A New Look at the Cosmological Argument

Robert C. Koons *
Department of Philosophy
University of Texas at Austin
koons@phil.utexas.edu

December 18, 1996

1 Introduction

The cosmological argument for God’s existence has a long history, but perhaps
the most influential version of it has been the argument from contingency. This
is the version that Frederick Copleston pressed upon Bertrand Russell in their
famous debate about God’s existence in 1948 (printed in Russell’s 1957 Why I
am not a Christian). Russell’s lodges three objections to the Thomistic argu-
ment:

- There is no intelligible form of necessity other than logical truth.
- There is no reason to suppose that any such thing as the “universe” exists.
- Even if there were such a thing as the “universe”, our empirical knowledge
gives us no good reason to assume that it has a cause.

Almost fifty years later, Russell’s objections seem quite dated, dependent
on a form of logical empiricism that has not weathered the intervening years
well. The logic and metaphysics of possibility and necessity have proved to be
a fruitful and rich area of investigation. Cosmology – the study of the universe
as a whole – has matured and gained respectability. The notion of causation
has taken root once again within philosophy, proving to be indispensable to
recent advances in semantics, epistemology and cognitive science. The theory of

*I would like to acknowledge my debt to Norman Geisler, Mortimer Adler, William Lane Craig
and Stanley Jaki, who are largely responsible for keeping the tradition of the cosmological argument
alive in the latter half of the twentieth century. I would also like to thank my colleagues in the
Philosophy Department at the University of Texas – Austin, Dan Bonevac, T. K. Seung, and Cory
Juhl, as well as Doug Geivett at Biola University, Professor Richard Swinburne of Oxford University,
and an anonymous reviewer, for their helpful comments on an earlier draft of this paper.
reasoning with and about causation has advanced far in recent years, resulting in a growing body of knowledge about so-called ‘defeasible’ or nonmonotonic inference systems. The time has come to take a new look at the cosmological argument, in light of the recent recovery within philosophy of so much of the classical elements of metaphysics.

I will follow closely the classical argument from contingency, with its origins in Aristotle’s *Metaphysics* Lambda 6 and developed by the *falsafa* movement of Arabic philosophy (al-Farabi and Ibn Sina). My argument closely resembles Maimonides’ fourth proof and Aquinas’ Second and Third Ways. The argument is rigorously empirical in character: I nowhere make claims to a priori knowledge (other than of the rules of classical logic). There is no claim of great originality to the argument presented here. What is original is the use of three logical resources that were not available to the classical authors: (1) mereology (the calculus of individuals – essentially a variant of Cantorian set theory adapted to aggregates of concrete things), (2) modern modal logic, and (3) nonmonotonic logic (the theory of defeasible reasoning). I lay out a successful defeasible argument for the existence of a necessary First Cause and discuss briefly its relevance to natural theology.

## 2 The Rehabilitation of Causality

In 1917, Bertrand Russell announced the demise of the concept of causality in his essay, “On the Notion of Cause” (in [38]). Subsequent developments in science and analytic philosophy have not supported Russell’s contention. Far from withering away, the notions of cause and effect have never held a more central position.

The notion of causality is absolutely central to recent philosophical work in semantics, the philosophy of mind and intentionality, epistemology, and philosophy of science. Work by Donnellan, Kripke [20], and Putnam [34] helped to make causal connections an indispensable part of our accounts of reference and signification. This in turn has generated causal theories of information and content ([11] and [13]). The Gettier problem led to the renaissance of causal theories of knowledge by Goldman [15], Armstrong [3], Pollock [33], and Plantinga [32]. Causality is put to much work in recent theories of personal identity and of the nature of mental states (as in the functionalism of Lewis [23] and Putnam [34]. Causation continues to figure prominently in philosophy of science (e.g., Wesley Salmon’s causal theory of evidence [41]) and in theoretical science, both within physics and outside.

Attempts to explain away causation or to replace it with some purely statistical regularity (whether or not supplemented by some kind of psychologistic decoration) have proved to be catastrophic failures. Every attempt to explain causal direction (surely one of the most fundamental features of causality) in terms of the nomological-deductive model has failed. Such models of causality
A New Look at the Cosmological Argument

have generated paradoxes far more rapidly than ad hoc solutions can be invented for them.

If a robust sense of reality leads us to recognize causal connections as first-class citizens of our ontological inventory, we must also make room for those special kinds of objects that can serve as relata for causal relations, whether we call these objects possible ‘facts’, ‘situations’, or ‘states of affairs’. These objects must be distinguished from propositions and from quasi-linguistic representations if we are to capture accurately the logical relations governing causal idioms. The restoration of such fact-like entities to respectability has also been a common theme of recent work in philosophy, including philosophical linguistics from Vendler to Asher, and the Stanford situation theory of Barwise and Perry.

3 A Modal Mereology of Facts

My formal framework will be a modal logic

supplemented by the Leśniewski-Goodman-Leonard calculus of individuals (“mereology”) [21].

There is no one-to-one correspondence between true propositions and facts. First of all, where a proposition \( p \) is verified by a fact \( a \), \( \neg p \) is typically not verified by any fact, but instead by what I shall call a “negative condition”. Similarly, if \( p \) and \( q \) are verified by facts, \( p \lor q \) (their disjunction) will not correspond to a third, disjunctive fact. Instead, if the disjunction is true, it will be true by corresponding to one or the other or both of the facts verifying its disjuncts. Secondly, supervenient truths, such as semantical, ethical, psychological, and some logical and mathematical truths, do not introduce additional facts: they correspond to very same facts to which their base truths correspond. For example, both ‘snow is white’ and ‘that snow is white is true’ correspond to the same fact, the whiteness of snow.

By way of modal logic, I need only the axioms of rules of \( T \). I will assume a fixed domain of possible facts; hence, the logic will include the Barcan and converse Barcan axioms.

I will use the two usual predicate symbols of mereology, \( \sqsubseteq \) and \( \circ \), representing part-of and overlap, respectively. I need three mereological axioms:

Axiom 1 \( x \sqsubseteq y \leftrightarrow \forall z \ (z \circ x \to z \circ y) \)

Axiom 2 \( \exists x \phi(x) \to \exists y \forall z \ (z \circ y \leftrightarrow \exists u \ (\phi(u) \& u \circ z)) \).

