being a nearly massless electrically neutral particle of spin 1/2 that is not subject to the strong force (one of the four known fundamental forces of nature), then the statement that a neutrino exists can be rendered as $(\exists x)Fx$. In his ontological argument, Gödel took some care to approach his basic existential question in the same way. Rather than inquire into the existence of God, he inquired into the existence of a *God-like individual*. This sort of language may have insufficient piety for some. However, it is in keeping with his theory of essences. In the previous section, I argued that the weakest point in Hartshorne's ontological argument was to be found in proposition (H1), to wit, that God possibly exists. Translated into Gödel's framework, this becomes the statement that the attribute of "God-likeness" is possibly exemplified. To discuss the possible

[23] This is Goldbach's conjecture, a famous unsolved problem in arithmetic.
[24] I use the word "exemplfication" here to mean what others call "instantiation."

exemplification of attributes, we need to consider the entailment and consistency of attributes.

Let F and H be two predicates. We shall write $F \Rightarrow H$ and say that F *entails* H if

(C1) $\Box(\forall x)(Fx \rightarrow Hx)$

is a true statement. That is, an attribute F entails an attribute H if any individual with attribute F necessarily has attribute H. If F entails H, then F can be regarded as containing all aspects of H within itself, so to speak. Using this definition of entailment, we can define an *equivalence relation* on a collection $\mathcal{F}$ of predicates by defining

(C2) $F \Leftrightarrow H$ whenever $F \Rightarrow H$ and $H \Rightarrow F$.

Then the relation of entailment between predicates of $\mathcal{F}$ satisfies

(C3) $F \Leftrightarrow F$ (reflexivity)
(C4) $F \Rightarrow H$ and $H \Rightarrow F$ implies $F \Leftrightarrow H$ (antisymmetry)
(C5) $F \Rightarrow H$ and $H \Rightarrow K$ implies $F \Rightarrow K$ (transitivity)

Property (C3) is immediate from the definition of entailment, and property (C4) is really a definition in itself. Property (C5) follows easily from the rules of predicate logic and modal modus ponens. Thus the entailment relation is a *partial ordering* on the equivalence classes of $\mathcal{F}$. The partial ordering can be shown to be a *lattice* if we suppose that $\mathcal{F}$ is closed under the lattice operations $F \wedge H$ and $F \vee H$ where

(C6) $\Box(\forall x)(F \wedge H)x \leftrightarrow \Box(\forall x)(Fx \wedge Hx)$, and
(C7) $\Box(\forall x)(F \vee H)x \leftrightarrow \Box(\forall x)(Fx \vee Hx)$.

The predicate $F \wedge H$ entails both the predicate F and the predicate H. It is the minimal such predicate in the sense that if $K \Rightarrow F$ and $K \Rightarrow H$ then $K \Rightarrow (F \wedge H)$. Similarly, the predicate $F \vee H$ is entailed by both F and H and is the maximal such predicate in the sense that if $F \Rightarrow K$ and $H \Rightarrow K$ then $(F \vee H) \Rightarrow K$. We shall call $F \wedge H$ the *conjunction* of F and H, and $F \vee H$ the *disjunction* of F and H. We can easily extend the definitions of disjunction and conjunction to three or more predicates. However, we shall also have reason to consider the conjunction of possibly infinitely many predicates. Since we cannot write infinitely long sentences, we can use the maximality and minimality of disjunction and conjunction to define them in the infinite case. Suppose $\mathcal{F}$ is a collection of predicates. First, we shall write the statement that a predicate H entails every predicate in $\mathcal{F}$ as $H \Rightarrow \mathcal{F}$. It will also be convenient to let $\mathcal{F} \Rightarrow H$ denote that F is entailed by every member of the collection $\mathcal{F}$. That is, for all F in $\mathcal{F}$, we have $F \Rightarrow H$. We define the conjunction of $\mathcal{F}$ to be a predicate $\bigwedge \mathcal{F}$ with the property that

(C8) $\bigwedge F \Rightarrow F$,

which is minimal in the sense that

(C9) $(\forall H)[(H \Rightarrow F) \rightarrow (H \Rightarrow \bigwedge F)]$.

Similarly, the disjunction of F is a predicate $\bigvee \mathrm{F}$ with the property that

(C10) $F \Rightarrow \bigvee F$,

which is maximal in the sense that

(C11) $(\forall H)[(F \Rightarrow H) \rightarrow (\bigvee F \Rightarrow H)]$.

From these considerations, we see that there are two predicates, which we shall denote by Φ and Ω, which are maximal and minimal among all predicates, respectively. Note that if F and H are any two predicates, then the predicate $F \wedge (\neg F)$ entails H, because a contradiction entails all statements. Let Φ denote this contradictory predicate.[25] In addition, the predicate $F \vee (\neg F)$ is entailed by all predicates because a necessary truth is entailed by all statements. Let Ω denote this predicate.

A predicate F will be said to be *consistent* if $\neg(F \Rightarrow \Phi)$ is a theorem. A predicate F will be said to be *exemplified* if the statement $(\exists x)Fx$ is a theorem, and is said to be *possibly exemplified* if the statement

(C12) $\Diamond(\exists x)Fx$

holds. Intuitively, it is clear that if a predicate is possibly exemplified then it must be consistent. We would also expect the converse to be true. That is,

(C13) Theorem: $\Diamond(\exists x)Fx \leftrightarrow \neg(F \Rightarrow \Phi)$.

Proof: To prove this, we note that $(\neg Fx) \leftrightarrow (Fx \rightarrow \Phi x)$ is a theorem. Therefore, so is $\Box(\forall x)[(Fx \rightarrow \Phi x) \leftrightarrow (\neg Fx)]$. Pushing both the universal quantifier and the necessity operator across the equivalence, we get

(C14) $\Box(\forall x)(Fx \rightarrow \Phi x) \leftrightarrow \Box(\forall x)\neg Fx$,

which is equivalent to (C13). Q.E.D.

