A TOPOS-THEORETIC FORMALIZATION OF ABSTRACT OBJECT THEORY

HARRY TMETIC

Abstract. We present a categorical formalization of Edward Zalta's Abstract Object Theory (AOT) using topos theory. Our approach models the important distinction between encoding and exemplification of properties through a presheaf topos construction over a category of property-indexed families of sets. This framework naturally accommodates impossible objects (such as the round square), fictional entities, and mathematical abstracta within a consistent logical system. We establish that abstract objects correspond to subobjects of a special encoding object in the topos, while the comprehension principle for abstract objects arises from the topos-theoretic subobject classifier. The internal logic of our topos validates key principles of AOT while maintaining consistency even for objects encoding contradictory properties. We show applications to formal ontology, the analysis of fictional discourse, and provide a novel categorical perspective on ontological arguments.

1. Introduction

From Frege's pioneering attempts to reduce arithmetic to logic to today's type-, set-, and category-theoretic systems, the deeper challenge of justifying the existence and knowability of abstract objects has persisted. While set theory has provided one successful foundation for mathematics, it struggles to accommodate the full range of abstract objects that appear in philosophical discourse: fictional characters like Sherlock Holmes, impossible objects like the round square, and intentional objects of thought and language.

Edward Zalta's Abstract Object Theory (AOT), developed over four decades (Zalta 1983; Zalta 1988), offers a sophisticated logical framework that handles these diverse abstract objects through a crucial distinction: objects may *encode* properties (a mode of predication special to abstract objects) or *exemplify* properties (ordinary predication applicable to all objects). This dual predication allows AOT to maintain that the round square encodes both roundness and squareness without exemplifying either, thus avoiding contradiction while preserving our intuitions about impossible objects.

The key insight of AOT is that encoding is an internal, definitional mode of predication—abstract objects encode properties by their very nature—while exemplification is an external mode requiring actual instantiation. This separation permits consistent reasoning about impossible and fictional entities: Sherlock Holmes encodes the property of being a detective (since this is part of his definition in Conan Doyle's stories) without exemplifying it (since no actual detective named Sherlock Holmes exists at 221B Baker Street).

Despite AOT's philosophical sophistication and recent computational implementations (Kirchner, Benzmüller, and Zalta 2020), it has remained isolated from mathematical foundations. In this paper, we attempt to bridge this gap by providing a categorical formalization of AOT using topos theory. Our construction reveals that AOT's philosophical insights align naturally with the mathematical structures of presheaf topoi, while topos theory provides new tools for analyzing abstract objects.

1.1. **Main contributions.** Our primary contributions are:

(1) A categorical model of AOT as a presheaf topos $\mathcal{Z} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ where $\mathbf{C} = \mathbf{Set}^{\mathbf{P}}$ for a set $\mathbf{P}$ of properties, providing a mathematical foundation for dual predication that naturally separates internal properties from external instantiation.
(2) A proof that Zalta's comprehension principle for abstract objects—that for any condition on properties, there exists a unique abstract object encoding exactly those properties—follows naturally from the topos-theoretic subobject classifier, without requiring any additional axioms.
(3) A demonstration that impossible objects can be consistently accommodated within our framework as abstract objects lacking global elements, resolving long-standing puzzles about round squares and their kin through the natural distinction between internal and external existence in topoi.
(4) A categorical analysis of ontological arguments that clarifies the logical gap between encoding and exemplifying necessary existence, identifying what additional axiom is required to bridge from conceptual to actual existence.
(5) A unification of techniques from categorical logic, modal semantics, and formal ontology that could open new research directions at the intersection of mathematics and philosophy.

1.2. **Technical overview.** Our construction begins with a set $\mathbf{P}$ of properties, treated as atomic entities without presupposed relationships. We form the category $\mathbf{C} = \mathbf{Set}^{\mathbf{P}}$ of $\mathbf{P}$-indexed families of sets, where each object represents a possible configuration of property extensions. The presheaf topos $\mathcal{Z} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ then provides our model of AOT.

