

INSIGHTS

Catuṣkoṭi and Non-Binary Reasoning in Artificial Intelligence

CHHEM Kieth Rethy¹, MD, PhD (Edu), and PhD (His)

¹ *CHHEM Kieth Rethy is a Distinguished Professor of AI and Humanities, Center for Buddhist AI Ethics, Cambodia University of Science and Technology and Advisor to AI Forum*

Introduction

In the expanding field of Artificial Intelligence (AI), logic forms the foundational structure for computation, inference, and decision-making. Classical logic, heavily influenced by Aristotle's system of binary oppositions, predicates reasoning on the principle of the excluded middle: every proposition must be either true or false, but not both or neither. However, in real-world scenarios, particularly those involving ethics, cultural interpretation, and natural language, this binary framework can be too rigid.

Enter *catuṣkoṭi*, or the Tetralemma, a logical structure rooted in Nāgārjuna's Madhyamaka tradition of Buddhist philosophy. *Catuṣkoṭi* posits four possibilities for any given proposition: (1) it is true, (2) it is false, (3) it is both true and false, and (4) it is neither true nor false. This framework challenges Aristotelian logic and offers a richer paradigm for modeling ambiguity and paradox, which are crucial features of human cognition and ethical judgment. This chapter explores the relevance of *catuṣkoṭi* for AI design, particularly in natural language processing (NLP) and machine ethics, where classical logics often reach their limits.

Aristotelian Logic and Its Limits

Aristotle's contribution to formal logic, especially through his *Organon*, established the dominant paradigm in Western thought. The principle of non-contradiction (a statement cannot be both true and false at the same time) and the law of the excluded middle (a statement must be either true or false) have become axioms in classical logic systems, including those used in programming languages and most AI architectures (Smith, 2003). While this system sustains much of computational logic, it struggles in domains where ambiguity, contradiction, or incompleteness is not a bug but a feature.

Consider an AI tasked with moral decision-making in autonomous vehicles. When faced with an ethical dilemma such as choosing between harming one pedestrian or multiple passengers, classical binary logic cannot always provide a satisfying answer. There may be multiple overlapping truths or no clear truth at all. Similarly, in NLP tasks like sentiment analysis, a sentence can convey conflicting emotions such as both joy and sorrow, which resists binary classification.

The *Catuṣkoṭi* Framework

The *catuṣkoṭi* framework originates from Nāgārjuna's *Mūlamadhyamakakārikā*, a foundational text of the Madhyamaka school of Mahāyāna Buddhism. The logic operates on four poles:

1. P (a proposition is true)
2. $\neg P$ (the proposition is false)
3. $P \wedge \neg P$ (the proposition is both true and false)
4. $\neg(P \vee \neg P)$ (the proposition is neither true nor false)

Rather than affirming a reality behind propositions, *catuṣkoṭi* emphasizes the limits of conceptualization and encourages a deconstruction of absolute beliefs (Garfield, 1995). It challenges the ontological assumptions that underpin binary logic, replacing determinacy with interdependence and emptiness (*śūnyatā*). In computational terms, this opens up a space for multi-valued logic systems that can accommodate contradictory, incomplete, or culturally contingent inputs.

This kind of non-binary approach also aligns with mindfulness-based cognitive flexibility. As Anālayo (2021) notes, the tetralemma reflects an awareness of complexity and conditionality that extends beyond rigid dualities, fostering insight into intuitive awareness.

Application in Natural Language Processing

NLP provides fertile ground for implementing *catuṣkoṭi*-inspired models. Languages are inherently fuzzy and context-dependent. A single utterance can carry irony, sarcasm, ambiguity, or layered meanings that defy binary classification.

Take the sentence: “I’m fine.” In classical sentiment analysis, the model might categorize it as positive. However, depending on intonation, context, or cultural norms, it could signal distress, resignation, or even anger. A *catuṣkoṭi*-based NLP model could allow the following classifications simultaneously:

- True (positive sentiment)
- False (actually negative sentiment)
- Both (mixed emotions)
- Neither (neutral, ambiguous, or contextually unclear)

Such a system could use probabilistic weights or fuzzy logic to rank each corner of the tetralemma, enabling more nuanced interpretation and response. Research in non-classical logic already points in this direction (Beziau, 2012). *Catuṣkoṭi* offers a philosophically grounded and culturally enriched logic to enhance these approaches.

Ethical Reasoning in AI

AI ethics, especially in high-stakes fields like healthcare, warfare, and autonomous transport, demands more than deterministic computation. It requires interpretative judgment. Traditional logic falters when dilemmas cannot be resolved with right or wrong binaries.

Consider AI in medical triage. Should the AI prioritize the young over the elderly? Or the person with a better prognosis over one with greater social responsibility? Such decisions involve values, contradictions, and cultural expectations that vary significantly. A *catuṣkoṭi* model could accommodate ethical pluralism. It might allow an AI to recognize that a decision is simultaneously justifiable and unjustifiable, depending on the moral framework invoked. Instead of resolving the contradiction, the AI would surface it, flagging the ethical ambiguity for human oversight or adaptive learning. This aligns with post-normal science, which calls for decision-making under conditions of uncertainty, complexity, and value conflict (Funtowicz & Ravetz, 1993).

Toward Culturally Sensitive AI

One major criticism of generative AI development is its epistemological bias. Many AI systems are trained and modeled using Western logic, datasets, and moral assumptions. Integrating *catuṣkoṭi* into AI design contributes to the broader project of epistemic pluralism. It not only diversifies computational models but also invites cross-cultural perspectives on what it means to think, decide, and act.

For Southeast Asian cultures influenced by Buddhism or for Indian traditions shaped by Nyāya and Jain logic, reasoning is not necessarily exclusionary. By embedding non-binary frameworks like *catuṣkoṭi*, AI can better resonate with the lived ethical and epistemic realities of these societies. This has implications for education, diplomacy, and cross-border AI regulation.

Conclusion

Catuṣkoṭi offers an elegant yet radical alternative to the binary logic that dominates AI today. Rooted in Buddhist philosophy and refined through centuries of epistemological reflection, the tetralemma provides a powerful framework for modeling ambiguity, contradiction, and ethical complexity. Its integration into AI, especially in NLP and ethical decision-making, can lead to systems that are not only technically sophisticated but also culturally respectful and philosophically enriched.

As AI expands into every domain of human life, drawing on non-Western logics like *catuṣkoṭi* is not merely an academic exercise. It is a necessary step toward truly inclusive, ethical, and intelligent machines.

References

- Anālayo, B. (2022). Beyond the limitations of binary thinking: Mindfulness and the tetralemma. *Mindfulness*, 13, 1410–1417.
- Béziau, J. -Y. (2011). A new four-valued approach to modal logic. *Logique et Analyse*, 54(213), 109-121.
- Funtowicz, S. O., & Ravetz, J. R. (1993). Science for the post-normal age. *Futures*, 25(7), 739- 755.
- Garfield, J. L. (1995). *The fundamental wisdom of the middle way: Nāgārjuna's Mūlamadhyamakakārikā*. Oxford University Press.
- Smith, R. (2022). Aristotle's logic. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2022 ed.). Metaphysics Research Lab, Stanford University.

** Edited with ChatGPT for clarity; all views expressed are solely those of the authors.*

** The opinions expressed are the authors' own and do not reflect the views of AI Forum.*