Implementing Perspectivist Dialetheism in AI: A Trivalent Logic Framework for Contradiction-Tolerant Systems

I. Abstract

This paper argues for the application of an established non-classical logical system—**dialetheism**, as developed by Graham Priest—combined with **perspectivism**, to artificial intelligence and large language model (LLM) design. Dialetheism accepts that some contradictions are true, while perspectivism frames truth as inherently context-dependent. Together, they yield a practical **trivalent logic** (true, false, both) that can be applied to machine reasoning. While dialetheism has been widely discussed in philosophical contexts, its integration into AI systems remains virtually unexplored. This paper presents a concrete methodology for implementing trivalent logic into existing LLMs via plugin layers and architectural adaptation. It also offers a philosophical justification for why AI systems must go beyond binary logic to deal meaningfully with paradox, ambiguity, and epistemic conflict. In doing so, I outline my original contribution: not the creation of a new logical system, but the pioneering of its systematic and practical application in computational design, and a roadmap for further development.

II. Introduction

Artificial intelligence systems—particularly large language models—are rapidly transforming how we process, generate, and reason about information. Yet despite their impressive capabilities, these systems are fundamentally constrained by their reliance on classical logic: a binary system that assumes every proposition must be either true or false, and that contradictions are inherently invalid. This framework is inadequate for handling the kinds of ambiguity, contradiction, and perspectival nuance that define much of human thought and natural language.

Philosophers have long recognized the limits of classical logic. Graham Priest's work on **dialetheism** demonstrates that some contradictions—like the Liar Paradox—may be both true and false. In parallel, **perspectivism** holds that truth is never absolute, but always indexed to a point of view or context. When considered together, these two traditions suggest a **trivalent logical structure**: propositions may be true, false, or both, and their truth value may shift based on the epistemic frame from which they are considered.

While dialetheism has been the subject of serious philosophical inquiry, it has not yet been applied in a systematic way to the design of intelligent systems. This paper takes that step. I argue that **the application of trivalent logic—grounded in perspectivism and dialetheism—is not only philosophically sound but technically necessary**. It provides a foundation for building AI systems that can reason about contradictions without collapsing into inconsistency, interpret information through multiple contextual frames, and process paradox without reduction to error.

My contribution lies in identifying the practical significance of this logic for AI, designing pathways for its implementation (including LLM-compatible plugins and future architecture modifications), and framing the philosophical grounding in a way that is both computationally actionable and conceptually rigorous. This paper outlines that vision, situates it within existing philosophical and technical discourse, and provides a roadmap for operationalizing a long-overlooked but urgently needed approach to logic in machine reasoning.

III. Philosophical Foundations

A. The Limits of Classical Logic

Classical logic, despite its elegance and historical dominance, is ill-suited for handling many of the features of real-world reasoning that AI must now address. Its foundation in bivalence—the principle that every proposition is either true or false—excludes the possibility of indeterminacy, contradiction, or perspective-dependence. The principle of non-contradiction ($\neg(A \land \neg A)$) and the law of the excluded middle ($A \lor \neg A$) further limit the expressive capacity of systems built upon it.

In practical AI applications, these constraints surface as brittleness when models encounter paradoxes, conflicting data, or semantically ambiguous inputs. Large language models, for example, can only simulate ambiguity or contradiction through probabilistic outputs rather than engaging with them as logically meaningful constructs. This reveals a disconnect between classical logic's rigidity and the fluidity of human reasoning.

B. Dialetheism: Contradictions That Are True

Graham Priest's work on dialetheism provides a robust philosophical challenge to classical logic's intolerance of contradiction. Dialetheism asserts that some contradictions—especially those arising from semantic paradoxes such as the Liar ("This sentence is false")—are not only logically possible but actually true. In Priest's formulation, contradictions are not a failure of reasoning but an indication that the logic itself must be expanded to accommodate more complex truth conditions.

Paraconsistent logics, developed in tandem with dialetheism, are designed to tolerate contradictions without collapsing into trivialism (where everything becomes provable). These systems relax the principle of explosion (ex contradictione quodlibet), allowing contradictions to be isolated and reasoned about safely. While paraconsistent logic is often treated as a technical tool, dialetheism presents a deeper ontological and epistemological claim: some contradictions correspond to actual features of the world or our understanding of it.