Within this topos, we identify a special *encoding object* E whose elements at each stage are property profiles—assignments of properties to elements. The key insight is that abstract objects are subobjects of E, with encoding and exemplification distinguished by different logical criteria:

- An object *encodes* a property if all its property profiles assign that property (an internal, universal statement in the topos)

- An object *exemplifies* a property if it has a global element (actual realization) that assigns that property (requiring external witness)

This separation allows impossible objects to encode contradictory properties (having property profiles that assign both) while exemplifying neither (having no global elements). The comprehension principle emerges from the bijection between subobjects of E and characteristic morphisms $E \to \Omega$, where Ω is the subobject classifier of $\mathcal{Z}$.

1.3. **Philosophical and mathematical significance.** Our work shows that the philosophical distinction between encoding and exemplification corresponds precisely to the mathematical distinction between internal properties and global elements in a topos. This correspondence is not merely formal—it reveals that topos theory provides the natural mathematical setting for theories of intentionality and abstract objects.

The use of presheaf topoi is particularly significant. Unlike the category of sets, presheaf topoi naturally accommodate "varying" or "staged" existence, where objects may exist at some stages (contexts) but not others. This matches our intuitions about fictional and impossible objects: they exist as concepts or within certain contexts (stories, theories) without existing globally.

2. Preliminaries

We follow the notation and conventions of Mac Lane and Moerdijk (1992).

2.1. **Topos theory essentials.**

Definition 2.1. An *elementary topos* is a category $\mathcal{E}$ that:

(1) Has finite limits
(2) Is cartesian closed (has exponential objects)
(3) Has a subobject classifier Ω

The subobject classifier is the key feature distinguishing topoi from other categories. It provides an internal notion of truth values and enables the interpretation of logic within the category.

Definition 2.2. A *subobject classifier* in a category $\mathcal{E}$ with finite limits consists of an object Ω and a morphism $\mathbf{true} : 1 \to \Omega$ such that for every monomorphism $m : S \rightarrowtail X$, there exists a unique characteristic morphism $\chi_m : X \to \Omega$ making the following diagram a pullback:

$$\begin{tikzcd} S \arrow[r, tail, "m"] \arrow[d] & X \arrow[d, "\chi_m"] \\ 1 \arrow[r, "\mathbf{true}"] & \Omega \end{tikzcd}$$

This universality property establishes a bijection between subobjects of X and morphisms $X \to \Omega$, enabling logical operations on subobjects.

Definition 2.3. For a small category $\mathbf{C}$, the *presheaf topos* $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ consists of:

- Objects: Contravariant functors $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{Set}$
- Morphisms: Natural transformations between functors

Proposition 2.4. *The category* $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ *is an elementary topos for any small category* $\mathbf{C}$.

The subobject classifier in a presheaf topos has a concrete description:

Lemma 2.5. *In the presheaf topos* $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$*, the subobject classifier* Ω *is given by:*

$$\Omega(C) = \{S \mid S \textit{ is a sieve on } C\}$$

where a sieve on C *is a collection of morphisms with codomain* C *closed under precomposition. The morphism* $\mathbf{true} : 1 \to \Omega$ *picks out the maximal sieve at each stage.*

2.2. **Abstract Object Theory overview.** Zalta's AOT (Zalta 1983; Zalta 2019) is built on the following philosophical principles:

(1) **Dual predication**: There are two fundamental modes of predication—encoding and exemplification. Abstract objects can encode properties they don't exemplify, while ordinary objects exemplify properties without encoding them.
(2) **Comprehension for Abstract Objects**: For any expressible condition ϕ on properties, there exists a unique abstract object that encodes exactly those properties satisfying ϕ. This principle generates a rich universe of abstract objects without paradox.
(3) **Modal distinction**: Ordinary objects necessarily fail to encode properties (encoding is reserved for abstract objects), while abstract objects may encode properties without exemplifying them. An object is abstract if it couldn't possibly exemplify concrete existence.

The formal language of AOT includes:

- Individual variables $x, y, z, \dots$ ranging over objects (both ordinary and abstract)
- Property variables $F, G, H, \dots$
- Two atomic formula types: Fx (x exemplifies F) and xF (x encodes F)

The comprehension principle for abstract objects is formalized as the schema:

$$\exists! x(A!x \wedge \forall F(xF \leftrightarrow \phi(F)))$$

where $A!x$ means "x is abstract" and $\phi(F)$ is any formula in which x doesn't occur free. This asserts that for each condition on properties, there exists a unique abstract object encoding exactly those properties satisfying the condition.

