

Mapping the Mind: A Ten-Dimensional Coordinate System for Classifying Philosophies of Artificial Intelligence

Di Zhang^{1*}

¹*School of AI and Advanced Computing, Xi'an Jiaotong-Liverpool University, No. 111, Taicang Avenue, Suzhou, 215400, Jiangsu, China PR.

Corresponding author(s). E-mail(s): di.zhang@xjtlu.edu.cn;

Abstract

The field of Artificial Intelligence (AI) is characterized by profound philosophical disagreements that transcend technical implementation and strike at fundamental questions about the nature of mind, intelligence, and knowledge. While traditional taxonomies such as “symbolic vs. connectionist” or “strong vs. weak AI” provide useful heuristics, they often obscure the complex multidimensional nature of philosophical commitments in AI research. This paper proposes a comprehensive analytical framework—the *Ten-Dimensional Coordinate System for AI Philosophies*—to systematically map the philosophical landscape of AI. Drawing on contemporary debates in philosophy of mind, cognitive science, and AI ethics, we identify ten fundamental dimensions spanning ontology, epistemology, methodology, functional architecture, core competency, learning paradigm, consciousness, development, ethics, and cosmology. We demonstrate how this framework can precisely characterize major AI research programs (e.g., GOFAI, connectionism, embodied AI, functionalism) and contemporary developments like large language models. The coordinate system reveals deep philosophical affinities and tensions that cut across traditional disciplinary boundaries, provides a diagnostic tool for analyzing interdisciplinary conflicts, and serves as a generative framework for identifying underexplored research directions. By offering a more nuanced vocabulary for philosophical positioning, this framework advances the discourse in AI philosophy beyond reductive dichotomies and provides essential conceptual scaffolding for navigating the complex ethical and conceptual challenges of artificial intelligence.

Keywords: Philosophy of Artificial Intelligence, Taxonomy, Symbolicism, Connectionism, Embodied Cognition, Functionalism, Strong AI, Philosophy of Mind, AI Ethics, Machine Learning

1 Introduction: Beyond Binary Thinking in AI Philosophy

The philosophical foundations of Artificial Intelligence have been contested terrain since the field’s inception at the Dartmouth Conference in 1956. While technical progress in machine learning has been explosive, the philosophical discourse surrounding AI has often remained trapped in what [14] critiqued as a series of false dichotomies: **mind vs. body**, **reason vs. perception**, **symbolic vs. subsymbolic**, and **biological vs. artificial**. These binary oppositions, while historically productive,

have increasingly proven inadequate for characterizing the complex philosophical positions that underlie contemporary AI research.

The limitations of these traditional taxonomies become particularly evident when examining hybrid approaches like neuro-symbolic AI [4], embodied cognitive systems [30], or the philosophical implications of large language models [3]. As [6] notes, the philosophical questions raised by AI are not merely academic but have profound implications for how we design, evaluate, and govern intelligent systems.

This paper addresses this conceptual gap by introducing the *Ten-Dimensional Coordinate System for AI Philosophies*. Our central thesis is that the philosophical commitments underlying different approaches to AI can be systematically characterized through their positions along ten relatively independent dimensions. This framework builds on earlier multidimensional analyses of cognitive science [45] while extending them to address specifically AI-related philosophical concerns, including recent debates about superintelligence [7], AI ethics [16], and the nature of machine consciousness [34].

Unlike previous taxonomies that primarily focus on technical implementation, our coordinate system emphasizes the underlying philosophical commitments that often implicitly guide research programs. By providing a more nuanced mapping of the conceptual space, we aim to facilitate more productive dialogue between different AI research communities and provide a framework for systematically exploring the philosophical implications of new AI developments.

2 Theoretical Foundations and Related Work

2.1 Historical Taxonomies in AI and Cognitive Science

The classic tripartite division of AI into **symbolicism, connectionism, and behaviorism** [8] has been remarkably durable, but it primarily categorizes approaches by their computational mechanisms rather than their philosophical foundations. Similarly, the distinction between **strong AI** (machines can truly think and understand) and **weak AI** (machines can simulate thinking) [37] focuses on a single philosophical dimension while neglecting others.

[26] unified theories of cognition attempted to bridge these divides but remained largely within the symbolic paradigm. More recently, [48] has proposed a framework for understanding different conceptions of intelligence, but his analysis focuses more on cognitive capabilities than philosophical foundations.

2.2 Philosophical Dimensions in Contemporary AI Research

Recent work in AI philosophy has highlighted the need for more sophisticated conceptual frameworks. [11] emphasizes the importance of priors and embodiment in intelligence, touching on what we categorize as the Ontology and Learning Paradigm dimensions. [21] argument for building machines that learn and think like people integrates concerns across multiple dimensions, particularly Core Competency and Developmental View.