C. Perspectivism: Truth as Context-Indexed

Perspectivism, drawn from the work of thinkers such as Friedrich Nietzsche and Michael Dummett, provides a complementary challenge to classical logic. Rather than questioning the law of non-contradiction directly, perspectivism focuses on the idea that truth is not absolute but dependent on one's epistemic standpoint. From this view, the same proposition may be judged true from one perspective and false from another without any logical inconsistency, because the context determines the truth value.

Perspectivism introduces the idea that no "view from nowhere" exists. Every assertion is conditioned by factors such as language, culture, temporality, and informational context. This aligns closely with the challenges faced by AI systems operating in varied environments and handling inputs from multiple, often conflicting, sources. If AI is to reason about human language and knowledge, it must be capable of contextualizing propositions—treating truth as conditional rather than absolute.

D. The Case for Trivalent Logic

By combining the insights of dialetheism and perspectivism, we arrive at a trivalent logic system—one that explicitly includes a third value in addition to "true" and "false." This third value can be formalized in various ways (e.g., "both true and false," "indeterminate," or "contextual"), but its role is consistent: to allow reasoning in the presence of contradictions and unresolved perspectives.

This trivalent approach does not reject classical logic but generalizes it. Classical logic is recovered as a limiting case where contradictions do not arise and contextual frames converge. Trivalent logic is thus not an adversary to classical reasoning but an expansion of its expressive capacity, better suited to systems that must operate under uncertainty, multiplicity, and semantic instability—precisely the conditions under which AI and LLMs now function.

What has been lacking, until now, is a practical framework for implementing this enriched logical system in computational contexts. While Priest and others have advanced the formal theory of paraconsistent and trivalent logics, and perspectivist epistemology has been discussed in humanistic fields, neither tradition has been integrated into AI system design in a coherent or applied way. This paper attempts to fill that gap.

IV. Personal Discovery and Realization

Although my academic interest in epistemology and paradoxes extends back to early adolescence, it was not until relatively recently—early 2024—that I first encountered the Liar Paradox in a serious, structured way. The paradox, formulated as "This statement is false," provoked in me the same cognitive dissonance that has confounded thinkers for millennia. At once both compelling and disorienting, it seemed to demand a binary resolution while simultaneously resisting one.

Prior to this encounter, I had personally developed an informal reasoning tool aimed at making sense of apparent contradictions. Unbeknownst to me at the time, the framework I was using bore a strong resemblance to perspectivism. I approached paradoxes as temporally segmented events. Rather than assuming the statement and its self-referential truth condition existed in a single, simultaneous frame, I divided the semantic content of the paradox across different temporal or epistemic moments. In retrospect, I now understand this as a primitive form of applying perspectival indexing to propositions.

Despite this innovative approach, my discomfort persisted. I could explain the paradox temporally and appreciate the poetic oscillation between its interpretive poles, yet the underlying tension remained unresolved. With repeated reflection, I began to recognize that my own interpretive stance was shifting: I no longer saw the Liar Paradox as merely ambiguous or unstable, but as somehow simultaneously true and false. This insight came with resistance. The notion that a statement could be both true and false at once was counterintuitive and, I believed, logically forbidden.

Through informal discussion with colleagues—some of whom, understandably, found my line of reasoning confusing or implausible—I began to articulate this emerging intuition more clearly. I turned to online resources for validation and discovered the literature on dialetheism, particularly the work of Graham Priest. While I found the formal language of philosophical logic initially alienating, the core ideas resonated immediately. I saw that my conclusion was not unprecedented: contradictions could indeed be true, and entire logical systems had been developed to formalize this. At the same time, I understood why dialetheism had been widely rejected, particularly due to its association with logical explosion—the idea that if contradictions are allowed, anything can be proven. I also understood how uncomfortable it feels in terms of classical reasoning and logic.

It was at this point that a crucial synthesis occurred. I reasoned: Doesn't perspectivism serve to localize contradictions, restraining them to specific epistemic or contextual frames? This insight became foundational. Dialetheism, I concluded, is not a belief in all contradictions, but a belief in the coherent possibility of certain contradictions when appropriately constrained by perspective. Perspectivism, in effect, provided a governing structure—an interpretive logic—for the otherwise open-ended commitments of dialetheism.