3. The Topos-Theoretic Model

We now develop our categorical formalization of AOT. The key insight is that property-indexed families provide natural "stages" or "contexts" for evaluating abstract objects, while presheaves over these stages capture the varying nature of abstract existence.

3.1. Basic construction.

Definition 3.1. Let $\mathbf{P}$ be a set of properties, treated as atomic entities. Define:

(1) The base category $\mathbf{C} = \mathbf{Set}^{\mathbf{P}}$, the category of $\mathbf{P}$-indexed families of sets
(2) The Zalta topos $\mathcal{Z} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$, the presheaf topos over $\mathbf{C}^{\mathrm{op}}$

Let us explain this construction. An object of $\mathbf{C}$ is a functor $X : \mathbf{P} \to \mathbf{Set}$ (viewing $\mathbf{P}$ as a discrete category), which amounts to a family $\{X_p\}_{p \in \mathbf{P}}$ of sets. We interpret X_p as the extension of property p at stage X—the set of entities that exemplify property p in context X.

A morphism $f : X \to Y$ in $\mathbf{C}$ is a natural transformation, which consists of a family of functions $\{f_p : X_p \to Y_p\}_{p \in \mathbf{P}}$. Composition and identities are defined componentwise, making $\mathbf{C}$ a well-defined category. In fact, $\mathbf{C}$ is equivalent to the product of $|\mathbf{P}|$ copies of $\mathbf{Set}$, and hence is itself a Grothendieck topos (the presheaf topos over the discrete category $\mathbf{P}$).

Remark 3.2. The choice of $\mathbf{C} = \mathbf{Set}^{\mathbf{P}}$ reflects the idea that properties can have varying extensions across different contexts or possible worlds. A morphism $f : X \to Y$ represents a transformation that preserves property structure—if an element has property p in context X, its image has property p in context Y.

The presheaf topos $\mathcal{Z} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ then consists of contravariant functors from $\mathbf{C}$ to $\mathbf{Set}$. Objects of $\mathcal{Z}$ assign to each "possible world" $X \in \mathbf{C}$ a set of entities that may exist at that world, with restriction maps for world transformations.

3.2. The encoding object.

The central construction in our formalization is the encoding object, which represents property profiles—the ways entities can be characterized by properties.

Definition 3.3. Define the encoding object $E \in \mathcal{Z}$ by:

$$E(X) = \mathbf{P}^{\bigcup_{p \in \mathbf{P}} X_p}$$

where $\mathbf{P}^Y$ denotes the set of functions from Y to $\mathbf{P}$.

Equivalently, $E(X) = \prod_{x \in \bigcup_p X_p} \mathbf{P}$, the product of copies of $\mathbf{P}$ indexed by all elements appearing in any X_p. An element $e \in E(X)$ is a property profile that assigns to each element in the stage X some property from $\mathbf{P}$.

Lemma 3.4. *E is a well-defined presheaf with functorial action given by restriction of property profiles.*

Proof. For a morphism $f : Y \to X$ in $\mathbf{C}$, we need to define $E(f) : E(X) \to E(Y)$. Given a profile $e : \bigcup_p X_p \to \mathbf{P}$, we define $E(f)(e)$ as the composite $e \circ \bigcup_p f_p : \bigcup_p Y_p \to \mathbf{P}$. This is well-defined since f maps elements of Y to elements of X componentwise. Functoriality follows from the fact that composition of functions is associative and preserves identities. $\square$

Remark 3.5. The encoding object E captures all possible ways of assigning properties to elements. Not every profile in $E(X)$ needs to respect the actual properties of elements in X—E includes "inconsistent" or "imaginary" profiles. This generality is crucial for modeling abstract objects that encode properties they don't exemplify.

3.3. **Encoding and exemplification relations.** We now formalize the fundamental distinction between encoding and exemplification through relations in the topos.

Definition 3.6. Define the universal encoding relation $\varepsilon \subseteq E \times y(\mathbf{P})$ in $\mathcal{Z}$, where $y(\mathbf{P})$ is the constant presheaf with value $\mathbf{P}$. At each stage X:

$$\varepsilon(X) = \{(e, p) \in E(X) \times \mathbf{P} \mid p \in \text{range}(e)\}$$

In the internal logic: $(e, p) \in \varepsilon$ if and only if the profile e assigns property p to some element.