Axiom 3 \( x = y \leftrightarrow (x \sqsubseteq y \& y \sqsubseteq x) \)

Axiom 1 defines the part-of relation in terms of overlap, and Axiom 2 is an aggregation or fusion principle: if there are any facts of type \( \phi \), then there is an aggregate or sum of all the \( \phi \) facts. Axiom 3 guarantees that the part-of relation is reflexive and anti-symmetric.
There is only one principle linking the modal and mereological languages. Here I need to introduce a new predicate, $A$. Where $b$ is a possible fact, $Ab$ can be used to state that $b$ actually obtains.

Axiom 4 $x \subseteq y \rightarrow (Ay \rightarrow Ax)$.

Axiom 4 ensures that aggregation of facts is a form of conjunction: a whole necessitates all of its parts.

There is one special notion to be defined: that of being “wholly contingent”, represented by ‘$\nabla$’.

Definition 1 $\nabla x \leftrightarrow (Ax \& \forall y (y \subseteq x \rightarrow \neg Ay))$

A wholly contingent fact is an actual fact none of whose parts are necessary. I am not assuming that there are any necessary facts: the existence of necessary truths does not entail the existence of necessary facts (since our logic lacks a comprehension principle). As we shall see, if there are any necessary facts, they are facts of a very special kind.

It is very important not to conflate facts with true propositions. Facts are things in the world that make certain propositions true and others false. As Russell [39], Hochberg[17] and others have argued, there is no need to posit conjunctive or disjunctive facts corresponding to conjunctive or disjunctive propositions. If the atomic fact $a$ and $b$ are sufficient to make true the atomic propositions $A$ and $B$, respectively, then the sum of $a$ and $b$ is sufficient to make true the conjunction $A \& B$. Similarly, the truth of a disjunction can be grounded in a fact corresponding to either disjunct: no “disjunctive” fact is needed. For similar reasons, there is no special category of fact corresponding to existential generalizations. Such generalizations are made true by the sum of the truth-makers of their instances. (Negative and universally generalized propositions pose more difficult questions; fortunately, nothing in this paper necessitates any particular answer to these questions.)

Similarly, mathematically elaborated propositions do not require similarly elaborate facts as their truth makers. If fact $a$ makes it true that I have three coins in my pocket, then it is also sufficient to make it true that the number of coins in my pocket is the positive square root of nine. If moral truths supervene on the non-moral, then there is no need to postulate moral facts in addition to the non-moral ones.

Facts may also differ from true propositions in their identity conditions. There may not be a one-to-one correspondence even between true atomic propositions and atomic facts (assuming there are such). The proposition the chair is red is arguably atomic, but there may be a non-denumerable class of possible facts that could make this proposition true (one corresponding to each possible shade of red). Even if the properties mentioned in an atomic proposition were perfectly precise, it might well be the case that the causal antecedents of a fact
are essential to its identity, but not to the identity of the corresponding true proposition.

At the same time, there is no clear reason for distinguishing (at least in this context) between “facts”, “events” and “states of affairs”, so long as each of these are thought of as concrete parts of the world. There is certainly a difference in natural language between expressions that pick out facts in the standard way (i.e., by using a complement “the fact that” clause) and expressions that introduce events (gerundives and other nominalizations, action verbs). Nonetheless, this can be accounted for as different ways of picking out entities from the same basic ontological category. Typically, we use event language to pick out thick, complex facts, like the death of Caesar or the Civil War, by means of high-level or indirect descriptions.

4 Principles of Causation

The causal relation will be represented by a primitive binary operator, ‘$\triangleright$’. I do not in fact believe that causation is an absolutely primitive relation: my favored conjecture is that a satisfactory definition of the causal relation in terms of modal facts is possible. For present purposes, however, it will be convenient to treat causation as a primitive.

There are a number of logical properties of causation that can be expressed, for instance, the transitivity and asymmetry of causation. I will, however, need only three facts about causation for present purposes:

- **Axiom 5** Veridicality: $(x \triangleright y) \rightarrow (Ax \& Ay)$
- **Axiom 6** Separate Existence: $(x \triangleright y) \rightarrow \neg(x \circ y)$
- **Axiom 7** Universality: $\forall x (\exists y (y \triangleright x))$

Axiom 5 stipulates that only actual facts can serve as causes or effects. Axiom 6 is intended to capture Hume’s insight that a cause and its effect must be “separate existences”. The language of mereology, when applied to facts, enables us to state Hume’s principle precisely: a cause must not overlap its effect. It is very important to bear in mind that Axiom 6 does not require that a cause must not overlap its effect in space or time: it is only mereological overlap (the having of a common part) that is ruled out. Axiom 7 expresses the universality of the causal relation: every wholly contingent fact has a cause. Axiom 7 does not entail determinism, in any of its usual senses, since I have not stated that causes are sufficient conditions for their effects. I am not assuming that every event is necessitated by its causes; in fact, I believe that this is not typically the case. Causal laws are always exception-permitting or defeasible generalizations. It is quite possible for $C$ to be in every sense the cause of $E$, even though it was possible for $C$ to occur without being accompanied by $E$. 
(For this reason, this account of causation is compatible with, although it does not entail, indeterministic theories of human freedom.)

The evidence for Axiom 7 is essentially empirical. Every success of common sense and science in reconstructing the causal antecedents of particular events and classes of events provides confirmation of Axiom 7.

5 The universality of causation

5.1 The role of defeasible reasoning

Even though we have excellent empirical evidence for the generalization that wholly contingent facts have causes, it is hard to see how any amount of data could settle conclusively the question of whether or not this generalization (Axiom 7) admits of exceptions. This is a legitimate worry, but I would respond by insisting that, at the very least, our experience warrants adopting the causal principle as a default or defeasible rule. This means that, in the absence of evidence to the contrary, we may infer, about any particular wholly contingent fact, that it has a cause.

This is, however, all that is needed for the cosmological argument to be rationally compelling. The burden will be shifted to the agnostic, who must garner evidence of a positive sort for the proposition that the cosmos really is an exception to the rule. Merely pointing out the defeasible nature of the inference does not constitute a cogent rebuttal.