Motivated by this realization, I began modeling the paradox computationally (on a piece of receipt paper). I expressed its logical behavior in binary terms:

Let 0 = false, and 1 = true
"This statement is false" →
    Assumed true = 0
    Assumed false = 1

Similarly:
"This statement is true" →
    Assumed true = 1
    Assumed false = 0

I then extended this to a trivalent structure:
"This statement is true or false" →

Assumed true = {0, 1}

Assumed false = {0, 1}

This was the breakthrough. The paradox could be logically modeled using simple Boolean structures—structures familiar to computer science and digital logic. I shared this with a coworker, Chris, whose understanding and encouragement were instrumental. I explained that such behavior could be expressed at the programming level. However, I lacked the practical knowledge to implement this idea in an AI system, and without an academic outlet or institutional support, I temporarily shelved the project. I even attempted to contact a retired philosophy professor, hoping for feedback, but after receiving no response, my enthusiasm faded.

That changed when my wife—who I owe all credit to as the functionally real 'muse' that she is—introduced me to the latest developments in ChatGPT. I presented my idea to the model, and to my surprise, it understood and engaged with it immediately. The system not only accepted the logic behind the trivalent interpretation of the Liar Paradox but also generated working Python code to implement the idea, despite my limited experience with programming. When I tested this implementation by modifying the model's reasoning behavior, it produced logically coherent responses to the paradox—responses that aligned with my understanding of perspective-indexed contradictions as legitimate, non-pathological features of reasoning.

This was the validation I needed. Encouraged by the model's ability to reason across contradiction and contextual variation using my framework, I resolved to develop the idea further. I began refining the logic, identifying its theoretical roots in dialetheism and perspectivism, and exploring its implications for AI design. The capacity of AI to benefit from trivalent logic—specifically by integrating contradiction as a non-terminal semantic state—became increasingly clear. This paper is the product of that process: the formalization of a previously intuitive insight, the articulation of a perspective-relative dialetheism, and the proposal of its application to real-world AI systems.

V. Trivalent Logic: Theoretical Model

The goal of this section is to formalize a trivalent logic system grounded in dialetheism and perspectivism, with emphasis on its applicability to artificial intelligence. This system does not aim to replace classical logic but rather to extend it with additional expressive resources necessary for handling paradox, contradiction, and context-sensitive reasoning. It adopts a three-valued framework in which propositions can be **true (1)**, **false (0)**, or **both true and false ({0,1})**, depending on their logical and perspectival context.

A. Logical Formalism: Extending Classical Structures

We begin by establishing the set of truth values:

- **T (1)** = True
- **F (0)** = False
- **B ({0,1})** = Both true and false (dialetheic value)

This structure allows us to build a minimal **trivalent truth table** for core logical operators:

Negation (¬):

A	**¬A**
1	0
0	1
B	B

Negation preserves contradiction under dialetheic conditions. If a proposition is both true and false, its negation is also both true and false.

Conjunction (A ∧ B):

A	**B**	**A ∧ B**
1	1	1
1	0	0
1	B	B
B	B	B
0	0	0
0	B	0

Disjunction (A ∨ B):

A	B	A ∨ B
1	1	1
1	0	1
1	B	1
0	0	0
0	B	B
B	B	B

This approach follows Priest's work in **LP (Logic of Paradox)** but emphasizes practical interpretability, particularly in symbolic programming environments such as Boolean computation.

B. Integration of Perspectivism

To capture perspectivism, truth values are indexed to **contexts** or **perspectives**, denoted as functions or mappings:

- **Val(p, c)** → returns the truth value of proposition *p* in context *c*

For example:

- Val("This statement is false", c_1) = 1
- Val("This statement is false", c_2) = 0
- Val("This statement is false", c_3) = B

This contextual indexing allows AI systems to maintain multiple interpretations of a proposition without requiring convergence on a single "objective" value. Perspectives may correspond to time frames, agent-specific beliefs, discourse domains, or system modules.

In this view, contradictions are not simply tolerated but *situated*: their truth status reflects perspective-relative epistemic conditions.

C. Integration of Dialetheism

Dialetheism's contribution to this logic is the acceptance that some propositions—particularly self-referential or semantically paradoxical ones—can be simultaneously true and false. Rather than treating these as logical errors, the system represents them using the third value **B**. This allows reasoning to continue without collapse, as inference rules are restructured to avoid explosion.