The relation ε captures when a property profile includes a given property. This provides the foundation for defining encoding and exemplification.

Definition 3.7. An *abstract object* in our model is a subobject $A \rightarrowtail E$. We say:

(1) A *encodes* property p if $\vdash_{\mathcal{Z}} \forall e \in A. (e, p) \in \varepsilon$
(2) A *exemplifies* property p if there exists a global element $\sigma : 1 \to A$ with $(\sigma, p) \in \varepsilon$ in the point determined by σ

The crucial distinction here cannot be overstated: encoding is an *internal* universal statement in the topos logic (all profiles in A include property p), while exemplification requires an *external* witness (an actual global element having that property). This precisely captures Zalta's philosophical distinction: abstract objects encode properties by their very nature (internally), while exemplification requires concrete instantiation (global elements).

Definition 3.8. An *ordinary object* is an abstract object $O \rightarrowtail E$ that:

(1) Has at least one global element $\sigma : 1 \to O$
(2) Encodes no properties (i.e., encodes exactly the empty set)

This definition aligns with Zalta's characterization: ordinary objects are those that exist concretely (have global elements) and acquire properties only through exemplification, not encoding. The fact that they encode no properties distinguishes them from abstract objects, which are characterized by what they encode.

4. The Comprehension Principle

The heart of AOT is the unrestricted comprehension principle for abstract objects. We now show this principle holds in our topos model.

4.1. **Construction of comprehending objects.**

Theorem 4.1 (Comprehension)**.** *For any subset $\Phi \subseteq \mathbf{P}$, there exists a unique abstract object A_Φ that encodes exactly the properties in Φ.*

Proof. Define the subobject $A_\Phi \rightarrowtail E$ via its characteristic morphism $\chi_\Phi : E \to \Omega$. For each $X \in \mathbf{C}$ and profile $e \in E(X)$:

$$\chi_\Phi(X)(e) = \begin{cases} \top_X & \text{if } \forall p \in \mathbf{P} : (p \in \text{range}(e) \iff p \in \Phi) \\ \bot_X & \text{otherwise} \end{cases}$$

where $\top_X$ is the maximal sieve on X (all morphisms into X) and $\bot_X$ is the empty sieve.

This defines a natural transformation $\chi_\Phi : E \to \Omega$ since the condition is preserved under restriction of profiles. By the universal property of the subobject classifier, this determines a unique subobject $A_\Phi \rightarrowtail E$.

To verify that A_Φ encodes exactly Φ: By construction, a profile e belongs to $A_\Phi(X)$ if and only if $\text{range}(e) = \Phi$. Therefore:

- If $p \in \Phi$: Every profile in A_Φ has p in its range, so $(e, p) \in \varepsilon$ for all $e \in A_\Phi$, thus A_Φ encodes p.
- If $p \notin \Phi$: No profile in A_Φ has p in its range, so there exists $e \in A_\Phi$ with $(e, p) \notin \varepsilon$, thus A_Φ doesn't encode p.

Uniqueness follows from the uniqueness of characteristic morphisms: if two subobjects encode exactly the same properties, they must have the same characteristic morphism and hence be equal. □

Corollary 4.2. *The comprehension schema*

$$\exists! x \forall F (xF \leftrightarrow \phi(F))$$

is valid in the internal logic of $\mathcal{Z}$ for any formula $\phi(F)$ expressible as membership in a subset of $\mathbf{P}$.

This establishes that our topos model fully validates Zalta's comprehension principle. Every conceivable collection of properties determines a unique abstract object encoding exactly those properties.

Remark 4.3. The comprehension principle in our model is stronger than Zalta's original formulation. While Zalta requires ϕ to be expressible in his formal language, our construction works for arbitrary subsets of $\mathbf{P}$. This corresponds to a second-order comprehension principle in the internal logic.

4.2. **Impossible objects.**

An important test of our framework is its treatment of impossible objects—entities that encode contradictory or incompatible properties.

Definition 4.4. Properties $p, q \in \mathbf{P}$ are *incompatible* if no element can simultaneously exemplify both in any context. Formally: for all $X \in \mathbf{C}$ and all $x \in \bigcup_r X_r$, if $x \in X_p$ then $x \notin X_q$.