Statement (C13) is a basic result in Gödel's ontological argument. To prove the existence

[25] It is easily checked that any two contradictory predicates are equivalent.

of a God-like individual, the argument will first show that the attribute of being God-like is consistent, and can therefore be possibly exemplified.

5. Positivity and Privation. In the last section we considered an algebraic theory of predicates which imposed a partial ordering among predicates by means of entailment. In this section, we shall consider Gödel's positivity operator, which will allow an ordering of individuals by means of positive attributes. Gödel's argument defines an operator **Pos**. In much the way that predicates provide a truth-functional assignment when applied to individual constants and variables, so the positivity operator **Pos** provides a truth-functional assignment on predicates. We shall say that **Pos**(F) is true provided that the predicate F is, in fact, a positive attribute.

In ordinary language, we might be inclined to say that one thing is greater than another if the former has some positive attribute that the latter lacks. We might disagree with each other as to the ranking of things or objects according to their value, but we must inevitably make such judgements, whether we regard them as objective or not. In view of the ambiguity of such concepts, it is important to understand what Gödel meant by a positive attribute. In his own words he said that the operator **Pos** could be interpreted in a *moral-aesthetic sense*, or in the sense of *pure attribution*. The concept of a predicate being positive in a moral or aesthetic sense, provides no difficulty, at least initially. Clearly, if Fx means that x is beautiful, we would be willing to accept that F is positive in an aesthetic sense, even if we disagree in our judgements about beauty. If Fx means that x is virtuous, we might grant the same, even if we have no idea what virtue is. But what is meant by "pure attribution?" By "pure attribution," Gödel states that we are to understand that a predicate attributes some quality to an individual, and that the quality contains no element of "privation." Similar ideas are to be found in the writings of Leibniz and Spinoza and can be traced back at least as far as the work of the neoplatonic philosophers. For example, in the *Enneads*, Plotinus wrote

> An absence is neither a Quality nor a qualified entity; it is the negation of a Quality or of something else, as noiselessness is the negation of noise and so on. A lack is negative; Quality demands something positive.[26]

By a Quality in this passage, we are perhaps to understand a simple attribute that is pure in the sense of containing no negation or privation. In the neoplatonic order, the simple attributive aspects of being emanate from God; privation, which is the negation of an attribute or attributes, comes from the distance between God as the source of emanation and the thing itself. Evil is the absence of the Good, and is itself a type of privation. However, for the purposes of our argument here, we cannot assume the entire neoplatonic framework for interpreting such ideas, as it is the very existence of God which is the question at issue.[27] We can either restrict the interpretation of the operator **Pos** to the

---

[26] *Enneads* II.4. Here I use the translation of Stephen MacKenna, which is available through Penguin Classics.

[27] Nor should we assume that Gödel accepted Plotinus' view that God, in the form of the first hypostasis, is beyond description.

moral-aesthetic sense, or accept an ontology that is compatible with the concept of pure attribution. Certainly, Plato and later neoplatonists had no difficulty incorporating a moral-aesthetic definition within a view of being emanating from a source. Both the Good and the Beautiful were identified with the higher realms of being in Plotinus' vision.[28]

In turn, the neoplatonic influence on Leibniz should also be noted, in view of the influence of Leibniz on Gödel's thinking. For example, Leibniz' monadology has a hierarchy of monads.[29] In Leibniz' system, the monads, the basic units of substance, do not interact with each other. Instead, they are harmonised with each other, and each reflects, although incompletely, the totality of all other monads. This leads to a partial ordering of monads in which the immaterial component of each monad is proportional to the clarity with which it reflects all other monads and their attributes. The material component of each monad is represented by a privation: the failure of the monad to provide a perfect representation of the totality of all other monads. This is reminiscent of the neoplatonic principle of emanation. The concept of pure attribution in Gödel's metaphysics has its parallel in the affirmative simples of Leibniz, who also tried to support the ontological argument by arguing for the possibility of God's existence.[30]

We shall now turn to the axiomatic framework for the operator **Pos**. The first axiom asserts that positivity of a predicate is independent of the accidental structure of the world. That is,

(P1) $\quad \mathbf{Pos}(F) \rightarrow \Box\, \mathbf{Pos}(F).$

The similarity between (P1) and (H2) is striking. While (P1) looks innocuous, it does rule out any relativism in ethics, aesthetics or ontology. If moral laws were solely the result of social contracts they would presumably be contingent truths rather than necessary ones. If beauty were simply in the eye of the beholder, then (P1) would also be an unwarranted conclusion. The next axiom tells us that a simple attribute is positive if and only if the negation of that attribute is not. This can be written as

(P2) $\quad \mathbf{Pos}(F) \leftrightarrow \neg\, \mathbf{Pos}(\neg\, F).$

---

[28] See *Enneads* I.6 for Plotinus' much admired treatise on Beauty, and *Enneads* I.8 for his discussion of Evil as a form of privation, and the Good. Plotinus wrote "The Good is that on which all else depends, towards which all Existences aspire as to their source and their need, while itself is without need ... giving out from itself the Intellectual-Principle and Existence and Soul and Life and all Intellective-Act." His writing on Beauty contains the passage: "What is beyond the Intellectual-Principle we affirm to be the nature of Good radiating Beauty before it."

[29] *Monadology* §49-50.