Considerable progress has been made in recent year in developing formal systems of defeasible or nonmonotonic reasoning that satisfy certain plausible meta-logical constraints. For example, in the Commonsense Entailment system of Asher and Morreau [4], a defeasible version of Axiom 7 could be expressed by using a default conditional connective, $\triangleright$

$$\forall x (\triangleright x \triangleright \exists y (y \triangleright x))$$

This version of Axiom 7 can be read as: normally, a wholly contingent fact has a cause. This defeasible Axiom 7 will allow us to infer that any given wholly contingent fact has a cause unless some positive reason can be given for thinking that the fact in question is an exception to the rule, for example, by showing that the fact belongs to a category of things that typically does not have a cause.

5.2 Is the universality of causation merely heuristic?

In his debate with Copleston, Russell insisted that there is a difference between claiming that scientists should always look for a cause and claiming that there is always a cause there to be found. Russell followed Kant’s suggestion that the universality of causation be seen as a canon or prescriptive rule for reason, and
not as a description of mind-independent reality. The cosmological argument depends on using the principle of universality as a descriptive generalization.

I have two principal responses. First, it is hard to see why the abundant success of empirical science in finding causes for contingent facts does not provide overwhelming empirical support for the generalization to all contingent facts. The category of wholly contingent facts is not an unnatural, gerrymandered kind like ‘grue’ or ‘bleen’. Are we to believe that it is merely a coincidence that time and time again we find causes for contingent facts?

Second, the denial of the universality of causation as a descriptive generalization constitutes a very radical form of skepticism. All of our knowledge about the past, in history, law and natural science, depends on our inferring causes of present facts (traces, memories, records). Without the conviction that all (or nearly all) of these have causes, all of our reconstructions of the past (and therefore, nearly all of our knowledge of the present) would be groundless. Moreover, our knowledge of the future and of the probably consequences of our actions depends on the assumption that the relevant future states will not occur uncaused. The price of denying this axiom is very steep: embracing a comprehensive Pyrrhonian skepticism.

6 The Cosmological Argument

Besides the logical principles presented above, the cosmological argument depends on only one factual premise: that there exists a contingent fact. For example, suppose there are an odd number of molecules in my pencil at the present moment: surely there could have been an even number. A single contingent fact of this kind is all that I need, although I believe that nearly every fact with which we are acquainted is contingent. I would go so far at so say that every physical fact is contingent.

6.1 The nature of modality

In saying that a fact is contingent, I am saying much more than merely that the corresponding proposition is neither logically true nor logically false. A contingent fact is one that is actual but could have been non-actual, where the relevant notion of possibility is that of broadly metaphysical possibility. Broadly metaphysical possibility is the fundamental form of possibility, of which all other kinds (physical, historical, legal, etc.) are qualifications or restrictions.

Attempts since the days of logical positivism to reduce metaphysical possibility to logical consistency (or logical consistency with all definitional or “analytic” truths) have failed. First, it has proved impossible to specify the “analytic” truths without making reference to possibility and necessity. Second, nothing is gained in clarity unless we insist on using first-order logic, which, as John Etchemendy has argued [12], is an implausible construal of logical consistency.
Finally, the attempt to avoid the supposed “mysteries” of metaphysical possibility in this way leads to the much more serious difficulties of set-theoretic platonism, with the attendant mysteries of how these transcendent mathematical entities connect to the rest of reality and, most crucially, of how we can obtain reliable knowledge of them. Recent efforts at making sense of mathematical reality make use of the notion of metaphysical modality (as in the “possible structures” of Hellman [16]), indicating that the proper order of explanation starts with modality, not with mathematical entities.

If we deny that there are any contingent facts, then we must conclude that we live in a world in which all three modalities – possibility, actuality, and necessity – collapse together. This is tantamount to denying that these modalities can do any interesting work. Such a denial runs athwart the growing body of philosophical work in which modality plays a central role.

6.2 A sketch of the proof

Lemma 1 All the parts of a necessary fact are themselves necessary.

Proof. By Axiom 4 and the K axiom of modal logic.

Lemma 2 Every contingent fact has a wholly contingent part.

Proof. Let \( a \) be a contingent fact. If \( a \) is wholly contingent, we are through, since \( a \) is a part of itself. Otherwise, \( a \) has a necessary part. By Axiom 2, there exists a fact \( \dot{x} (x \subseteq a & \square Ax) \) that consists of the aggregate of all the necessary parts of \( a \). Since \( a \) is contingent, \( a \) itself is not a part of \( \dot{x} (x \subseteq a & \square Ax) \). By Axiom 1, there is an \( b \) that overlaps \( a \) but not \( \dot{x} (x \subseteq a & \square Ax) \), hence there is a part of \( a \), say \( c \), that is not a part of \( \dot{x} (x \subseteq a & \square Ax) \).

We can show that \( c \) is wholly contingent. Suppose that \( d \) is a part of \( c \). Then \( d \) is part of \( a \) but \( d \) does not overlap \( \dot{x} (x \subseteq a & \square Ax) \). Hence, \( d \) is not necessary. Since \( d \) was an arbitrary part of \( c \), \( c \) is wholly contingent.

Definition 2 Let \( C \) be the aggregate of all wholly contingent facts.

By axiom 2, it follows that if there are any wholly contingent facts, then any fact overlaps \( C \) if and only if that fact overlaps some wholly contingent fact.

\[ \exists x \quad \nabla x \rightarrow \forall y \quad (y \amalg C \leftrightarrow \exists z \quad (\nabla z & y \amalg z)) \]

Lemma 3 If there are any contingent facts, \( C \) is a wholly contingent fact.

Proof. Suppose that there is at least one contingent fact. Then there is also a wholly contingent part, by the preceding lemma. To show that \( C \) is wholly contingent, we must show that every part of \( C \) is contingent. Let \( a \) be a part of \( C \). Since \( a \) is a part of \( C \), \( a \) overlaps \( C \), by Axioms 1 and 3. Hence, \( a \) overlaps
some wholly contingent \( b \) (by the definition of \( C \)). It is a theorem of mereology that two facts that overlap have a common part. Hence, some \( d \) is part of both \( a \) and of \( b \). Since \( b \) is wholly contingent, \( d \) is contingent. By Lemma 1, if \( a \) were necessary, \( d \) would be necessary. Consequently, \( a \) is contingent. Therefore, since \( a \) was an arbitrary part of \( C \), \( C \) is wholly contingent.