Proposition 4.5. *Let $\Phi \subseteq \mathbf{P}$ contain incompatible properties. Then A_Φ has no global elements.*

Proof. Suppose $p, q \in \Phi$ are incompatible. A global element $\sigma : 1 \to A_\Phi$ would determine a profile $e \in A_\Phi(1)$ with range exactly Φ. The terminal object $1 \in \mathbf{C}$ has exactly one element in each component. For this element to be assigned both p and q by profile e would require it to belong to both 1_p and 1_q, contradicting incompatibility. Therefore $\mathrm{Hom}(1, A_\Phi) = \emptyset$. □

Example 4.6 (The Round Square)**.** Let $\Phi = \{\text{Round}, \text{Square}\}$ where these properties are incompatible. The object A_Φ satisfies:

- Encodes Round: Yes, since all profiles in A_Φ assign Round
- Encodes Square: Yes, since all profiles in A_Φ assign Square
- Exemplifies Round: No, since A_Φ has no global elements
- Exemplifies Square: No, since A_Φ has no global elements

Thus the round square encodes contradictory properties without exemplifying them, avoiding logical contradiction while preserving our ability to reason about impossible objects.

The round square exists as an abstract object (a subobject of E) and we can meaningfully assert that it encodes both roundness and squareness. However, it has no global element—no actual instantiation—preventing any contradiction from arising through exemplification.

5. Consistency and Key Properties

We now establish that our construction preserves the essential features of AOT while maintaining logical consistency.

5.1. **Functoriality and well-definedness.**

Proposition 5.1. *All constructions in our model are functorial and well-defined.*

Proof. We verify each component:

(1) E is a presheaf by Lemma 3.4, with restriction maps preserving the profile structure
(2) Subobjects of presheaves are presheaves, so each $A_\Phi \rightarrowtail E$ is in $\mathcal{Z}$
(3) The encoding relation ε is defined by a simple logical formula at each stage, making it a subobject of $E \times y(\mathbf{P})$

(4) Comprehension uses the universal property of the subobject classifier, which is functorial by definition

Working in an elementary topos guarantees that logical operations are interpreted functorially. ☐

5.2. **Separation of encoding and exemplification.** The following theorem establishes that our model correctly implements Zalta's fundamental distinction:

Theorem 5.2 (Separation)**.** *There exist objects in $\mathcal{Z}$ that encode properties they don't exemplify.*

Proof. The round square $A_{\{\text{Round},\text{Square}\}}$ provides an explicit example. By Theorem 4.1, it encodes both Round and Square. By Proposition 4.5, it has no global elements and hence exemplifies neither property. ☐

This separation is the cornerstone of AOT's ability to handle intentional phenomena. We can think about and reason with impossible objects (they encode properties) without being committed to their existence (they don't exemplify properties).

Example 5.3 (Sherlock Holmes)**.** Let $\Phi_{\text{Holmes}} = \{\text{Detective}, \text{LivesAt221B}, \text{Fictional}, \ldots\}$ be the properties attributed to Holmes in Conan Doyle's stories. Then $A_{\Phi_{\text{Holmes}}}$:

- Encodes all properties in Φ_{Holmes} (by construction)
- Has no global elements (fictional characters don't exist concretely)
- Can consistently encode properties that might be incompatible across different stories

Our framework naturally handles the problem of fictional inconsistency: Holmes can encode incompatible properties from different stories without logical contradiction, since encoding doesn't require consistency at the global level.

5.3. **Identity conditions.**

Proposition 5.4. *Two abstract objects are identical if and only if they encode exactly the same properties.*

Proof. If $A_{\Phi_1} = A_{\Phi_2}$ as subobjects of E, they have the same characteristic morphism, which implies $\Phi_1 = \Phi_2$ by our construction. Conversely, if $\Phi_1 = \Phi_2$, then A_{Φ_1} and A_{Φ_2} have the same characteristic morphism and hence are equal as subobjects. ☐

This validates Zalta's identity criterion for abstract objects and shows our construction preserves the intended individuation of abstracta.

6. Internal Logic Formulation

The internal logic of $\mathcal{Z}$ provides a formal language for reasoning about abstract objects. We now make this explicit.