The ethical dimensions of AI have also received increasing attention. [16] proposes a philosophy of technology that addresses the ethical status of AI systems, relevant to our Ethics and Value Alignment dimension. Similarly, work on AI safety and alignment [36] engages with questions about the ultimate potential and controllability of AI systems, connecting to our Cosmology dimension.

Our ten-dimensional framework synthesizes these diverse strands of inquiry into a unified analytical tool that captures the full spectrum of philosophical commitments in contemporary AI research.

3 The Ten-Dimensional Coordinate System

Each dimension in our coordinate system represents a spectrum between two contrasting philosophical positions on a fundamental aspect of intelligence or AI. The dimensions are designed to be relatively orthogonal, allowing for independent variation, though some natural affinities exist between certain positions.

3.1 Ontology (X1): Formal vs. Embodied Intelligence

This dimension concerns the fundamental nature of intelligence. The **Formal** pole, exemplified by [33] computational theory of mind, views intelligence as an abstract information-processing capacity that can be realized in any sufficiently powerful computational system. This position is closely associated with functionalism in philosophy of mind [32] and underlies much of classical AI.

The **Embodied** pole, drawing from phenomenological philosophy [46] and embodied cognition [39], contends that intelligence is fundamentally shaped by and dependent on physical embodiment. Proponents argue that cognition emerges from the dynamic interaction between an organism and its environment [10], and that abstract intelligence divorced from bodily experience is impossible or severely limited.

3.2 Epistemology (X2): Symbolic vs. Connectionist Knowledge

This dimension addresses how knowledge is represented and processed. **Symbolicism**, the foundation of Good Old-Fashioned AI (GOFAI) [27], views knowledge as explicit, discrete symbols that can be manipulated according to formal rules. This approach emphasizes compositionality, systematicity, and explainability.

Connectionism, revitalized by the parallel distributed processing approach [35] and contemporary deep learning [22], represents knowledge as patterns of activation distributed across neural networks. This approach emphasizes gradient-based learning, fault tolerance, and similarity-based reasoning.

3.3 Methodology (X3): Engineering vs. Simulation

This dimension distinguishes between AI as engineering and AI as cognitive science. The **Engineering** approach, exemplified by [8] behavior-based robotics, focuses on building systems that work effectively, regardless of whether their mechanisms resemble human cognition. The motto “wheels, not wings” captures this pragmatic orientation.

The **Simulation** approach aims to replicate human cognitive processes to understand the mind [43]. This tradition includes cognitive architectures like ACT-R [2] and neural models that seek biological plausibility.

3.4 Functional Architecture (X4): Modular vs. Holistic

This dimension concerns the structural organization of intelligent systems. The **Modular** view, influenced by [17] modularity of mind hypothesis, posits that intelligence consists of specialized, domain-specific processing modules. This view is evident in systems with separate components for perception, reasoning, and action.

The **Holistic** view, associated with connectionist approaches and unified cognitive architectures [25], sees intelligence as emerging from highly integrated, interactive processing systems where knowledge and capabilities transfer across domains.

3.5 Core Competency (X5): Logical Reasoning vs. Pattern Recognition

This dimension identifies the fundamental operation considered most essential to intelligence. The **Logical Reasoning** view, central to the logicist tradition in AI [24], sees intelligence as fundamentally involving abstract inference and symbolic manipulation.

The **Pattern Recognition** view, emphasized in statistical learning approaches [19], sees intelligence as fundamentally involving the detection of statistical regularities in data and generalization from examples.

3.6 Learning Paradigm (X6): Top-Down vs. Bottom-Up

This dimension concerns how knowledge is acquired. **Top-Down** learning involves explicit instruction, programming, or knowledge engineering, where structured information is provided to the system [15].

Bottom-Up learning involves knowledge construction through experience with data or environment [31]. This includes both supervised learning from labeled examples and reinforcement learning through environmental interaction [44].

3.7 Consciousness and Intentionality (X7): Functionalism vs. Biological Naturalism

This dimension addresses the hard problems of consciousness and meaning. **Functionalism**, the dominant view in philosophy of mind and classical AI [5], holds that mental states are defined by their causal roles and can be realized in multiple physical substrates.

Biological Naturalism, defended by [38], insists that consciousness and intentionality are biological phenomena specific to certain biological systems, particularly human brains, and cannot be replicated in digital computers.

3.8 Developmental View (X8): Designism vs. Evolutionism

This dimension concerns the path to advanced intelligence. **Designism** believes intelligence can be directly engineered through careful design and programming [41].

Evolutionism argues that genuine intelligence requires extended developmental processes involving interaction with rich environments [23], possibly through evolutionary algorithms [42] or cumulative cultural learning.