**Lemma 4** If there are any contingent facts, \( C \) has a cause.

**Proof.** An immediate consequence of Lemma 3 and Axiom 7, the Universality of Causation.

**Lemma 5** Every contingent fact overlaps \( C \).

**Proof.** Let \( a \) be a contingent fact. By Lemma 2, \( a \) has a wholly contingent part, say \( b \). By axiom 2 and the definition of \( C \), \( C \) and \( b \) overlap.

**Theorem 1** If there are any contingent facts, then \( C \) has a cause that is a necessary fact.

**Proof.** By Lemma 4, \( C \) has a cause. By Axiom 6 (Separate Existence), this cause does not overlap \( C \). By Lemma 5, every contingent fact overlaps \( C \). By Axiom 1 (Veridicality), the cause of \( C \) is actual. Hence, the cause of \( C \) must be a necessary fact.

Since we know that there is at least one contingent fact, we can identify \( C \) with the cosmos, and use Theorem 1 to conclude that the cosmos has a cause that is a necessary fact, a First Cause. It is legitimate to call this cause a “first cause” if we assume (as seems plausible) that all effects are contingent.

## 7 Identifying the First Cause

Demonstrating the existence of a First Cause is of course not the same thing as demonstrating the existence of God as conceived, for example, in biblical theology. Nonetheless, the result of the cosmological argument is quite useful to the project of natural theology, providing very helpful support to a number of important arguments for theism. In this section, I will first lay out what I take to be plausible corollaries of Theorem 1, and then I will discuss briefly the relevance of Theorem 1 to the argument from design.

**Corollary 1** The cause of the cosmos includes the existence of a necessary being.

I presume that every fact includes at least one being and at least one property of that being. In addition, I assume that a being cannot be involved in an actual fact without actually existing. Hence, a necessary fact entails the necessary existence of some being (or system of beings), which we might as well call “God”.
Corollary 2  God (the necessary being included in the cause of the cosmos) is not a mere composite or aggregate object.

A composite or aggregate object cannot exist necessarily, since a constituent part of a mere aggregate can exist in the absence of the rest of the aggregate. Hence, an aggregate must have parts that exist only contingently, which means that the aggregate as a whole must exist contingently.

Corollary 2 does not entail that God is absolutely simple. God could have parts, as long as God is not merely the aggregate of those parts, that is, as long as the parts cohere together in an essential (not merely aggregative) unity.

Corollary 3  God has all of its basic attributes by necessity.

By a (basic) attribute of a thing, I mean all of the causally and ontologically fundamental properties of the thing, such that all of the other properties of a thing supervene on or are caused by its basic attributes. God’s basic attributes are all included in the First Cause and hence God possesses all of these attributes necessarily.

Corollary 4  All of the parts of God have all of their attributes by necessity.

If God has any parts, then its having each of these parts is an attribute of God. By Corollary 3, God has all of its parts by necessity. I will assume that if God has a part, than any attribute of that part corresponds to an attribute of God. Consequently, since God has all of its attributes by necessity, so must any parts of God.

Corollary 5  God has only immeasurable attributes.

Any attribute that is measurable participates in the structure of the more and the less. The more and the less constitute a continuous spectrum. Consequently, it seems reasonable to assume that for any measurable attribute $A$, where $A$ consists in having determinable $D$ to degree $\mu$, and any being $x$ that has $A$, there is some degree $\epsilon$ such that it is possible for $x$ to have $D$ to degree $\mu - \epsilon$ or $\mu + \epsilon$. Therefore, no measurable attribute can be had by necessity.

This means that if God has a size, it must be infinitely large. If God has an age, it must be infinitely old. If it makes sense to attribute power or intelligence to God, then that intelligence and power must be absolutely infinite in quality. The only finite properties that God could have essentially are those that involve whole integers, such as existing as three persons. It doesn’t make sense to exist as 3.01 persons, so this attribute counts as immeasurable.

Corollary 6  God is not essentially located in space or time.
If God were located in space or time, it would be located either in a part or the whole. If it were located in a part, its location would be measurable and hence contingent. The location of an essentially spatio-temporal being is one of its basic attributes, so it would have a contingent attribute, contrary to the Corollary 3. If it were essentially located in the whole of space and time, it would have spatio-temporally located parts, each of which would have a contingent location, and hence a contingent attribute, contrary to Corollary 4.

**Corollary 7** God is not essentially a physical object, nor is it essentially constituted by physical objects.

A physical object, or something constituted by physical objects (like Aristotle’s bronze statue) would be located in space-time, contrary to the last corollary. Consequently, neither God nor any of its parts can be (in their very essence) physical objects.

### 7.1 The cosmological argument as support for the teleological argument

Teleological arguments and arguments from design all start with various observable features of the cosmos and use these features as evidence of the existence of an intelligent designer/creator of the cosmos. These observable features include the anthropic and biothropic characteristics of the fundamental physical constants and of the nature of the Big Bang ([6]), the intelligibility ([24]) or order and lawfulness ([44]) of nature, or the apparent improbability of various significant events, such as the original emergence of life or the eventual emergence of consciousness ([5], [44]).

From a logical point of view, these arguments from design take quite different forms, depending on whether or not they rely on the cosmological argument as a preparatory step. Since Paley, it has been customary for theists to present the teleological argument as separate from and wholly independent of the cosmological argument. This Paleyian argument relies heavily on an argument from analogy. The cosmos resembles human artifacts, human artifacts have designer/creators, therefore it is likely that the cosmos has a designer/creator. It is misleading to call this Paleyian argument an argument from design, since it is really an argument to design.

As was famously exploited by Hume, such arguments from analogy are highly vulnerable to the observation of significant disanalogies between the two cases, and there are surely plenty of these. In addition (and this was also pointed out by Hume), the analogical argument to design seems to lead to a problematic infinite regress. Human creators/designers are highly complex and integrated systems; hence, it is very likely that the cosmic creator is also a complex, integrated system. As such, the cosmic creator resembles human artifacts. Therefore, it seems likely that the cosmic creator has a creator, and so on to infinity.
Suppose, however, that we think about the teleological argument in close connection with the cosmological argument, as Aquinas did. In this case, we already know that the cosmos has a First Cause, and that this cause is necessary and involves a necessary being, whom we call “God”.