[30] We shall discuss this at greater length later. It is interesting to speculate that Gödel's ontological argument was intended to provide the same kind of foundation for Leibniz' work within modal logic that Hartshorne provided for Anselm. Leibniz' contributions to the ontological argument have not been as influential as those of Kant, perhaps because his monadology has been difficult for many people to accept. Like Hartshorne, Leibniz recognised that the second argument from Anselm, from existence to necessary existence, was ontologically the more interesting. Leibniz, like Gödel but unlike Hartshorne, felt that the argument for God's possible existence could have a rigorous foundation.

This is Gödel's axiom. There is no doubt that it looks a little strange because it does not admit the existence of neutral attributes, those which are not positive and whose negations are not positive. Anderson has proposed a somewhat weaker proposition which allows for neutral attributes.[31] This criticism of (P2) is too hasty in my view. However, I shall defer consideration of these objections until later.

The next axiom is designed to capture the idea that the concept of positivity is "pure" in the sense that it contains no privation. Any predicate that is entailed by a positive predicate is itself positive. Therefore,

(P3) $[\mathbf{Pos}(F) \wedge (F \Rightarrow H)] \rightarrow \mathbf{Pos}(H)$.

Axiom (P3) needs careful consideration, because its consequences are difficult to determine. Some people have maintained that any proposition which is affirmative in its meaning must entail something negative. If this were true, then the class of predicates which are purely attributive in Gödel's sense might be void. Consider the apparently positive predicate F where Fx means that x was present at the Statue of Liberty in New York at noon on July 4, 2000. While F is affirmative in nature, it would be obvious that anyone who was present at the Statue of Liberty on that day at noon was not present at the Pyramid of Cheops, say, at the same time. The predicate would thereby seem to entail something which is negative in the same sense that the original predicate is affirmative. However, such an objection cannot be sustained. If the attribute of being present at a given location in space and time inherently means not being somewhere else, then the attribute is certainly not positive. However, the attribute need not entail any such negation. The fact that people cannot be in two places at once is a limitation on people and not on the attribute of being present in a place. Such a predicate can be positive provided it contains no privation and is compatible with the possibility of being present at other locations as well. This is certainly something than we would expect of God. Similarly, the predicate "knows the capitol city of North Dakota" can be regarded as affirmative in nature provided that it is compatible with all other forms of knowledge. If (P3) is true when **Pos** is interpreted in a moral sense as "good," then **Pos** must obviously mean "purely good" as opposed to "good overall on balance." So axioms such as (P3) appear to rule out any utilitarian interpretation of what is good, which would require a calculus of the greatest good for the greatest number.

The last axiom that we shall need states that positivity is preserved under conjunction. That is, if every attribute in a collection is positive, then the attribute which is the conjunction of all these attributes is also positive. We can write this as

(P4) If $\mathbf{Pos}(F)$ holds for all F in some collection $\mathcal{F}$, then $\mathbf{Pos}(\bigwedge \mathcal{F})$.

[31] See Anderson (1990).

Collectively, these axioms tell us that if we start with a collection of positive simple attributes and combine them using the rules of entailment, conjunction and weak disjunction, then the resulting family of attributes so obtained is also positive. This is the idea behind Leibniz' idea that positive attributes can obtained by combining simple affirmative attributes.

We shall conclude this section with Gödel's modal version of the result of Leibniz, namely that positive attributes are consistent, and thereby possibly exemplified.

(P5) <u>Theorem:</u> $\mathbf{Pos}(F) \rightarrow \Diamond(\exists x)Fx$.

<u>Proof:</u> To prove this result we need only note that if F is a positive attribute which is not possibly exemplified, then F is inconsistent by theorem (C13) above. So $F \Rightarrow \Phi$. That is, F entails the inconsistent predicate. Using (P3) we see that $\mathbf{Pos}(\Phi)$. But for any predicate H we have $\Phi \Rightarrow H$ and $\Phi \Rightarrow \neg H$. Applying (P3) again, we obtain $\mathbf{Pos}(H)$ as well as $\mathbf{Pos}(\neg H)$. But this contradicts axiom (P2). Q.E.D.

<u>6. Essences.</u> It will be convenient to let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, etc., denote the collection of all attributes that belong respectively to given individuals x, y, and z, etc. It seems natural to assume that each such collection is closed under the operations of entailment, conjunction and disjunction.[32] An implicit assumption in Gödel's system that the class of attributes of an individual x may be said to determine x. So by the *essence* of x, Gödel meant the conjunction $\bigwedge \mathcal{X}$ of all such attributes of x. It will also be convenient to denote the essence of an individual x, by the predicate X. Similarly $Y = \bigwedge \mathcal{Y}$, $Z = \bigwedge \mathcal{Z}$ etc. Of course, it is natural to ask whether an individual is determined by its essence in the sense that distinct individuals have distinct essences. In proposing his principle of the *identity of indiscernibles*, Leibniz apparently felt that something like this is true. This principle can be formalised as

(E1) $(X \Leftrightarrow Y) \rightarrow (x=y)$.

Some comments can be made about (E1). The first is that the essence of an individual x is uniquely defined up to equivalence of predicates. It is easily seen that if $F \Leftrightarrow H$ then H is an essence of x if and only if F is. So we shall speak of *the* essence of x without any real ambiguity. Secondly, a consequence of (E1) is that at most one individual can have

[32] I use the algebraic concept of closure here. The collection of attributes of an individual forms a lattice in its own right. Starting with the attributes of an individual x and using the operations of conjunction, disjunction or entailment, one produces attributes which are all attributes of x as well. In contrast to these operations, the negation of an attribute of an individual is never within the collection of attributes. A technical note here is that the extensional counterparts of the collection of attributes of an individual forms what topologists would call a fixed ultrafilter.

all positive attributes.[33] To prove this, suppose that X and Y both contain every positive attribute. Then by (P2), any attribute F which is which is not positive is such that $\neg F$ is in X. But F and $\neg F$ cannot both be attributes of x. Thus the attributes of x are precisely the positive attributes and no other. By similar reasoning, the same must be true of y. Property (E1) then implies that $x=y$.