6.1. **Types and terms.** In the internal logic of $\mathcal{Z}$, we have the following types:

- prop: The type of properties (interpreted as the constant presheaf $\mathbf{P}$)
- obj: The type of objects (interpreted as $\mathrm{Sub}(E)$, the lattice of subobjects of E)
- $\mathsf{profile}$: The type of property profiles (interpreted as E)

6.2. **Axioms in internal logic.** The fundamental axioms of AOT can be expressed in the internal logic as follows:

Definition 6.1 (Encoding axiom)**.**

$$\forall a : \mathsf{obj}.\forall p : \mathsf{prop}.\, \mathsf{encodes}(a, p) \leftrightarrow \forall e : \mathsf{profile}.\, (e \in a \to (e, p) \in \varepsilon)$$

This axiom formally defines encoding as a universal quantification over profiles—an object encodes a property if and only if all its profiles include that property.

Definition 6.2 (Exemplification axiom)**.**

$$\forall a : \mathsf{obj}.\forall p : \mathsf{prop}.\, \mathsf{exemplifies}(a, p) \leftrightarrow \exists \sigma : \mathtt{global}(a).\, (\sigma, p) \in \varepsilon$$

where $\mathtt{global}(a)$ denotes the type of global elements of a.

This axiom requires actual instantiation for exemplification—an object exemplifies a property only if it has a concrete realization possessing that property.

Definition 6.3 (Comprehension schema)**.**

$$\forall \phi : \mathsf{prop} \to \Omega.\, \exists! a : \mathsf{obj}.\, \forall p : \mathsf{prop}.\, \mathsf{encodes}(a, p) \leftrightarrow \phi(p)$$

This schema asserts that every condition on properties determines a unique object encoding exactly those properties satisfying the condition.

Theorem 6.4. *These axioms are valid in the internal logic of $\mathcal{Z}$.*

Proof. The encoding and exemplification axioms are valid by our definitions in Section 3.7. The comprehension schema is valid by Theorem 4.1, extended to arbitrary internal predicates $\phi : \mathsf{prop} \to \Omega$ using the internal logic of the topos. $\square$

6.3. **Logical properties.** The internal logic of $\mathcal{Z}$ is intuitionistic, as is typical for presheaf topoi. This has interesting philosophical implications:

Remark 6.5. While Zalta's original AOT uses classical logic, our intuitionistic setting could be more appropriate for reasoning about abstract objects. The failure of excluded middle corresponds to the idea that some properties of abstract objects may be indeterminate—neither definitely encoded nor definitely not encoded. This aligns with intuitions about vague or incomplete fictional objects.

If classical logic is desired, one can consider the double-negation sheafification or work with Boolean presheaf topoi. However, the intuitionistic setting provides additional expressive power for modeling partial and contextual existence.

7. Application: Ontological Arguments

To demonstrate the philosophical applications of our framework, we analyze ontological arguments for the existence of God. This analysis shows precisely what additional assumptions are needed to move from conceptual to actual existence.

7.1. **Setup.** Following traditional formulations, we consider:

Definition 7.1. Let $D \subseteq \mathbf{P}$ be the set of divine perfections, including:

- Omniscience, omnipotence, moral perfection, etc.
- Necessary existence $N \in D$

We assume D is consistent (no incompatible properties), following Leibniz's requirement that the perfections be compossible.

7.2. **The argument structure.** The ontological argument in our framework proceeds as follows:

Proposition 7.2. *By comprehension (Theorem 4.1), there exists a unique object $G = A_D$ encoding exactly the divine perfections.*

This object G represents the concept of God—by construction, it encodes all perfections including necessary existence. In the internal logic: $\mathsf{encodes}(G, N)$ holds.

However, this alone doesn't establish that God exists. We have only shown that the concept of God includes necessary existence, not that God actually exists. The issue is the gap between encoding and exemplification.

7.3. **The bridging principle.** To complete the ontological argument, one needs an additional principle:

Definition 7.3 (Necessary Existence Bridge)**.**

$$\forall a : \mathsf{obj}.\, \mathsf{encodes}(a, N) \to \exists \sigma : \mathtt{global}(a)$$

"Any object encoding necessary existence must have a global element."

This principle asserts that necessary existence is a special property whose mere encoding guarantees actual instantiation. It's not derivable from our framework—it must be assumed as an additional axiom.