The fact that a set of facts has been ordered to some purpose is empirically verifiable and does not logically entail (although it may suggest) the existence of any personal intentionality. A teleological law is simply a projectible, empirical generalization, which can be used to explain a set of facts by reference to their common effects (not their causes). Teleological generalizations do not compete with or contradict causal laws: instead, they partly supervene on them (in the case of certain anthropic generalizations, they entirely supervene on causal laws). This supervenience of the teleological on the causal does not make the teleological reducible to the causal, nor any less real or less explanatory than the causal.

For the sake of this argument, let us presume that we have discovered such teleological generalizations at the level of the cosmos, such as: all physical constants and Big Bang conditions are such as to make possible complex life forms. The cosmos, so characterized, is the effect of the First Cause. We attribute intelligence to human beings because of the teleological generalizations that characterize the actions of normal human beings. Since the effects of the First Cause are strongly analogous to the effects of human action in exactly this respect, we have the strongest possible reason for attributing to God something analogous to intelligence.

In the Paleyian argument, we start with only three terms, human artifacts, human creators, and the cosmos, and we argue that since the cosmos resembles human artifacts, and human artifacts are caused by humans, it is probable that the cosmos has been caused by something similar to humans. Dissimilarities between the cosmos and human artifacts are quite relevant to this argument. In Figure 1, we start with the information on the lefthand side (the production of artifacts by human intelligence), plus the resemblance of the cosmos to human artifacts. Everything in the upper-right rectangle (including the existence of God and His causing of the cosmos) is part of the conclusion to be drawn.

In the Thomistic argument, we start with four causally related terms: humans as cause of human actions, and the First Cause as cause of the cosmos. We notice that the cosmos shares the very feature of human actions upon which we base our attribution of intelligence to humans. We conclude that the First Cause is in some sense intelligent. Dissimilarities between the cosmos and human actions are irrelevant to this inference. In addition, there is here no threat of an infinite regress, since the teleological argument was not our reason for positing the cause of the cosmos, only our reason for characterizing it as intelligent. In Figure 2, the cosmological argument supplies us with the existence of God and His causing of the cosmos. The teleological argument supplies us with the resemblance (in respect of purposiveness) between the cosmos and human artefacts. The new conclusion to be drawn consists only in the resemblance (in
respect of intelligence and purposefulness) between humans and God.

8 Objections

8.1 Don’t we need to observe the origin of many universes?

Hume argued [19] that we can infer that the world has a cause only by collecting a large sample of worlds and observing that nearly all of them have causes. Hume is assuming that the cosmological argument depends on a premise of the form: all (or nearly all) worlds have causes. However, this is obviously unnecessary. Hume seems to be assuming that in order to apply a well-supported generalization to a new case, I must know that the generalization applies to every possible reference class to which the case belongs (or to all the most specific reference classes). To know that a rubber ball dropped on a Tuesday in Waggener Hall by a red-headed tuba player will fall the ground, I must have observed a large sample of such balls dropped by such tuba players at just this location on a Tuesday. This is clearly absurd.

We know that all (or nearly all) wholly contingent facts have causes, the world is such a wholly contingent fact, and therefore we may conclude that the world has a cause, unless some relevant consideration pointing to the opposite conclusion can be produced.
8.2 Isn’t causation valid only for the phenomenal world?

In the first Critique, Kant argues that causation pertains only to the apparent or “phenomenal” world, not to the real or “noumenal” world. His argument depends on assuming that the fundamental causal principles are known prior to experience, and that nothing substantial or material about the real world can be known by us prior to experience. Kant’s objection is relevant only to a priori arguments for God’s existence, like those of Scotus or Leibniz. It is not relevant to an argument like mine that rigorously appeals only to empirical, a posteriori arguments. I am not claiming that the axioms of causality I am appealing to are known by us prior to their application to the world of experience. Instead, I appeal to our success in finding causal explanations as empirical evidence for these generalizations.

8.3 What about quantum mechanics?

Quantum mechanics is sometimes taken to provide abundant counter-evidence to the universality of causation. Quantum mechanics raises two problems for our understanding of causality: the indeterminism of wave collapse (under the Copenhagen interpretation), and the Bell inequality theorems. The indeterminism of quantum transitions during observation does not contradict Axiom 7. I have not assumed that causes necessitate their effects: in
fact, I strongly suspect that such an assumption is incoherent (if “necessitate” is understood in a strong sense). According to the Copenhagen version of quantum mechanics, every transition of a system has causal antecedents: the preceding quantum wave state, in the case of Schrödinger evolution, or the preceding quantum wave state plus the observation, in the case of wave packet collapse.

The Bell inequalities demonstrate that the data described by quantum mechanics forces us to reject one of the following three principles:

- Causal influences never travel backwards in time.
- Causal influences never travel faster than the velocity of light.
- Every reliable (projectible) correlation has a causal explanation.

In discussions of the Bell inequalities, the third principle is sometimes labelled a law of “causality”. It is, however, much stronger than my Axiom 7. I have nowhere assumed that (as the third principle implies) a cause always “screens off” (in Reichenbach’s sense) its effects from non-posterior states.

The Bell inequalities are merely another demonstration of the impossibility of reducing causation to some sort of statistical relationship. They raise no difficulties for a causal realist such as myself. In my opinion, the most reasonable response to the Bell inequalities would be to restrict one or more of the three principles above to macroscopic (large-scale or classical) phenomena and to restate them as defeasible (exception-permitting) rules.

Axiom 7 implies that each wholly contingent fact is caused: it does not imply that every correlation can be explained. For example, suppose that you and I meet (coincidentally) in the market. Our meeting has a cause, which is simply the sum of the cause of your being there at that time and the cause of my being there at that time. There may well be no explanation of why we are both there at the same time. It could even be that we reliably meet there time and time again, and yet there is no causal explanation of our repeated coordination. Similarly, the statistical correlations described by quantum mechanics may (for all I know) remain forever mysterious. We may find no causal explanations of their truth. Nonetheless, the individual observations making up each correlated pair still certainly have causes.

8.4 Is the argument compatible with creation as a free act?

It is not immediately clear how the conclusion of this cosmological argument sits with the traditional conception of God as a person who creates the world as a free, undetermined action. Since the relata of the causal relation are facts, the account of causation seems to exclude the sort of libertarian agency favored by those esposing some form of “agent causation”.