Does every individual have an essence? Many existentialists might well say no, if that individual is a human being rather than a tree, say. This resolution of this issue is not required for the ontological argument. However, it would be somewhat strange if God were found to have an essence while humans do not. Gödel apparently believed that people have essences. In a letter dated August 14, 1961, he wrote

> Among all possible beings, "I" am precisely this combination of properties whose nature is such and such.[34]

The identity of indiscernibles, also known as Leibniz Law, is well known in quantum mechanics, where it is impossible to ask whether two subatomic particles are really the "same." If they have the same properties in the sense of being the same type of subatomic particle then the question ceases to be scientifically meaningful beyond that. If the principle of the identity of indiscernibles is false, then the door is left open to a kind of polytheism in the ontological argument which follows.

7. Necessary Existence. Gödel's definition of necessary existence follows from his characterization of individuals through their essential properties. An individual is said to have the property of *necessary existence* if it is necessary that the essence of that individual is exemplified.[35] In formal terms, this can be written as

(N1) $$NE(x) \leftrightarrow \Box(\exists y)Xy \; .$$

Gödel added the additional axiom

(N2) **Pos**(NE),

which is in the spirit of Anselm's argument. While (N2) looks innocuous, it makes a large jump in our interpretation of positivity. The concept of necessary existence introduced in (N1) is not an ordinary attribute that can be verified in each possible world. To know that John Smith has red hair in the actual world, one need merely check the state of affairs in this world. But to check whether John Smith necessarily exists, one has to check the

---

[33] This is a very important point for any ontological argument that makes a claim to be monotheistic.

[34] As quoted in Wang (1996), p. 106.

[35] This definition has some affinity to Frege's interpretation of existence as a second-order predicate. In the statement "Socrates is wise" the predicate "is wise" is first-order. On the other hand, if one says that "wisdom is rare" then the predicate "is rare" is second-order, predicating the concept of wisdom. In a similar way, the argument here is that existence is saying something about the exemplification of essences, here interpreted as a predicate.

situation in all other possible worlds as well.

If **Pos** is interpreted in a moral-aesthetic sense, then it is by no means clear that (N2) is true. If **Pos** denotes pure attribution then (N2) would seem to be reasonable. Among all properties that are generally affirmative in nature, existence, and particularly necessary existence, should be included as we understand it. Indeed, nonexistence is a form of privation almost by definition. This is not to say that the consequences of (N2) are unambiguous. In isolation (N2) seems perfectly reasonable. But it is hard to determine its compatibility with axioms that have been described earlier.

8. Completing Gödel's Ontological Argument. With these ideas in place, the basic building blocks of the argument are now available. Indeed, the proof is almost complete. The remaining steps follow fairly quickly from the definition of God that will now be introduced. Let $\mathcal{G}$ denote the collection of all positive properties. A *God-like* individual is defined as one that possesses the conjunction of all positive attributes. We define a predicate G, that of being God-like, such that

(G1) Definition: $G = \bigwedge \mathcal{G}$

From (P4) it follows that $\bigwedge \mathcal{G}$ is positive. So

(G2) **Pos**(G).

Therefore G is a member of the collection $\mathcal{G}$ A God-like individual has every positive property. Moreover, by axiom (P3) and by (G2) above, it follows that every property that is entailed by G is positive. Now by (N2), we see that NE is a positive property. Therefore NE is a member of the collection $\mathcal{G}$ So

(G3) $(\forall x)[Gx \rightarrow NE(x)]$,

which can be regarded as analogous to Hartshorne's basic argument. Since G is a positive predicate, by Theorem (P5) we also have

(G4) $\Diamond[(\exists x)\ Gx]$.

Combining (G3) and (G4) together, we get

(G5) $\Diamond(\exists x)[Gx \wedge NE(x)]$.

But an individual that possesses all positive attributes can possess no attributes which are not positive. If this were not the case, then the individual would possess some attribute which is not positive and its negation, which is. But this would be a contradiction. So if Gx is true, then $G{=}X$. That is, a God-like individual must have God-likeness as its essence. We can write this as

(G6) $\quad Gx \leftrightarrow (G{=}X)$.

Substituting (N1) into (G5) and using the identity (G6),[36] we get

(G7) $\quad \Diamond(\exists x)\{Gx \wedge \Box[(\exists y)\, Gy]\}$.

From (G7) we get

(G8) $\quad \Diamond\Box[(\exists y)\, Gy]$.

The contrapositive of Becker's postulate (M6) tells us that

(G9) $\quad \Diamond\Box[(\exists y)\, Gy] \rightarrow \Box[(\exists y)\, Gy]$.

Applying (G9) and (G8) finally yields

(G10) $\quad$ Theorem: $\Box[(\exists y)\, Gy]$.

which completes the proof. Q.E.D.

By Leibniz' Law, here formulated as (E1), there can only be one individual which is God-like.

9. Sobel's Objection. Sobel (1987) has objected that Gödel's argument leads to modal collapse. So this objection needs to be examined in detail. A modal logic is said to suffer from modal collapse if every true statement in the system becomes necessarily true. While this is not in itself inconsistent, it does appear to undermine the point of the logic by negating any distinction between accidental and necessary truth. Let φ be any true proposition. We define F to be the predicate "is such that φ is true." Suppose x is God-like. Then x has attribute F because, to put it simply, φ happens to be true. Since God-likeness G is the essence of x, it follows that G entails F, by (G6) and (C8). But by (G10) there necessarily exists a God-like individual. So there necessarily exists an individual with property F. It follows that φ is necessarily true.