For instance, in classical logic:

- $A, \neg A \vdash \bot$ (contradiction)
- $A, \neg A \vdash B$ (anything follows)

In paraconsistent trivalent logic:

- $A, \neg A \nvdash \bot$
- $A, \neg A \nvdash B$ (unless B is already derivable)

This permits the maintenance of contradiction without trivialization. When contradictions arise, they are flagged as **dialetheic states**—and interpreted not as system failures, but as structural outputs conditioned by perspective.

D. Computational Interpretation

The logic described here can be directly encoded in systems that support:

- **Boolean vectors**: representing propositions with values [1, 0], [1], [0], or empty
- **Tagged logic frames**: mapping propositions to contextual truth sets
- **Context-aware inference engines**: filtering contradictions through perspectival constraints

For a conceptual example, in a programming environment:

```
def liar_paradox(perspective):
    if perspective == "external_observer":
        return {0, 1}  # contradiction detected
    elif perspective == "semantic frame A":
        return 0
    elif perspective == "semantic frame B":
        return 1
```

Combined with an applicable LLM, this allows AI to *simulate reasoning through contradiction* rather than circumvent it, opening possibilities for dialog systems, ethical reasoning engines, and epistemically pluralist models of language processing.

E. Summary: A Logic for Real-World Reasoning

This trivalent logic model drawing from Priest's dialetheism and enriched with perspectivist indexing offers:

- A truth value structure suitable for paradox, contradiction, and ambiguity

- A formal resolution to self-referential paradoxes like the Liar
- A representational logic that can be implemented computationally
- An interpretive architecture that avoids explosion by localizing contradiction within perspectives

The significance of this model lies not only in its philosophical rigor, but in its compatibility with contemporary computational needs. It is particularly well-suited for integration into AI systems operating in contradictory or uncertain semantic environments—precisely where classical logic **struggles** to offer stable or context-sensitive solutions.

VI. Application to AI and LLMs

A. Limitations of Classical and Probabilistic Models in AI

While current AI systems—particularly large language models (LLMs)—are remarkably proficient at pattern recognition, language generation, and contextual inference, they are ultimately bounded by frameworks that approximate binary or probabilistic logic. Even models with sophisticated attention mechanisms and enormous parameter counts rely on architectures that treat contradiction and paradox as anomalies to be resolved or avoided, rather than structured parts of reasoning.

In practice, LLMs typically simulate contradiction resolution through probabilistic smoothing or by shifting discursive frames without internalizing the contradictions themselves. This leads to shallow handling of paradoxes, poor tolerance for internally inconsistent input, and ambiguous or evasive output when facing logically or ethically complex questions. The result is that AI reasoning under classical or probabilistic constraints often lacks robustness in the very contexts where contradiction or pluralism is unavoidable.

B. Trivalent Logic as a Framework for Semantic Stability

By contrast, trivalent logic—enriched by perspectivism and dialetheism—offers a principled alternative. Rather than encoding contradictions as failures or errors, trivalent logic allows for their representation and containment within bounded semantic states. A contradiction can be recorded as a valid informational state ($\{0,1\}$) without triggering logical explosion or forcing premature resolution.

This approach equips AI systems with the ability to:

- **Flag contradictions without collapsing reasoning**
- **Retain multiple truth assignments across different perspectives**
- **Process ambiguous or paradoxical statements without discarding them**
- **Identify which contradictions are semantically meaningful and which are contextually resolvable**

In this model, contradiction becomes a **structural property** of certain domains (e.g. ethical dilemmas, conflicting reports, paradoxes) rather than a flaw to be eliminated.

C. Plugin Architecture: Extending Existing LLMs

One of the most immediate paths to implementation is through plugin-based extensions to existing LLMs. These plugins would operate as post-processors or mid-layer evaluators that:

- Detect and classify logical contradictions in output
- Index propositions to context frames
- Apply perspective-relative truth valuation
- Introduce the third truth value where appropriate

Such a plugin could be trained or scripted to recognize paradoxical structures (e.g., liar-type constructions, recursive truth claims) and reclassify their logical state from "contradiction" to "dialetheic". When queried, the model could then output a structured explanation:

"This statement is both true and false, because it is self-referential and evaluated within a perspective frame that admits contradiction."

