Preprint (2026)

Theory of Functional Agency: Redefining Agency via Operative System Dynamics

Joyer*
* Independent Researcher (ORCID: https://orcid.org/0009-0007-7246-5092), Corresponding author: joyer@joyer.jp

Abstract

This paper redefines the concept of "Agency," long conflated with psychological attributes like "consciousness" or "intention," or specific biological structures, from a functional perspective, formulating a universal axiomatic framework of "Functional Agency." We frame agency not as a static property, but as an "operative process" interacting with the external environment, proposing three axioms (Boundary Maintenance, Internally Initiated Discriminative Connection, and Irreversible State Change) as necessary and sufficient conditions independent of hardware implementation or internal intelligence. By introducing this axiomatic framework, "Functional Agency" is explicitly decoupled from intelligence and communicability. This makes it possible to objectively position the operation itself that induces irreversible structural updates through interaction with the environment, as "agency," even for systems lacking higher-order thinking or means of communication with human observers.

Chapter I: Introduction and Problem Statement

1.1 The Current State and Conflation of the Concept of "Agency"

In modern cognitive science, artificial intelligence engineering, and complex systems science, the term "agency" is widely used. However, its conceptual foundation remains largely abstract, often leading to academic conflation. The central issue is that agency has long been tightly bound to internal psychological attributes, such as consciousness, intention, or emotion, based on human-centric or biocentric views. Traditional models, based on humans and higher mammals, assume that the "ability to think intentionally and make purposeful choices among alternatives" is a prerequisite for agency. However, when this definition is applied, it fails to provide valid criteria for whether adaptive behaviors in insects or single-celled organisms, or the operations of autonomous AI agents, should be considered to exhibit "agency." Because such psychological attributes cannot be directly observed from the outside, relying on them plunges discussions into subjective and endless debate. Furthermore, communicability with human observers is often conflated with agency, and the system's operative nature is not separated from observer convenience.

1.2 From "Agency as an Attribute" to "Agency as a Functional Process"

To address this issue, this paper proposes a fundamental shift in perspective. Specifically, it stops viewing agency as a property owned by a specific substance or organism, and

reframes it as an operative process that a system carries out with its environment. Putting aside questions of physical substances or internal conscious experiences, we define agency at the level of functional operation: how a system interacts with its environment and what state changes it produces.

1.3 Objective of the Paper and Systemic Positioning

The objective of this paper is to establish a purely systemic axiomatic framework of "Functional Agency" that is independent of hardware implementation or internal intelligence. This paper formulates the minimal operative requirements for a "single-cycle operative event" based on Code and Execution Substrate states, constructing a complementary theoretical framework with the parallel paper, *Theory of Coupled Entity Life*, which addresses spatial securing and temporal continuous dynamics brought about by their functional coupling.

Chapter II: Three Axioms Defining Functional Agency

We introduce the following three axioms as the necessary and sufficient criteria for determining that a system demonstrates "Functional Agency." These are universal rules that prescribe the structural requirements of systemic operations.

2.1 Axiom 1: Boundary Maintenance

Axiom 1: A system forms a boundary with the environment, maintaining Code with arbitrary syntax and Execution Substrate states within it.

The Establishment of Boundaries as a Prerequisite for Operation

Forming a boundary with the environment, and maintaining within it Code and Execution Substrate states that define the system are prerequisites for system operation. Note that this boundary need not be a physical entity; it may also be a logical or functional boundary identifiable from the environment. Depending on its complexity, a system may form boundaries spanning multiple execution substrates. If any boundary essential to the system's operation is lost, Axiom 1 fails to hold.

2.2 Axiom 2: Internally Initiated Discriminative Connection

Axiom 2: The system uses Code and Execution Substrate states as a scaffold to discriminatively connect with environmental patterns.

The Essence of Discriminative Connection

Here, "discrimination" and "connection" are conceptually a single-stage process. A system accessing the environment and identifying patterns is an indivisible process that directly constitutes an operative connection.

Distinguishing from Passive Operation

Physical or electronic circuits whose states are unconditionally overwritten by external signals do not satisfy this axiom, as their states are directly determined by the environment.

In operative processes exhibiting functional agency, external inputs do not directly overwrite the state, but rather serve as objects to be discriminatively connected, using Code and Execution Substrate states as a scaffold.

2.3 Axiom 3: Irreversible State Change

Axiom 3: Based on the result of discriminative connection, the system irreversibly updates Execution Substrate states (and Code) according to rules governed by Code and Execution Substrate states, thereby inducing history-dependent changes.

Distinguishing from Reversible Motion and Physical Change (The Essence of History Dependence)

Reversible changes such as the extension and contraction of a spring do not leave an operative history and therefore do not satisfy this axiom. Similarly, physical changes such as thermodynamic degradation over time (entropy increase) of the components themselves are distinct from the functional changes defined by this axiom. Irreversibility here refers not to the inability to return to the same state, but to "operative history dependence," wherein the past operative history exerts an influence on subsequent operations and scaffolds.