A New Look at the Cosmological Argument
The two accounts of causation can, however, be reconciled. When one substance (such as an agent) causes a change in the world through exercise of one of its powers on some substance (either itself or another) with a corresponding liability to be changed, there are clearly two facts: (1) the composite fact that the agent substance possessed the relevant power and that the patient substance possessed the relevant liability at the relevant point in time, and (2) the fact that the change occurred. The first fact is at least part of the cause of the second fact. Nothing in the account given above requires that the effect be necessitated by the cause. Hence, there is no requirement of determinism: the actual exercise of the power in producing the change might not have happened, even if the cause (and all concomitant facts) had existed without alteration. The account of causation given here is neutral on the question of determinism, and also on the question of the compatibility of determinism with human (or divine) freedom.

In the case of creation, it is clear that we are dealing with an undetermined action. If God's being necessitated His creating this world, then this world would be necessary after all, contrary to the arguments above. Thus, we have at least one instance of undetermined, intentional action.

8.5 Doesn't the cosmological argument assume the impossibility of an infinite regress?

Leibniz was the first to realize that the cosmological argument does not depend on any assumption about the impossibility of infinite regresses. Even if there are infinite regresses of causes within the totality of contingent facts, the totality itself must have a cause that is outside it and, hence, a cause that is necessary. The crucial assumption is Axiom 2, the assumption that any non-empty set of facts can be aggregated into a single fact. This corresponds to the pre-modern denial of infinite regress, since it in effect denies that any such totality is what Cantor termed an “absolute” or improper totality (like the set of all sets, or the set of ordinal numbers).

There is little if any reason to think that there is anything improper about the totality of all wholly contingent facts. We are talking only about ontologically basic facts, not about mathematical or semantical truths that supervene upon them. We are simply aggregating concrete particulars, and we are not running afoul of Russell’s vicious circle principle in the process. There is no reason to postulate any facts that somehow involve or presuppose the totality of all facts, or of all contingent facts.

8.6 Doesn’t the cosmological argument commit the fallacy of composition?

Russell accused Copleston of committing the fallacy of composition, arguing that because each of the parts of the world is caused, the whole must be caused.
The cosmological argument includes no such error: it is demonstrated that the cosmos is itself a wholly contingent fact, and for that reason must have a cause.

8.7 Isn’t necessary existence an impossibility?

A number of twentieth century philosophers follow Hume in holding that only logical truths can be necessary, that the very notion of a necessary fact is incoherent.

Two replies. First, we have not assumed the existence of a necessary fact: this was the conclusion, not a premise, of the argument. Thus, this so-called objection simply fails to engage the argument. The objector is content merely to deny the conclusion without bothering with the premises or the reasoning.

Second, the Humean principle being relied upon is self-defeating. Is it supposed to be true by definition that only logical or definitory truths are necessary? Surely in saying this, Hume, Russell, et al. intended to be saying something informative. How could such a principle be contingent? What sort of contingent facts about the actual world make it the case that there are no non-logical necessities? What empirical justification have the anti-essentialists provided for their claim?

In response, the objector must simply deny that he can make any sense of this notion of modality, except insofar as it is replaced by the clear and well-behaved notion of logical consistency. This sweeping denial of modality is simply obscurantist, undermining fruitful philosophical research into the nature of natural law, epistemology, decision, action and responsibility, and a host of other applications.

8.8 Doesn’t the cosmological argument presuppose the ontological argument?

The argument presented above in no way presupposes any version of the ontological argument. It does presupposes that the notion of necessary fact is coherent, and this may be what Kant meant when he tried to tie the cosmological argument to the ontological argument.

8.9 Don’t contingent facts typically have contingent causes?

This is probably the most promising line of rebuttal to the cosmological argument. It is an instance of a wider strategy: focus on some unique feature of the First Cause, and point out the cause of the world’s having that feature is an exception to some well-established generalization. Indeed, for the most part, contingent facts do have contingent causes. They also have causes with finite attributes and causes that can be located in space and time, unlike the hypothesized First Cause. Once we have established that the cosmos is relevantly unusual, we seem to be faced with two equally unattractive options: supposing
that the cosmos has only a very unusual kind of cause, or supposing that it has no cause at all. Thus, we end in a stalemate.

The defender of the cosmological argument must respond with substantial reasons for thinking that, although the First Cause is unique in a number of respects, each of these unique features can be adequately explained by extrapolating from tendencies already observable in ordinary cases of causation. For instance, I would conjecture that, in some precise sense, a cause is always more necessary or less contingent than its effect.

One very simple definition of relative necessity would be the following:

\[ a \text{ is more necessary than } b \leftrightarrow \exists y \left[ \Box(\neg a \rightarrow \neg Ab) \land (\Box a \land \neg \Box \neg Ab) \right] \]

In other words, a fact \( a \) is more necessary than fact \( b \) just in case \( a \) holds in every world in which \( b \) holds, but not vice versa. This follows from the identity conditions of facts. The causes of a fact are essential to its identity: had the very same truth been verified by a fact caused in a different way, we would not have had the same fact as verifier. The corresponding thesis involving effects is not plausible: a fact’s identity does not include the eventuality of all its effects. The contingency of the evolution of the world depends on this asymmetry: a fact’s holding necessitates the holding of its causes, but not of its effects.

This principle (an effect necessitates the existence of its causes) does not imply that the content of an effect necessitates the content of its causes. For example, the fact of Caesar’s death could not have existed had not all of its causes, including Brutus’ knife-thrust, existed. This of course does not mean that Caesar wouldn’t have died unless Brutus and the other senators had killed him. The truth ‘Caesar died’ would have been verified by a different fact in all of those worlds in which Brutus does not help in inflicting the fatal set of wounds. The fact that actually verifies the truth ‘Caesar died’ would not have existed had any of its causes failed to exist.

There are several additional reasons (besides the one involving the identity conditions of facts) for thinking that causes are more necessary than their effects. First, there is the authority of Aristotle and the Aristotelian tradition. Second, it is clear that we need some account of causal priority that explains the transitivity and asymmetry of this relation. An account of causal priority in terms of relative contingency nicely satisfies this desideratum. Finally, this account enables us to specify exhaustively the ‘potential causes’ of a given fact: \( a \) is a potential cause of \( b \) if and only if \( a \) is less contingent than \( b \). Such a specification is necessary if we are to account for the statistical properties of causal connections, the so-called “Markovian principles” developed by Salmon [41] and Suppes [43] and studied recently by Pearl and Verma [31] and Spirtes, Glymour and Scheines [42].