If we accept that idea that F is a valid attribute for an individual, then the logic of Sobel's argument is inexorable. However, we must be careful to qualify the discussion by saying that not every such construction should be accepted as a valid attribute for an individual. As Gödel's notes clearly demonstrate, he considered attributes to be *the cause for the difference between things*. The attribute F defined above is an attribute that all

[36] Strictly speaking, we are using not (G6) but the fact that (G6) is necessarily true for all x.

individuals must share, and is incapable of finding the difference between things as Gödel intended. I suspect that appropriate attributes should be intrinsic, and not something for which the relationships are more superficial. However, this does not provide a complete answer to Sobel. Unless the distinction between the right and the wrong sort of attributes can be clearly delineated, Sobel's objection will hover over Gödel's argument ready to exploit any ambiguities in the logic. By encouraging emendations such as in Anderson (1990), Sobel's objection may lead to increasingly more variations on Gödel's original proof. As in mathematics or logic, one interesting proof often leads to another.

Sobel's objection is one of a class of problems that arise with ontological and cosmological arguments for the existence of God. If an argument for God's existence is interpreted as a grounding of the accidental attributes of the world in the act of creation of a necessary being, then it flirts with modal collapse: it is not obvious to see why a necessary being would create a world with accidental attributes. The traditional solution to the problem is to suppose that God freely chose to create this particular world and that its accidental attributes are consequences of that freedom. As is often the case, with solutions involving free will, this solution is also open to criticism. If God's actualisation of this world was a free choice, then the act of creation seems to be capricious or random. However, despite Einstein's opinion that "God does not play dice," many modern physics have become comfortable with the idea that the basic structure of the universe is governed by the laws of quantum probability. It is conceivable that God could have rolled quantum dice, so to speak, before decide which of the possible universes to create. Another solution to the problem was given by Leibniz, who contended that God created the best of all possible worlds.[37] A third explanation is provided by the philosophical stance known as *modal realism*. Modal realists, such as David Lewis, argue that our world is not to be distinguished from possible worlds by being actualised, as Leibniz suggested. Rather all possible worlds exist, rather like parallel universes in science fiction, or Everitt's many-worlds interpretations of quantum mechanics.

10. Anderson's Objection. C. Anthony Anderson has objected to axiom (P2) on the grounds that it does not permit neutral attributes.[38] For example, being of average height would seem to be a neutral attribute. However, axiom (P2) demands that either it is positive or that not being of average height is a positive attribute. There seems to be something wrong with the claim that either is positive. Anderson recommended that axiom (P2) be replaced by

$$\text{(A1)} \qquad \mathbf{Pos}(F) \rightarrow \neg\, \mathbf{Pos}(\neg\, F).$$

Axiom (A1) is compatible with the idea that some attributes are neither positive nor

[37] Leibniz' argument that this world is the best of all possible worlds was parodied by Voltaire in his play *Candide*, and has been widely misunderstood ever since. The idea that everything that happens is for the best was not Leibniz' contention. Leibniz did not claim that every aspect of the world is optimal when taken in isolation. Rather that among all possible worlds this world is the best when all aspects are considered together.

[38] See Anderson (1990).

negative. However, it imposes a *consistency* condition on positivity, namely that it is impossible for both an attribute and its negation to be positive. Gödel's formulation in (P2) is much stronger: it adds a *completeness* assumption to the collection of positive predicates. The collection is complete in the sense that it is impossible to extend his class of positive predicates without obtaining an inconsistency. Anderson has forcefully argued that completeness is too strong.

This criticism seems to be obvious, and might well make us wonder why a logician of Gödel's stature made such an elementary mistake about the appropriateness of a completeness assumption on a class of predicates. But I do not believe that Gödel was mistaken. Rather, I believe that if Gödel was at fault here, it was in failing to communicate his intentions more clearly in the choice of his axioms (P2) and (P3). Suppose F and H are two attributes that are generally agreed to be positive and negative respectively. As $F \Rightarrow F \vee H$, axiom (P3) implies that $F \vee H$ is a positive property. Many people would object to this because the predicate $F \vee H$ is completely symmetrical in its positive and negative components. So any argument that $F \vee H$ is positive based upon F should be mirrored by an argument that $F \vee H$ is negative based upon H. Did Gödel miss this?

I believe he did not. Part of the problem is that Gödel's definition of the positive is not this common sense one. A more explicit and careful explanation of his intentions here is to say that $\mathbf{Pos}(F)$ denotes something less restrictive than pure positivity as Anderson interpreted it. Let me argue that Gödel meant $\mathbf{Pos}(F)$ to denote the *logical consistency* of the attribute F with all purely positive attributes. With this interpretation, axiom (P2) becomes more reasonable. If pure positivity is a consistent notion, then either a given attribute F will be consistent with pure positivity or its negation $\neg F$ will be consistent. This is simply a formulation of axiom (P2). The statement $\mathbf{Pos}(F \vee H)$ is true because $F \vee H$ is entailed by F, which is consistent with the purely positive. If logical consistency of F with pure positivity is sufficient to ensure $\mathbf{Pos}(F)$ then we have the axiom

(A2) $\quad \mathbf{Pos}(F) \vee \mathbf{Pos}(\neg F)$.

At this point, the reader may object that (A2) has been defended at the expense of (A1). Is it possible that both F and $\neg F$ might *both* be logically consistent with pure positivity? While this statement looks strange, it simply asserts the logical independence of F from the class of positive attributes. For example, it is well known that the continuum hypothesis of set theory is independent in this way of the standard axiomatization.[39] One can add either the continuum hypothesis or its negation to the axiom system and retain a consistent set of axioms. Can F be independent of positive attributes in this way? The answer is no provided that the class of purely positive attributes is *sufficiently rich*. To

[39] The continuum hypothesis states that there is no set whose cardinality (number of elements) is strictly between the cardinality of the set of natural numbers and the set of real numbers. The standard axiomatization of set theory is not sufficiently rich, i.e., strong, to prove either the continuum hypothesis or prove its negation. This concept of logical independence lies at the heart of Gödel's work on the foundations of mathematics.

determine this, we will examine the concept of the purely positive in greater detail in the next section.