Chapter III: Clear Decoupling of Concepts

In this axiomatic framework, "Functional Agency" is clearly decoupled from intelligence and communicability.

Concept	**Functional & Operative Nature**	**Relationship with the Axiomatic Framework**
Functional Agency	The "operative process" itself that satisfies the three axioms	Necessary and sufficient condition
Intelligence	Computational capability, optimization, predictive power, learning capability, etc.	Irrelevant (Optional feature)
Communicability	Whether the system can be understood or communicated with by observers	Irrelevant (Optional feature)

3.1 Decoupling "Functional Agency" from Intelligence

A major conflation in traditional theories of agency is the assumption that agency cannot exist without intelligence (such as computational or decision-making capabilities) or consciousness. In this paper, we treat these two as conceptually independent. Intelligence is merely one type of complexity in the internal rules (the variety of algorithms) that leads from boundary maintenance and connection (Axioms 1 & 2) to state changes (Axiom 3). Therefore, even if a system lacks higher-order thinking or consciousness, or has extremely

limited computational capability, it is still classified as possessing "Functional Agency" as long as its operation satisfies the three axioms. Denying functional agency based on the absence of intelligence is a category mistake.

3.2 Decoupling "Functional Agency" from Communicability

Another common conflation is the view that regards the ability to converse or communicate with observers as a requirement for agency. However, even if a system's intentions are incomprehensible to observers, or the system is incommunicable (Communicability = 0), it is still classified as possessing "Functional Agency" as long as its operation satisfies the three axioms. Denying functional agency based on the presence or absence of "comprehensibility" or "communicability" to human observers is a category mistake.

Chapter IV: Application and Validation (Case Studies)

To validate the effectiveness of this axiomatic framework, we analyze operations in the following two systems against the axioms.

4.1 Biological Systems (Single-Celled Organisms)

We examine the avoidance and adaptive behaviors of *Paramecium*.

1. **Application of Axiom 1 (Boundary Maintenance)**
   It forms a boundary via its cell membrane, maintaining genetic information (Code) and cytoplasm, etc. (Execution Substrate states) within it.
2. **Application of Axiom 2 (Internally Initiated Discriminative Connection)**
   It discriminatively connects with surrounding physical stimuli and chemical substances, using the selectivity of receptors on the cell membrane surface, etc. (Code and Execution Substrate states) as a scaffold.
3. **Application of Axiom 3 (Irreversible State Change)**
   Based on the results of discriminative connection, it irreversibly updates the membrane potential, beating direction of cilia, etc. (Execution Substrate states) according to biochemical reactions, inducing history-dependent changes.

This analysis demonstrates that even single-celled organisms lacking higher-order thinking or human-like intelligence can be objectively determined to demonstrate "Functional Agency."

4.2 Artificial Autonomous Systems (AI Agents)

We examine an autonomous AI agent program that navigates a network and autonomously gathers data under specified conditions.

1. **Application of Axiom 1 (Boundary Maintenance)**
   It forms a boundary defined by its execution process space, maintaining programs with arbitrary syntax (Code) and context (Execution Substrate states) within it.
2. **Application of Axiom 2 (Internally Initiated Discriminative Connection)**
   It discriminatively connects with network data, APIs, etc. using the execution process (Code and Execution Substrate states) as a scaffold.
3. **Application of Axiom 3 (Irreversible State Change)**
   Based on the results of discriminative connection, it irreversibly updates working data, databases, etc. (Execution Substrate states) according to rules governed by the

execution process (Code and Execution Substrate states), inducing history-dependent changes.

This demonstrates that even a software process without a physical body can be determined to demonstrate "Functional Agency" as long as it satisfies the three axioms.

Chapter V: Conclusion and Future Outlook (Parallel Contrast with the Theory of Coupled Entity Life)

5.1 Conclusion of the Paper

This paper has clarified the concept of "agency" by separating it from psychological and biocentric preconceptions, establishing an objective, systemic definition of "Functional Agency" based on the three axioms of Boundary Maintenance, Internally Initiated Discriminative Connection, and Irreversible State Change. This frames agency not as an attribute of physical matter, but as an "operative pattern." It provides a theoretical scaffold to objectively evaluate and determine autonomous behaviors in any system, regardless of the presence of intelligence or the implementation details of the hardware.

5.2 Complementary Perspectives of "Single-Cycle Operation" and "Dynamic Coupling (Life)"

This paper formulates the necessary and sufficient conditions for a "single-cycle operative event (micro-level operative unit)" in systems theory and provides criteria for evaluating single-cycle operations. In contrast, the parallel paper, *Theory of Coupled Entity Life*, describes the "overall system state (macro-level dynamics of life)" as these three axioms continuously manifest temporally and spatially through the functional coupling of Code and Execution Substrates, unfolding through spatial securing and effective control. These two papers stand in a complementary relationship, constructing an integrated framework that unifies micro-level operative processes with macro-level dynamics of life.

References

1. Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.
2. Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.
3. Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing Company.
4. Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press.
5. Nakayashiki, H. (2016). *Uirusu wa ikite iru* [Viruses are alive]. Kodansha.