3.9 Ethics and Value Alignment (X9): Instrumentalism vs. Subjectivism

This dimension addresses the moral status of AI systems. **Instrumentalism** views AI as tools that derive their value entirely from human purposes [9].

Subjectivism holds that sufficiently advanced AI systems could become moral patients or even moral agents with interests and rights that must be considered [13].

3.10 Cosmology (X10): Carbon Chauvinism vs. Cosmic Computationalism

This dimension concerns the ultimate potential and place of intelligence in the universe. **Carbon Chauvinism** expresses skepticism about the possibility of machine intelligence matching or exceeding human capabilities [29].

Cosmic Computationalism views intelligence as a fundamental property that can emerge in various physical systems and potentially exceed human capabilities by many orders of magnitude [7, 20].

4 Mapping Major AI Research Programs

Applying our coordinate system reveals the complex philosophical profiles of major AI approaches (see Table 1).

Table 1 Philosophical Profiles of Major AI Research Programs

Research Program	Pro-	Philosophical Profile
GOFAI		Strongly?? Formal (X1) , Symbolic (X2) , Logical Reasoning (X5) , Top-Down (X6) . Typically Modular (X4) , Functional (X7) , Designist (X8) , Instrumentalist (X9) .
Connectionism		Core commitment to Connectionist (X2) , Pattern Recognition (X5) , Bottom-Up (X6) . Generally Formal (X1) , Holistic (X4) , Engineering-oriented (X3) , Instrumentalist (X9) .
Embodied AI		Defined by Embodied (X1) . Typically allied with Connectionist (X2) , Bottom-Up (X6) , Evolutionist (X8) , and often sympathetic to Biological Naturalism (X7) .
Neuro-Symbolic AI		Attempts to bridge Symbolic (X2) and Connectionist (X2) positions. Typically maintains a Formal (X1) ontology while integrating Logical Reasoning (X5) and Pattern Recognition (X5) capabilities.
Artificial General Intelligence		Emphasizes Holistic (X4) architecture and Evolutionist (X8) development. Often combines Formal (X1) ontology with ambitions toward Cosmic Computationalism (X10) .

5 Contemporary Applications and Case Studies

5.1 Large Language Models: A Multidimensional Analysis

The philosophical implications of large language models (LLMs) like GPT-4 [28] become clearer through our coordinate system:

- **X1 (Formal)**: LLMs operate purely in linguistic space, entirely disembodied from physical reality.
- **X2 (Connectionist)**: Their architecture is fundamentally based on artificial neural networks, specifically the Transformer [47].
- **X3 (Engineering)**: The primary goal is practical utility rather than cognitive simulation, though they incidentally model some aspects of human language.
- **X5 (Pattern Recognition)**: Their core operation is statistical pattern prediction, though they can simulate logical reasoning.
- **X7 (Functionalism)**: The intense debate about whether LLMs “understand” language reflects the tension between functional and biological conceptions of intentionality.

This analysis explains why LLMs generate both excitement and skepticism: they produce functionally intelligent behavior while lacking several features traditionally associated with intelligence (embodiment, symbolic manipulation, conscious understanding).

5.2 AI Safety and Alignment Research

AI safety research [1] also exhibits distinct philosophical profiles. Work on technical alignment tends toward **Formal (X1)** and **Instrumentalist (X9)** positions, viewing AI as powerful tools that must remain under human control. In contrast, some discussions of AI rights and moral patienthood [40] explore **Subjectivist (X9)** positions.

6 Theoretical Contributions and Utility

6.1 Diagnosing Interdisciplinary Conflicts

Our coordinate system helps explain why certain interdisciplinary collaborations in AI research prove difficult. For example, the tension between classical AI and embodied robotics often reflects opposition along multiple dimensions: X1 (Formal vs. Embodied), X2 (Symbolic vs. Connectionist), and X6 (Top-Down vs. Bottom-Up). Recognizing these fundamental philosophical differences can help researchers either find productive compromises or understand why certain integrations may be philosophically incoherent.

6.2 Generating Novel Research Directions

The framework systematically identifies underexplored regions of the philosophical space. Promising research directions include:

- **Embodied Symbolic Systems (X1-Embodied + X2-Symbolic):** Can symbolic reasoning be grounded in physical interaction? This approach might address the symbol grounding problem [18] through embodied cognition rather than statistical correlation.
- **Conscious AI Ethics (X7-Functionalism + X9-Subjectivism):** If we take functionalism about consciousness seriously, what ethical obligations might we have toward AI systems that exhibit consciousness-like properties?
- **Evolutionary Design (X8-Evolutionism + X3-Engineering):** Can we engineer environments that foster the open-ended evolution of intelligence, blurring the distinction between design and growth?

7 Limitations and Future Directions

While comprehensive, our framework has limitations. The dimensions, though designed for relative independence, exhibit some natural correlations (e.g., Embodied positions often align with Connectionist and Evolutionist views). The framework also primarily addresses philosophical foundations rather than technical implementation details.