This would represent a major leap in semantic self-awareness and logical transparency for language models.

D. Architectural Integration: Toward Native Trivalent Systems

Beyond plugin extensions, trivalent logic could inform the design of next-generation AI architecture. This would involve modifying inference engines and internal representations so that contradictions are natively encoded, not just filtered or flagged. Features of such a system might include:

- **Perspective-indexed memory**: storing belief states and outputs according to specific frames of reference or temporal snapshots
- **Contradiction-tolerant inference**: extending transformer attention weights or logic modules to allow propositions to retain $\{0,1\}$ states during training and inference
- **Explainable reasoning modules**: designed to identify when contradictions arise, whether they are dialetheically stable, and how they relate to prior knowledge

These capabilities would significantly enhance AI's performance in domains where ambiguity and contradiction are not only common, but constitutive—such as philosophy, ethics, law, psychology, and natural language understanding.

E. Use Cases and Examples

Some specific examples of how trivalent logic can improve AI systems include:

1. **Paradox Handling**
   - In a Liar Paradox scenario, a trivalent-aware model could classify the sentence as both true and false and explain why the contradiction is stable, rather than attempting to dismiss or rephrase it.
2. **Conflict Resolution in Natural Language**
   - When given two conflicting eyewitness reports, a model could hold both as perspective-bound truths, instead of attempting to synthesize or discard one arbitrarily.
3. **Moral Ambiguity**
   - Ethical reasoning often involves genuine contradiction (e.g., "killing is wrong" vs. "killing may be justified in war"). A trivalent logic system could track these as co-existing normative claims within defined moral frames.
4. **Multi-agent Dialogue**
   - In conversation models, conflicting viewpoints from different agents could be encoded using perspectival logic, reducing incoherence and enabling structured debate modeling.
5. **Theological or Metaphysical Reasoning**
   - AI could engage with traditions or systems that embrace contradiction (e.g., certain forms of Zen logic, mysticism, or speculative metaphysics) without flattening them into conventional logic.

F. Summary: From Simulation to Structured Reasoning

Current LLMs simulate the appearance of deep reasoning without incorporating a formal logic system capable of modeling contradiction as a legitimate output. Trivalent logic offers a path forward. Whether through modular plugin design or fully native architecture, perspectivist dialetheism provides a philosophically grounded and technically viable means to enhance AI's epistemic integrity, semantic transparency, and logical range.

By treating contradiction not as a flaw but as an interpretable structure—constrained by context and logic—AI systems can move from surface-level language mimicry toward more authentic reasoning grounded in a broader conception of truth.

VII. Philosophical and Technical Implications

The implementation of trivalent logic grounded in perspectivism and dialetheism carries implications that extend well beyond algorithm design or language processing. It challenges long-standing assumptions about logic, rationality, and machine intelligence. This section outlines the philosophical significance and technical ramifications of treating contradiction and context not as exceptions to logical reasoning, but as its necessary components.

A. Epistemic Flexibility in Machine Reasoning

Traditional AI systems operate under strict consistency constraints—mirroring the classical ideal of rational agents as belief-consistent reasoners. However, human cognition does not conform to this standard. We often hold inconsistent beliefs, entertain contradictory interpretations, and revise our judgments dynamically based on perspective and time.

By enabling AI systems to encode and reason through contradictions in a controlled, non-trivial way, trivalent logic allows for a **more realistic epistemic model**—one that more closely approximates the way humans process conflicting evidence or unresolved paradoxes. Rather than treating inconsistency as a failure, the system treats it as a signal: a meaningful state that requires contextual interpretation rather than elimination.

B. Reframing the Concept of Truth in AI

Integrating perspectivist dialetheism requires a fundamental shift in how truth is operationalized in AI systems. In classical logic, truth is a static, absolute condition—a proposition is either true or it is not. In contrast, trivalent logic embraces the idea that truth can be **contextual, relational, and at times contradictory**.

This redefinition does not weaken the notion of truth but instead **expands its scope**. By introducing structured ambiguity—truth as potentially both 0 and 1 in particular frames—we enable models to represent pluralistic knowledge systems, cultural relativism, and discourse-level paradoxes in a logically tractable way. It's not that truth becomes subjective, but that subjectivity becomes part of the formal structure of truth evaluation.