11. The Nature of Perfection. It might well be said that Gödel's ontological argument stands or falls on the coherence and the interpretation of his concept of the positive. So it is of particular importance that we examine Gödel's intentions in this matter, as well as other possible interpretations. To do so, we shall consider Gödel's axioms of positivity and privation in light of two semantic systems, which I shall call *Leibnizian semantics* and *Plotinian semantics*. The two semantic interpretations are roughly analogous to the One and the Many as described in Plato's *Parmenides*, the debate being resolved in favour of the God as conjoined complexity in Leibnizian semantics, and undifferentiated unity in Plotinian semantics. Both interpretations are ontological, rather than moral-aesthetic.

In Leibniz' metaphysics, the collection of *all attributes* is constructed by combining together *simple attributes* using the rules of conjunction, disjunction or negation. Simple attributes are always positive, although positive attributes need not be simple. When simple attributes are "combined" using conjunction, positivity is preserved. Gödel would agree with this, and would add that when attributes are "relaxed" by disjunction, positivity is preserved whenever at least one of the attributes is simple. In this system, all attributes, including those which are not positive can be written as a Boolean combination of simple positive ones. Privation enters the picture because a Boolean combination can include the negation of the simple attributes.[40]

This architectural approach looks promising. Gödel, an admirer of Leibniz, undoubtedly had Leibniz' interpretation in mind in developing his argument. But there are problems with using Leibnizian semantics on Gödel's argument. One of the problems is that it is difficult to know what simple attributes are. According to Leibniz, simple attributes are those that are conceived through themselves and not through other attributes. For example, the attribute of being human is complex, because a human is a featherless biped, and can be conceived as the conjunction of being featherless and being a biped.[41] On the other hand, attributes such as having extension in space or being conscious, may be simple. We might argue that simple attributes are purely semantic in nature, and have no syntactic aspects. To avoid circularities in decomposing attributes into their defining components, we must ensure that attributes are conjunctions of more fundamental ones which are at a lower level. At the lowest level simple attributes should have the characteristic that they cannot be explained to someone who has no experience of them. For example, the *qualia*[42] of consciousness appear to be simple. To someone who is completely colour-blind, it is impossible to provide any explanation of what it is to experience redness: it is an irreducible personal experience that cannot be defined in

[40] I am following the interpretation of Leibniz' metaphysics as described by Blumenfeld (1995).

[41] See Plato's *Definitions*, a work that is likely due to a follower of Plato. This infamous definition seemed strange to people in Antiquity as well. Silly as the definition is, it probably reflects their inability to define humanity any better. It is likely that a definition such as "hairless primate" will one day be just as dated.

[42] By *qualia*, we are to understand fundamental qualitative conscious experiences.

simpler terms that a colour-blind person knows. Indeed, there seems to be something intrinsically subjective about Leibniz' definition of a simple attribute. David Chalmers suggests that

> Trying to define conscious experience in terms of more primitive notions is fruitless. One might as well define *matter* and *space* in terms of something more fundamental. The best we can do is give illustrations and characterisations that lie at the same level.[43]

We should note the similarity between Leibniz' definition of a simple attribute and the classical definition of a substance. For example, Definition I.3 of Spinoza's *Ethics* is that "substance" is "that which can be conceived through itself alone." We are to take the concept of a substance as primitive in the same way that a point or a line are primitive in Euclid's system of geometry. Space, time and consciousness are argued to be substances. Simple attributes are the essences of substances. Both Leibniz and Spinoza concurred in the belief that simple attributes or substances cannot clash with each other. If the conjunction of two simple attributes created a contradiction, they would have to have conflicting component attributes, which would contradict the principle that they are simple.

Another difficulty with Leibnizian semantics is that necessary existence is positive in Gödel's argument. So it should follow that necessary existence is an conjunction of simple attributes. Naturally, the most straightforward solution is that necessary existence is itself a simple attribute, but this is not intuitively clear. Gödel's argument places existence as a second-order property: an attribute of attributes. An attribute exists if it is exemplified by some individual. Necessary existence is also a second-order property. An attribute exists necessarily if it is exemplified in every possible world. Necessary existence is then transferred to individuals by using Leibniz' Law (E1) to identify individuals with their essences. So necessary existence seems to be decomposable into the more primitive ideas of modality and essence. If this is the case, necessary existence is far from simple. Nevertheless, necessary existence can be formulated as an conjunction of existential statements over possible worlds. If each of these existential statements is positive in character then an argument for **Pos**(*NE*) could be assembled from these pieces. However any formal logic in which positivity could be conjoined over possible worlds would require a metalogic in which possible worlds could be treated as individuals, statements within each world as attributes of those worlds, and the modes of possibility and necessity as existential and universal quantifiers over possible worlds.

We shall now turn to an alternative semantic interpretation of Gödel's argument. If Leibnizian semantics builds from the bottom up, Plotinian semantics works from the top down. For Plotinus, the deepest levels of reality are characterised, not by their complexity, but by their simplicity. If simple attributes cannot be decomposed into simpler ones, it is because, for Plotinus, the simple attributes inherit their unitary character from God — undifferentiated, eternal being with no inherent multiplicity. Plotinus accepted the

---

[43] Chalmers (1996), page 4.

idea that to exist and to be one are the same. If I break the coffee cup that stands on my desk, the cup ceases to be because it has dissolved into parts. Plotinus gave the example of an army, which, when deprived of unity ceases to exist.[44] God, as creator of all existing things, is the universal provider of all unity. As such, God is the One. God is beyond predication, because to predicate an individual is to limit that individual and thereby to allow for the possibility of something outside of it. This, argued Plotinus, would make Absolute Reality complex, not simple, and can be ruled out by his ontological postulate that to come into existence is to become a unity from many things.