Theorem 7.4. *Given the Necessary Existence Bridge, God exists.*

Proof. From $G = A_D$ and $N \in D$, we have $\mathsf{encodes}(G, N)$. By the bridging principle, there exists $\sigma : 1 \to G$. This global element witnesses that G exemplifies all divine perfections, including existence. □

7.4. **Critical analysis.** AOT and our framework make transparent what is often obscured in ontological arguments:

(1) The comprehension principle alone only establishes that the concept of God is coherent and encodes necessary existence
(2) Moving from encoding to exemplification requires an additional metaphysical principle that is not logically necessary
(3) This principle essentially states that "necessary existence is an existence-entailing property"—precisely what critics of the ontological argument contest

Remark 7.5. The bridging principle is consistent with our framework (assuming D is consistent), but it's a substantial metaphysical assumption. Without it, G remains an abstract object encoding necessary existence without exemplifying it—a "merely possible God" in Meinongian terms.

This analysis doesn't settle the philosophical debate but could clarify its logical structure. The ontological argument requires not just comprehension for abstract objects but a special principle linking the encoding of necessary existence to actual existence.

8. Mathematical Properties and Extensions

8.1. **Relationship to other topoi.** Our construction $\mathcal{Z} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ where $\mathbf{C} = \mathbf{Set}^{\mathbf{P}}$ has some interesting mathematical properties:

Proposition 8.1. *$\mathcal{Z}$ is equivalent to the topos $\mathbf{Set}^{\mathbf{P}\times\mathbf{P}^{\mathrm{op}}}$.*

Proof. Using the equivalence $[\mathbf{Set}^{\mathbf{P}}, \mathbf{Set}]^{\mathrm{op}} \cong \mathbf{Set}^{\mathbf{P}^{\mathrm{op}}}$, we have:

$$\mathcal{Z} = [(\mathbf{Set}^{\mathbf{P}})^{\mathrm{op}}, \mathbf{Set}] \cong \mathbf{Set}^{\mathbf{P}\times\mathbf{P}^{\mathrm{op}}}$$

This shows $\mathcal{Z}$ is itself a presheaf topos over a product category. □

This alternative presentation reveals that our model tracks both covariant and contravariant dependencies on properties, reflecting the dual nature of encoding versus exemplification.

8.2. **Modal extensions.** Following Awodey, Kishida, and Kotzsch (2014), we can extend our framework to include modal operators:

Definition 8.2. A *modal extension* of $\mathcal{Z}$ replaces the subobject classifier Ω with a complete Heyting algebra H and interprets:

- Necessity $\Box$ as a comonad on $\mathcal{Z}$
- Possibility $\Diamond$ as the associated monad

This allows encoding of modal properties like "necessarily round" or "possibly existent" within the same framework. The modal structure interacts naturally with the encoding/exemplification distinction:

- An object might encode $\Box F$ (necessarily has property F in its concept) without exemplifying F in the actual world

- The necessary existence bridge becomes:

$$\mathsf{encodes}(a, \Box \text{Exists}) \to \Diamond \mathsf{exemplifies}(a, \text{Exists})$$

8.3. **Higher-order properties.** Our framework naturally extends to higher-order properties:

Definition 8.3. Define a hierarchy of encoding objects:

- $E^0 = E$ (first-order property profiles)
- $E^{n+1}(X) = \mathbf{P}^{E^n(X)}$ (profiles of n-th order properties)

This allows abstract objects to encode properties of properties, opening applications to:

- Self-referential abstract objects
- Paradox analysis (Russell-type properties)
- Formal semantics for natural language quantification over properties

9. Philosophical Implications

9.1. **Resolving classical puzzles.** Our topos-theoretic model provides solutions to several classical puzzles:

9.1.1. *The problem of non-being.* Non-existent objects exist as abstract objects without global elements. They have being (as subobjects of E) without existence (no global elements), avoiding Meinong's paradox while preserving intuitions about intentional reference.

9.1.2. *Fictional discourse.* Statements about fictional characters become statements about what properties they encode. "Holmes is a detective" is true if and only if the abstract object corresponding to Holmes encodes the property Detective. This provides truth conditions without ontological commitment to fictional entities.

9.1.3. *Impossible objects.* The round square and similar impossibilia are accommodated consistently. They encode contradictory properties (internally to the topos) without exemplifying them (no global elements), maintaining logical consistency while explaining our ability to reason about impossibilities.