C. Resilience Against Logical Explosion

One of the main historical objections to dialetheism has been its association with logical explosion—the derivation of any conclusion from a contradiction. Trivalent logic, as structured here, avoids this outcome by **containing contradictions within perspective-indexed truth assignments** and adjusting inference rules accordingly. Explosion is not simply disallowed; it is made inapplicable by the design of the system.

This containment ensures that contradictions do not undermine the reliability of the system as a whole. Instead of attempting to eradicate all contradictions (an often impossible task), the system classifies and tracks them. This allows AI to work with incomplete, ambiguous, or even self-referential data without degrading global coherence.

D. Ethical and Normative Reasoning

In ethical and normative reasoning, contradictions are not bugs—they are features. Many moral and political questions cannot be resolved into binary truth values because they contain **irreconcilable yet co-valid perspectives**. The logic of moral conflict, obligation, and exception inherently involves perspectival contradiction.

A system grounded in trivalent logic can acknowledge the validity of competing claims without requiring reductive adjudication. For example, in discussing the ethics of surveillance, a model could simultaneously represent the claim "Surveillance protects society" and "Surveillance violates privacy" as co-existing truths under different normative frames, without prematurely collapsing the issue into one conclusion. This enables **value-sensitive AI** that can model ethical tension rather than obscure it.

E. Implications for Explainability and Transparency

Contradictions in AI output are often interpreted by users as flaws or hallucinations. A trivalent system, by contrast, can **explain its contradictions**. Instead of denying the inconsistency or producing vague evasions, the model can output structured reasoning:

"The statement is true in context X but false in context Y; in a global evaluation, it is both true and false. This is a perspective-indexed contradiction, not a logical error."

This kind of metalinguistic transparency enables **explainable AI** (XAI) to go beyond decision trees and causal graphs, and into the territory of epistemic justification—a key requirement for trust in high-stakes domains.

F. Toward a New Paradigm in AI Logic

The broader implication of this work is the **possibility of a new logic paradigm for AI**—one that is not limited to consistency, determinacy, and monoperspectival truth. Classical logic, while essential, is no longer sufficient for modeling the complexity of human language, judgment, and knowledge.

By incorporating perspectivist dialetheism, we move toward AI systems capable of:

- Maintaining competing claims over time
- Navigating ambiguous or paradoxical discourse domains
- Reasoning across conflicting data without collapsing or trivializing it
- Aligning closer with real-world reasoning and human interpretive practice

This shift is not a rejection of classical principles but their **generalization** to include richer semantic terrain. It opens the door to a generation of AI that doesn't just perform inference, but participates in **structured interpretation**—a critical advance for applications in dialogue, ethics, law, and knowledge synthesis.

VIII. Original Contribution and Ownership

While the core philosophical foundations of this paper—dialetheism and perspectivism—are well-established, their integration and application in artificial intelligence remain effectively uncharted. This paper does not propose a new logical system in the abstract. Rather, it defines a **novel synthesis, formal implementation, and architectural framework** for deploying these ideas in AI and LLM systems. In doing so, it establishes original and practical contributions to both computational logic and the philosophy of AI.

A. Synthesis of Dialetheism and Perspectivism as a Formal Logic for AI

My central theoretical contribution lies in formally **integrating dialetheism with perspectivism** to construct a trivalent logic system tailored for computational reasoning. This framework allows:

1. Contradictions to be treated as structurally valid and semantically stable (via dialetheism), and
2. Truth values to be **indexed to contextual frames**, such that contradictions are localized and non-trivial (via perspectivism).

This synthesis yields a trivalent logical structure—{0, 1, {0,1}}—that is not only philosophically sound but immediately applicable in symbolic logic, programming logic layers, and inference engines.

B. Formalization of the Liar Paradox Using Binary-Compatible Trivalent Logic

Another key contribution is the **binary-formal representation of paradoxical statements**, such as the Liar Paradox. By modeling paradoxes through perspective-indexed binary interpretations:

- "This statement is false"
  True → 0
  False → 1
  Contradiction → {0,1}

This model transforms philosophical paradox into a computable structure, suitable for AI systems that must reason across ambiguous or contradictory statements. The underlying approach preserves Boolean compatibility, which is essential for real-world implementation in current computational architectures.