However relative contingency is defined, it is clear that the cosmos is a fact of absolutely minimal contingency. If fact \( a \) contains fact \( b \) as a part, then \( b \) is no less contingent (no more necessary) than \( a \), since \( a \) could not exist if \( b \)
did not exist. Since the cosmos contains every wholly contingent fact as a part, no wholly contingent fact can be less contingent than the cosmos. Since the cosmos is a fact of minimal contingency, it is not surprising that it should have no contingent cause, but it would still be very surprising if it had no cause at all.

8.10 Where did God come from?

If we’re right in thinking that causes must be strictly more necessary than their effects, it follows that necessary facts cannot be caused (at least, not in the ordinary sense).

Another reason for thinking that necessary facts cannot be effects is this: we know that the totality of all facts cannot be caused (since there is no fact that does not overlap it), and the best explanation of this fact is that this totality contains necessary facts, and necessary facts cannot be caused.

8.11 The James Ross objection: did God cause that He caused the world?

James Ross ([35], pp. 295-304) has argued that the principle of sufficient reason can be demonstrated to be false. His objection can be adapted into an objection to my Axiom 7 (the Universality of Causation) as follows. Consider the fact that the First Cause caused the cosmos. Call this fact $C^*$. $C^*$ is clearly a contingent fact, since if it were necessary, the cosmos itself would be necessary (by Axiom 5, veridicality). If $C^*$ is also wholly contingent, then it must be a part of the cosmos, and the First Cause must cause $C^*$, i.e., the First Cause must cause the fact that it causes the cosmos. The same argument can be repeated, showing that the First Cause must cause that it causes that it causes the cosmos, ad infinitum. This appears to be a vicious infinite regress.

The best answer to this objection is to point out that there is no reason to think that $C^*$ is wholly contingent. The fact that the First Cause causes the cosmos would appear to be composed of two facts: namely, the First Cause on the one hand and the cosmos on the other. The truth that the first caused the second does not represent a third fact in addition to the first two. Instead, such statements about single-case causal connections supervene upon the cause, the effect, and certain non-factual truths about the modal relationship between the cause and the effect. Therefore, the wholly contingent part of $C^*$ is simply the cosmos itself, and we are forced only to re-affirm that the First Cause does cause the cosmos.

This response entails that there are no facts, over and above facts about modality and other non-causal matters, corresponding to single-case causal nexi. That is, we are assuming that causal truths are supervenient on modal and other non-causal truths (including truths about objective chance or propensity, and about powers and liabilities). Causal connections between facts in a world are
to be explained entirely in terms of what has happened in that world, and what might or probably would happen in it and alternative worlds. This sort of modest ontological reduction is quite attractive, since the alternative is to posit causal nexi as brute facts, without any logical relationship to predictability or to statistical regularities. At the same time, this sort of modest reduction does not entail the eliminability of causal discourse, nor does it obviate in any way the necessity of positing facts as an ontological category. Causation is a relation between facts, not any kind of propositional operator, but any particular causal nexus between facts consists of some aggregation of other modal, stochastic and historical facts.

8.12 William Rowe’s objection

William Rowe ([37], pp. 108-110) has proposed a variant of Ross’s objection to the cosmological argument. Rowe asks us to consider the fact $a$ that corresponds to the true proposition: there are contingent (positive) facts. Most defenders of the cosmological argument will accept that $a$ is itself contingent. Therefore, the First Cause must cause $a$. However, the fact that the First Cause has caused $a$ is itself a contingent fact, so the First Cause would have to cause the fact that it caused $a$, and so on, ad infinitum.

The proper response to this objection is only slightly different from the response to the last objection. The proposition that there are contingent facts does not correspond to a single fact. Facts are not closed under existential generalization, as propositions are. From the existence of a fact that $n$ has $F$, it does not follow that there is a distinct fact that something has $F$. Consequently, the fact that makes Rowe’s $a$ true is simply the cosmos itself, and no infinite regress can be generated.

This is not simply an ad hoc response, since there are independent grounds for denying the existence of a special category of existential facts. Causation is transparent: that is, if the fact that there is an $F$ caused $a$, then there is some $n$ such that the fact that $n$ is $F$ caused $a$. Similarly, if the proposition that there is an $F$ has been made true by some fact $a$, then there is some instance of this generalization that has been made true by $a$. Thus, in neither case is there any reason to posit a special category of fact corresponding to the existential quantifier.

8.13 Can’t the anthropic principle be explained by positing infinitely many parallel universes?

A standard non-theistic response to the data underlying the anthropic principle is to suggest that there may be an infinity of parallel universes, representing every possible permutation of possible physical laws and initial Big Bang conditions. Only an infinitesimally fraction of these permit the development of life
and consciousness, but it is not surprising that we inhabit one of these vanishingly rare universes, since otherwise we would not be here to observe it.

For the sake of clarity, let’s stipulate a few definitions. Let us call each of the spatio-temporally complete, causally isolated histories a ‘universe’. The totality of all such universes I shall call the ‘cosmos’. The cosmos is thus a vast aggregate, composed of infinitely many parallel universes. I have argued above that an aggregate like the cosmos cannot exist necessarily: there are infinitely many sub-aggregates that could have existed in its place, each sub-aggregate being just like the actual cosmos except in lacking one or more actual universes. The cosmos thus requires a cause. Since the parallel universes are causally isolated, the First Cause must directly cause each of the constituent universes. In particular, the First Cause must cause our universe.

Since it is empirically verifiable that our universe is objectively ordered to an end (namely, the eventual existence of conscious life), we have good reason to characterize the First Cause as purposeful and intelligent. The fact that the First Cause may also have caused other universes for other purposes is beside the point. It may be that the apparent purposefulness of the First Cause is merely an illusion. It may be that the First Cause blindly caused every possible sort of universe, including very many ordered to no end at all. This bare possibility is not enough to rebut the teleological argument: we must have some positive evidence for the existence of such “junk” universes, but of course we have and could have no such evidence.