At first glance, such a metaphysical system would seem to be far removed from Gödel's argument. However, a careful examination of Plotinus' argument shows that he did not convincingly prove that the One is beyond predication or attribution. Lloyd Gerson has argued that we should interpret this argument of Plotinus as implying that God's necessary existence and God's essence are identical attributes. For example, Gerson states

> I think we shall better understand Plotinus' highly creative and nuanced response to this problem if we suppose that the One's being "beyond being" does not mean that it has no nature or essence at all or that it is a blank ontological place-holder or bare particular. Rather its essence is identical with its existence and is therefore unqualifiedly simple. By contrast, if in everything else essence or nature or "whatness" is really distinct from existence, then what each being is can be conceived of apart from its existence. If the One is identical to its existence, conceiving of it is impossible.[45]

If $G = NE$, as Gerson proposes, then the class of positive predicates is simply the collection of all those which are entailed by necessary existence, and axioms (P1), (P2) and (P3) become theorems: logical consequences of the fact that NE is an essence. Unfortunately, we have no clear way of determining whether a given a attribute F is entailed by NE. Formally, an attribute F is entailed by NE if it can be represented in the form $NE \vee H$, for some choice of H. However, there is no obvious way to determine whether such a representation exists.

Together, I would argue that Leibnizian and Plotinian semantics form a *via affirmativa* and a *via negativa* for theological discourse. The former puts an emphasis on those positive aspects of God which compose reality. The latter insists on God's transcendence, and the inability of ordinary predicates to provide an adequate description of God. If this interpretation is correct, then it is worth noting that European mysticism has consistently stated that both approaches are valid. The *via negativa* has been considered the higher path. However, the tension between the two has been dialectical, rather than antagonistic. Both semantic systems are neoplatonic in the broad sense. I leave it to the reader to consider their validity and their consistency with each other.

12. The Cosmology of a Neoplatonic World. There is a certain sense in which no proof

---

[44] *Enneads* VI.9.1.
[45] Gerson (1998), page 6.

can ever be deemed valid if its conclusions are sufficiently unpalatable. In the foundations of mathematics, the axiom of choice is a case in point. While the axiom looks innocuous, it has a number of unpalatable consequences, such as the Banach-Tarski Theorem.[46] Those who find the conclusions of the Banach-Tarski Theorem unacceptable have usually implicated the axiom of choice. On the other hand, those who find the conclusions of an argument particularly appealing have often been accepting of its premises, even when these are clearly fallacious. For example, we know that the axiomatization of geometry proposed by Euclid is unsatisfactory because it makes unwarranted assumptions about the interior and exterior of sets. This fact was overlooked for centuries because the results provided by Euclidean geometry were so satisfactory. This is not to say that mathematics has no external reality apart from our practice of it. However, mathematics is not the dry development of lemmas, corollaries and theorems from axioms and postulates. There is, rather, a more dynamic relationship between the premises of a mathematical argument and its conclusion.

If this is the case in mathematics, it is surely more so in theology and the philosophy of religion. The person who finds the conclusions of Gödel's argument unpalatable, has much to attack in its premises. The concept of necessity is arguably vague,[47] and the concept of positivity is more so. It is likely that each attack can be defended in some way, as ontological arguments have persistently maintained their presence in religious speculation despite some insightful criticisms. Throughout this essay, I have tried to keep a critical eye on each step of the argument to determine whether its premises, like the axiom of choice, are in accord with common sense. Those who find the assumptions of the ontological argument suspicious should ask themselves whether their suspicion is based, like the Banach-Tarski paradox, on an unwillingness to accept the conclusion of the argument. Those who find the assumptions of the ontological argument acceptable should ask themselves whether they have been too forgiving of the ambiguities of the logic. Either of these approaches would be a mistake. Each step of the argument should be judged on its own merits.

Although the premises of the ontological argument are suspect, there is much to learn by examining the consequences of certain ontological assumptions. Gödel's proof gives us an opportunity to examine the formal "architecture" of a universe that is neoplatonic in the broad sense. The definition of God, the embodiment of all positivity, is built from the

[46] The Banach-Tarski Theorem states that it is possible to cut up a solid body such as a sphere into a finite number of pieces and reassemble them to make a larger sphere. There are no holes in the larger sphere, and the pieces are moved rigidly without stretching or distortion. Surprisingly, the fact that this is possible leads to no contradiction in mathematics. The fact that such a dissection is possible in a nonconstructive sense is now taken for granted by mathematicians. Since the result used the axiom of choice in its proof, the Banach-Tarski Theorem has been used to argue that the axiom of choice allows too much freedom in the construction of sets. However, other mathematicians accept the Banach-Tarski Theorem as good mathematics.

[47] This is the point made Hick (1968) with reference to Hartshorne's proof. Whatever criticisms are valid for Hartshorne's argument must surely be equally valid for Gödel's argument, depending as it does on even more difficult issues in modal logic. Is logical necessity the right kind of necessity to discuss God? The issue is surely confounded by our inability to see things from the viewpoint of eternity.

collection of positive simple predicates, which are the “bricks” of the argument. The principle that the collection of positive predicates is closed under conjunction provides the “mortar” of the argument, cementing the bricks together. If there are no such things as positive properties then the building is a castle in the air. On the other hand it may be that positive properties do exist. Properties such as omniscience, omnipresence and omnipotence can be studied to determine whether they entail any privation, as we might understand it. Alternatively, we can take the axioms as given, and search for attributes that satisfy the axioms. This would be like turning the proof on its head, and using its logical structure to define what we mean by God.