Future work could develop quantitative measures for positions along each dimension, allowing for more precise comparison of research programs. The framework could also be extended to incorporate sociotechnical dimensions addressing how AI development is shaped by economic, political, and cultural factors [12].

8 Conclusion

The Ten-Dimensional Coordinate System for AI Philosophies provides a sophisticated conceptual toolkit for navigating the complex philosophical landscape of artificial intelligence. By moving beyond reductive dichotomies, this framework enables more precise characterization of philosophical commitments, helps diagnose interdisciplinary conflicts, and suggests novel research directions at the intersections of different philosophical positions.

As AI continues to advance and permeate society, clear philosophical thinking about its foundations becomes increasingly crucial. Our coordinate system contributes to this endeavor by providing a structured vocabulary for articulating and debating the deep philosophical questions that artificial intelligence raises about mind, intelligence, and our place in the universe.

References

References

- [1] Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., and Man, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.
- [2] Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51(4), 355.
- [3] Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623).
- [4] Besold, T. R., d'Avila Garcez, A., Bader, S., Bowman, H., Domingos, P., Hitzler, P., ... and de Penning, L. (2017). Neural-symbolic learning and reasoning: A survey and interpretation. *arXiv preprint arXiv:1711.03902*.
- [5] Block, N. (1980). What is functionalism? In N. Block (Ed.), *Readings in philosophy of psychology* (Vol. 1, pp. 171-184). Harvard University Press.
- [6] Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.
- [7] Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- [8] Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, 47(1-3), 139-159.
- [9] Brynjolfsson, E., and McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton and Company.
- [10] Chemero, A. (2009). *Radical embodied cognitive science*. MIT press.
- [11] Chollet, F. (2019). On the measure of intelligence. *arXiv preprint arXiv:1911.01547*.
- [12] Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- [13] Danaher, J. (2020). Welcoming robots into the moral circle: A defence of ethical behaviourism. *Science and Engineering Ethics*, 26(4), 2023-2049.
- [14] Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. Harper and Row.
- [15] Feigenbaum, E. A. (1977). The art of artificial intelligence: Themes and case studies of knowledge engineering. In *Proceedings of the 5th international joint conference on Artificial intelligence*

(pp. 1014-1029).

- [16] Floridi, L. (2021). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- [17] Fodor, J. A. (1983). *The modularity of mind*. MIT press.
- [18] Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3), 335-346.
- [19] Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science and Business Media.
- [20] Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin.
- [21] Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40.
- [22] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- [23] Lungarella, M., Metta, G., Pfeifer, R., and Sandini, G. (2003). Developmental robotics: a survey. *Connection Science*, 15(4), 151-190.
- [24] McCarthy, J. (1959). Programs with common sense. In *Proceedings of the Teddington Conference on the Mechanization of Thought Processes* (pp. 75-91).
- [25] Minsky, M. (1986). *The society of mind*. Simon and Schuster.
- [26] Newell, A. (1981). The knowledge level. *AI magazine*, 2(2), 1-1.
- [27] Newell, A., and Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3), 113-126.
- [28] OpenAI (2023). GPT-4 Technical Report. *arXiv preprint arXiv:2303.08774*.
- [29] Penrose, R. (1989). *The emperor's new mind: Concerning computers, minds, and the laws of physics*. Oxford University Press.
- [30] Pfeifer, R., and Bongard, J. (2007). *How the body shapes the way we think: a new view of intelligence*. MIT press.
- [31] Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press.
- [32] Putnam, H. (1967). Psychological predicates. In W. H. Capitan and D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37-48). University of Pittsburgh Press.
- [33] Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. MIT Press.
- [34] Reggia, J. A. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, 44, 112-131.
- [35] Rumelhart, D. E., McClelland, J. L., and PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). MIT press.
- [36] Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Penguin.
- [37] Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424.
- [38] Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.
- [39] Shapiro, L. (2011). *Embodied cognition*. Routledge.

- [40] Shipley, W., and Goldfarb, E. (2021). The artificial intelligence of the ethics of artificial intelligence: An introductory overview for law and regulation. In *The Oxford Handbook of Ethics of AI*.
- [41] Simon, H. A. (1996). *The sciences of the artificial*. MIT press.
- [42] Stanley, K. O., and Lehman, J. (2019). *Why greatness cannot be planned: The myth of the objective*. Springer.
- [43] Sun, R. (2008). *The Cambridge handbook of computational psychology*. Cambridge University Press.
- [44] Sutton, R. S., and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- [45] Thagard, P. (2005). *Mind: Introduction to cognitive science*. MIT press.
- [46] Varela, F. J., Thompson, E., and Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT press.
- [47] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... and Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).
- [48] Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, 10(2), 1-37.