This objection brings out the importance of considering evidence for design in the context of the cosmological argument. Without the cosmological argument, we would be forced to the conclusion John Leslie reaches[22], namely, that there are two equally good possible explanations of the anthropic data: a cosmic designer, and observer selection in a world of infinitely many, uncaused universes. However, once we know that the cosmos (including our universe) has a cause, the second hypothesis is excluded. It is still possible that the First Cause caused a junky cosmos, and that the evidence for intelligent design is illusory, but in the absence of positive evidence for these other universes, the reasonable inference to draw from the only universe we can observe is that the First Cause encompasses the existence of an intelligent designer.

There is another serious drawback to the junky cosmos hypothesis: if employed globally, it has the consequence that any form of induction is demonstrably unreliable. If we embrace the junky cosmos hypothesis to explain away every appearance of orderedness in the universe, then we should assume that the simplicity and regularity of natural law is also an artifact of observer selection. Universes would be posited to exist with every possible set of natural laws, however complex or inductively ill-behaved.

Now take any well-established scientific generalization. Among the universes that agree with all of our observations up to this point in time, the number that go on to break this generalization is far greater than the number that continue to respect it. The objective probability that every generalization we have observed
extends no farther than our observations is infinitely close to one. Thus, relying on induction in such a universe is demonstrably futile.

In short, the junky cosmos hypothesis is both the most flagrant possible violation of Occam’s razor and a death sentence to all other uses of that principle. This hypothesis postulates an infinity of entities for which there is absolutely no positive evidence, simply in order to avoid the necessity of explaining the anthropic coincidences we have observed. This is the height of metaphysical irresponsibility, far worse than the most extravagant speculations of medieval angelology. Moreover, it undermines all subsequent appeals to simplicity or economy of explanation. If the junky cosmos hypothesis is true, it is demonstrable that the simplest hypothesis of astronomy or biology is no more likely to be true of our universe than the most complicated, Rube-Goldberg constructions. We would have absolutely no reason, for instance, to believe that the Copernican hypothesis is more likely to be true than a fantastically complex version of Ptolemy’s system, elaborated as far as necessary to save the astronomical phenomena.

9 God and the Metaphysical Hyper-darwinists

Two great creation myths contend for dominance. According to the first, in the beginning was Chance. Chance was responsible for the origin of the cosmos and for its causal constitution. Hence, all subsequent events, including the origin of life and consciousness on earth, are ultimately attributable to the agency of Chance and of Chance alone. The impression that something more has been involved is an illusion generated by the operation of natural selection over long periods of time and (perhaps) by the statistical bias necessitated by the fact that conscious observers can only occur in very special kinds of universes.

According to the alternative myth, the cause of the cosmos is a positive reality that is not, with felicity, referred to under the name of “Chance”. Instead, this first cause bears some analogy to intelligence and purposefulness. We can even glean something of its quasi-purposes: the eventual emergence of complex life, consciousness and meaningful agency.

It is one of the ironies of history that the devotees of the first myth have claimed the authority of Science for their doctrines, attributing all conviction attaching to the competing myth to dogma, wishful thinking, and a variety of psychopathologies. In fact, there is nothing especially scientific about the first myth: in some form it antedates science by millennia, being one of the first explanations (or pseudo-explanations) of the cosmos. It is the second myth that has the far greater claim to the allegiance of Science, both because it was first formulated at the inception of science in antiquity, and also because it fostered the growth of scientific knowledge through promoting a faith in the contingent intelligibility of the universe. Pierre Duhem and Alfred North Whitehead have both argued persuasively that it is not accident that the scientific revolution of
the thirteenth through sixteenth centuries occurred in Christian Europe, where
the second myth had become for the first time deeply rooted in a society’s official
worldview.

Defenders of the Chance myth have in recent years resorted to the authority
of Charles Darwin ([26] [9] [10]). In this they have been assisted by some
misguided defenders of the alternate myth who have saddled it with pseudo-
scientific baggage, such as a 10,000 year old cosmos or a dogmatic rejection of
human evolution. Darwin himself was, at least in his public pronouncements,
a believer in the theistic creation myth. He believed that God was needed as
an explanation of the orderly cosmos we observe (see [8], pp. 395-6). He was
opposed, not to creation as such, but only to the theory of the special creation
of each of the world’s species.

Metaphysical hyper-darwinists (like Monod, Dawkins and Dennett) are helped
in their misappropriation of the support of Darwinism by a confusion between
randomness and Chance. Modern neo-Darwinian theory postulates that genetic
variation is the result of the random introduction of mutations. The relevant
notion of ‘randomness’ has a precise, mathematical characterization: it means
that the sequence of mutations is non-computable, and possesses thereby cer-
tain statistical properties. The randomness of a sequence does not entail that
any member of the species is wholly unintended. Many events studied in social
science form random sequences even though every single event is intended by
some person or other.

Moreover, the randomness of a sequence is compatible with the entire se-
quence being intended by an agent, so long as the agent has the capacity to
generate non-computable intentions. It is a matter of controversy whether or
not human beings have such capacities, but there is no reason to deny such a ca-
pacity to God. The plans of an infinite mind can constitute a random sequence
of events, in the precise sense of the word.

To demonstrate that the cosmos can be explained wholly in terms of chance,
therefore, it is not sufficient to show that life is the product of a random sequence
of mutations, filtered through natural selection. Instead, one must show that the
actual universe (including its history) belongs to a space of possible universes
and that, when some natural probability measure is defined over this space, the
measure of the region of universes relevantly like this one is reasonably high.
Recent evidence concerning the anthropic principle has decisively refuted this
conjecture. In addition, work by mathematicians concerning the likelihood of
the chance origin of life and the chance development of highly complex forms of
life points to the same conclusion (see for example, the symposium of biologists
and mathematicians hosted by the Wistar Institute [28], and the recent book
by Hubert Yockey [45]).

The future progress of science, and the successful defense of science against
its post-modernist and relativist opponents, depends on the severing of the mis-
taken connection between science and the Chance creation-myth. The inherent
hostility of the Chance creation-myth to the rationality to science is made clear
by the final, definitive form of the myth: the junky cosmos hypothesis. This hypothesis egregiously violates the fundamental canon of scientific rationality – the preference for economical explanations – and it undermines that canon in all of its applications.

References


A New Look at the Cosmological Argument