9.1.4. *Concrete individuals.* Ordinary objects (those with global instances, encoding nothing) pose no problem in our framework: an object like Balzac can be represented as an abstract object that encodes no property essentially. All of Balzac's properties are exemplified via some global element (his actual instantiation), and indeed by definition such an object encodes none. This is consistent because any property F Balzac has is not encoded (since we can find a profile in which Balzac lacks F), and Balzac's existence is accounted for by a global element in the topos (designating the actual Balzac).

9.2. **Advantages of the topos-theoretic approach.** Our categorical framework offers several advantages over traditional approaches:

(1) **Natural separation**: The encoding/exemplification distinction arises naturally from the difference between internal properties and global elements, not as an ad hoc stipulation
(2) **Logical flexibility**: The internal logic naturally accommodates partial existence, vague objects, and contextual truth
(3) **Geometric intuition**: Topoi provide geometric insights—abstract objects as "spaces of possibilities" varying over contexts
(4) **Systematic variations**: Different choices of **P** and **C** yield different theories of abstract objects, enabling systematic exploration of alternatives

9.3. **Comparison with set-theoretic approaches.** Traditional set-theoretic formalizations of AOT face several challenges our approach avoids:

- Set theory's extensionality conflicts with intentional phenomena
- The cumulative hierarchy imposes unwanted structure on abstract objects
- Classical two-valued logic forces artificial precision on vague objects
- Set-theoretic models typically require encoding as an additional primitive relation

In contrast, our topos model derives the encoding/exemplification distinction from the categorical structure itself, provides natural models for partial and contextual existence, and connects to established work in categorical logic and semantics.

10. Conclusion

We have presented a categorical formalization of Abstract Object Theory using topos theory. Our construction demonstrates that the philosophical insights of AOT—particularly the distinction between encoding and exemplification—align naturally with fundamental mathematical structures in category theory.

The main idea is recognizing that presheaf topoi provide exactly the right setting for theories of abstract objects. The separation between internal properties (what holds at each stage) and global elements (what exists absolutely) corresponds precisely to Zalta's distinction between encoding and exemplification. This is not merely a formal analogy but reveals that topos theory provides the natural mathematical home for intentional phenomena.

Our framework achieves several goals:

- Provides a consistent treatment of impossible objects without ad hoc restrictions
- Validates unrestricted comprehension for abstract objects while avoiding paradox

- Clarifies the logical structure of ontological and other philosophical arguments

Perhaps more importantly, this work shows the value of applying sophisticated mathematical tools to philosophical problems. Just as set theory revolutionized the foundations of mathematics in the 20th century, category theory—and topos theory in particular—may provide crucial insights for 21st-century philosophy. The work presented here shows that deep philosophical distinctions can correspond to elegant mathematical structures, suggesting that further categorical investigations of philosophical concepts may prove equally fruitful.

The theory developed here could provide philosophical insight. By showing that the encoding/exemplification distinction arises naturally from the structure of presheaf topoi, we gain confidence that this distinction tracks a fundamental feature of abstract objects rather than being an arbitrary formal device. The fact that impossible objects, fictional characters, and mathematical abstracta all find natural homes in our framework suggests we might have identified the right level of mathematical abstraction for modeling intentionality.

We hope this work inspires further research at the intersection of category theory and philosophy. Indeed, the right mathematics can illuminate philosophical problems in unexpected ways.

References

Awodey, Steve, Kohei Kishida, and Hans-Christoph Kotzsch (2014). "Topos Semantics for Higher-Order Modal Logic". In: *Logique et Analyse* 228, pp. 591–636.

Kirchner, Daniel, Christoph Benzmüller, and Edward N. Zalta (2020). "Mechanizing Principia Logico-Metaphysica in Functional Type Theory". In: *Review of Symbolic Logic* 13.1, pp. 206–218.

Mac Lane, Saunders and Ieke Moerdijk (1992). *Sheaves in Geometry and Logic: A First Introduction to Topos Theory.* Universitext. Springer-Verlag.

Zalta, Edward N. (1983). *Abstract Objects: An Introduction to Axiomatic Metaphysics.* Vol. 160. Synthese Library. D. Reidel.

– (1988). *Intensional Logic and the Metaphysics of Intentionality.* MIT Press.

– (2019). "Abstract Objects". In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Spring 2019. URL: `https://plato.stanford.edu/archives/spr2019/entries/abstract-objects/`.