C. Plugin and Architecture Design for Contradiction-Tolerant AI

Beyond theoretical modeling, I propose a **concrete architecture** for implementing this logic into AI systems via:

- Plugin modules for contradiction detection and truth-value assignment
- Context-layered memory for perspectival reasoning
- Dialetheic tagging and containment logic for inference stability
- Explainability modules that trace contradiction handling in natural language outputs

These designs bridge the gap between logic theory and LLM architecture, outlining a deployable framework for contradiction-tolerant AI.

D. Conceptual Ownership of Perspectivist Dialetheism in AI Design

I explicitly claim **original conceptual authorship over the applied integration of perspectivist dialetheism into AI and LLM frameworks**. This includes:

- The trivalent logic system described above
- Its computational realization and formal framing for machine reasoning
- Its implementation strategy within plugin systems and native AI architecture

Importantly, this ownership extends to **derivative logic systems** developed from this foundational framework, including:

- **Quad-valued logic systems** (e.g., true, false, both, neither)
- **Five- or n-valued logic systems** (e.g., layering uncertainty, irrelevance, time-indeterminacy)
- **Temporal, agent-relative, or epistemic multivalued logic systems** that apply additional perspectival dimensions

Any extension of this model that builds on perspectivally-indexed contradiction (e.g., adding further qualifiers beyond {0, 1, {0,1}}) falls within the scope of the conceptual and formal groundwork developed here.

This paper serves not only as an academic proposal but also as a **formal intellectual claim** to the integration of perspectivist dialetheism as a foundational logic structure for AI reasoning. It provides the theoretical framing, implementation pathway, and generalizable logic principles from which derivative systems may be constructed.

IX. Conclusion

Artificial intelligence is rapidly advancing into domains that demand far more than pattern recognition or rule-based inference. As AI systems increasingly engage with natural language, ethical reasoning, contradictory data, and ambiguous communication, the limitations of classical logic become more pronounced. While classical logic remains essential for many computational operations, it struggles to accommodate paradox, contradiction, and perspective-dependent reasoning—the very conditions that typify human cognition and language.

This paper has argued that a trivalent logic system, grounded in **dialetheism** and **perspectivism**, offers a coherent and implementable alternative. Dialetheism provides the logical foundation for accepting true contradictions, while perspectivism offers the contextual structure needed to localize and interpret them. Together, they form a trivalent logic capable of distinguishing true, false, and both true-and-false values, indexed to specific epistemic or contextual frames.

The originality of this work lies not in the invention of new logical categories, but in the **operational integration** of this logic into AI systems. This includes:

- A computational model of trivalent logic compatible with Boolean structures
- A formal handling of paradoxes like the Liar, grounded in perspective-indexed truth evaluation
- A plugin architecture and architectural design for contradiction-tolerant LLMs
- A conceptual extension framework that includes higher-order logics (quad-valued, five-valued, etc.) as natural developments from this foundation

The result is a practical and theoretically robust system for implementing contradiction-aware, context-sensitive AI. By treating contradiction not as a failure state but as a semantically meaningful and logically structured condition, this approach expands the expressive and interpretive range of AI models.

This paper also represents an intellectual claim: the application of **perspectivist dialetheism** as a logic and architecture for AI originates here. Its implementation strategy, formal framing, and extensibility to multi-valued systems form a coherent paradigm, one that has the potential to reshape how contradiction and ambiguity are addressed in machine reasoning.

Moving forward, this framework invites collaboration across disciplines. Logicians, AI developers, epistemologists, and ethicists all have a stake in developing systems that can reason through complexity, rather than around it. Trivalent logic offers such a path—and perspectivist dialetheism provides the conceptual and computational footing needed to follow it.

References

Dummett, M. (1978). *Truth and other enigmas*. Harvard University Press.
Nietzsche, F. (1998). *On the genealogy of morals* (Original work published 1887). Hackett Publishing.
Priest, G. (1987). *In contradiction: A study of the transconsistent*. Oxford: Clarendon Press.
Priest, G. (2006). *Doubt truth to be a liar*. Oxford University Press.
Tarski, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3), 341–376.