To what extent is the God of Gödel’s argument compatible with the modern scientific world view? The ontological argument attempts to ground all existing things in a necessary being, namely God. In a similar fashion, basic science attempts to ground the phenomena of this world in a deeper type of necessity. Physical laws are not logically necessary, but are necessary to a greater extent than simple facts because they are presumed to be true at all times and all locations.[48] So the search for a fundamental unifying physical theory can be regarded as similar in spirit to the search for a grounding of all creation in a necessary being. Before our understanding of the ontological mystery can be complete, we will need to have a better understanding of how to ground the world’s contingencies in essential truths. This is the goal of Gödel’s argument and the scientific quest for a final theory.

Does Gödel’s argument imply pantheism? That is to say, is maximal positivity — God — identifiable with the universe? The answer to this question depends as much on our understanding of the universe as it does on God. We are often quick to categorise thinkers as pantheist or not. However, the distinctions blur, because the growth of our understanding of relativity and quantum mechanics makes the universe appear much more transcendent in its fundamental attributes than was true in Isaac Newton's time. If the universe is defined as all that exists, then it is true that all theists are by definition pantheists. But this blurring is not useful. We should make a distinction between pantheism and *panentheism*, the latter being the belief that the universe is an aspect or epiphany of God but is not identical with God. Panentheism usually takes the viewpoint that the universe is that aspect of existence which is accessible to observation or ordinary experience. Gödel’s argument is certainly intended to be panentheistic. Whether it is pantheistic probably depends on whether there exist positive attributes which are unknowable by any observer.

---

[48] Physical laws are those which are invariant under an appropriate class of transformations of space and time. In this sense, the interpretation of the necessity operator in tense logic as being “true at all times,” has its relevance here. To illustrate the point, the necessity operator • φ can also be interpreted to mean that a certain proposition φ is true under all Lorenz transformations of space-time, say. The Lorenz group is the class of isometries of Minkowski space-time that is associated with the theory of special relativity. So the statement “mass-energy is conserved” is true at a more necessary level in this interpretation that the statement “mass is conserved.” The latter is true only in the accidental sense for inertial frames of reference which are relatively slow with respect to the speed of light, whereas the former is true for all inertial frames. Symmetries also arise in other physical theories such as quantum mechanics.

It takes a certain amount of courage to state publicly that you believe God's existence can be proved. Many scholars who admire Kurt Gödel as the greatest logician of the twentieth century have found themselves a bit embarrassed by his ontological argument. His Incompleteness Theorems are undoubtedly masterpieces of mathematics. His rotating universe is respectable in physics because its properties can be explained without reference to his belief that Becoming is in some sense illusory. By showing that the rotating universe is conceptually consistent with general relativity, he helped make academic discussion of time travel acceptable for later researchers such as Kip Thorne. As a proof theorist, he helped make modal logic respectable by demonstrating the utility of modal operators in the foundations of mathematics. By establishing the consistency of the continuum hypothesis with Zermelo-Fraenkel set theory, he helped make Cantor's transfinite arithmetic respectable. Yet the ontological argument remained on the margin of intellectual thought. It is now emerging from this intellectual wilderness to a more prominent place where it belongs. Gödel's research forms a unified body of thought in the Platonic tradition. Within that body of thought, his ontological argument is no aberration, but an essential part of one of the most remarkable thinkers of modern times.

References.

Adams, R. M. (1995) Introductory note to *1970. In *Kurt Gödel, Collected Works, Vol 3*. Oxford University Press, Oxford.

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

Blumenfeld, D. (1995). Leibniz's ontological and cosmological arguments. In *A Cambridge Companion to Leibniz*, edited By Nicholas Jolley, Cambridge University, Cambridge.

Chalmers, D. J. (1996). *The Conscious Mind*. Oxford University, Oxford.

Gerson, L. (1998). *Plotinus*. Routledge, London.

Gödel, K. (1933) Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematischen Kolloquiums*. Franz Deuticke, Vienna. English translation in *Kurt Gödel, Collected Works, Vol 1*. Oxford University Press, Oxford 1986, 300-302.

Gödel, K. (1995) *Kurt Gödel, Collected Works, Vol 3*. Oxford University Press, Oxford, p. 437.

Hartshorne, C. (1962) *The Logic of Perfection and Other Essays in Neoclassical Metaphysics*. Open Court Press, Lasalle, Illinois.

Hawking, S. and Penrose, R. (1996) *The Nature of Space and Time*. Princeton University Press, Princeton, NJ.

Hick, J. (1968) What did Anselm discover? In *The Many-Faced Argument*, ed. John Hick and Arthur McGill. Macmillan, London.

Knuuttila, S. and Hintikka, J. (1986) *The Logic of Being*. Reidel, Dordrecht.

Malament, D. B. (1995) Introductory note to *1946b. In *Kurt Gödel, Collected Works, Vol 3*. Oxford University Press, Oxford, 261-268.

Plantinga, A. (1974) *The Nature of Necessity*. Clarendon Press, Oxford.

Sobel, J. H. (1987) Gödel's ontological proof. In *On Being and Saying: Essays for Richard Cartwright*, ed. J. J. Thompson. MIT Press, Cambridge, Mass.

Stein, H. (1995) Introductory note to *1946/9. In *Kurt Gödel, Collected Works, Vol 3*. Oxford University Press, Oxford, 202-229.

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

Wang, H. (1996) *A Logical Journey: From Gödel to Philosophy*. MIT Press, Cambridge, Mass.