# Viable Action without State Transitions: A Social-Affordance-Centered Formal Model with Memory-Based Semantics

Y. Matsuda and ChatGPT5.2

January 20, 2026

**Abstract**

Classical models of robotic action rely on explicit internal states and their transitions. However, such formulations face fundamental limitations in human-interactive environments, where action viability depends not only on physical feasibility but also on social acceptability, history, and contextual memory. This paper proposes a formal model of viable robotic action centered on *physical affordance* and *social affordance*, while explicitly rejecting the notion of state as a transitionable entity. Instead, we redefine state as a reconstructed bundle of conditions, memories, and constraints that enable viable action. To operationalize this view, we introduce a memory-based semantic operator, referred to as CognitiveRAG, which retrieves context from interaction history and biases affordance evaluation via affection-related parameters. The resulting framework provides a mathematically grounded yet implementation-oriented alternative to state-transition-based robotics, suitable for human-robot interaction under social and emotional constraints.

## 1 Introduction

Robotic action has traditionally been modeled through explicit internal states and state transition functions. In such formulations, an action is selected as an output of a state-dependent policy, and the internal state is updated according to a transition rule. While effective in closed or fully observable environments, this paradigm becomes increasingly fragile in human-interactive settings, where the viability of an action cannot be determined solely from instantaneous physical conditions.

In interactions involving humans, actions are constrained not only by geometry and dynamics but also by social norms, perceived intentions, emotional reactions, and accumulated interaction history. These constraints are often implicit, context-dependent, and history-sensitive, making them difficult to encode as discrete or continuous state variables subject to deterministic or stochastic transitions.

In contemporary robotics, this tension has led to a gradual but decisive shift: state is no longer regarded as a faithful representation of the world, nor as a transitionable entity that governs action selection. Rather, action viability emerges from a complex interplay of conditions, memories, and constraints that are reconstructed at decision time.

This paper formalizes this shift by proposing a model in which *viable action* is defined through affordance relations and memory-based semantics, without introducing an explicit state transition equation. We focus on two complementary forms of affordance: physical affordance and social affordance, and show how their joint evaluation, biased by retrieved interaction history, constitutes the basis for viable robotic action.

# 2 Affordance-Centered Action Modeling

## 2.1 Physical Affordance

Let $A_{\text{phys}}$ denote the set of physical affordances available to the robot in a given situation. Physical affordances characterize what actions are physically feasible, safe, or executable under the laws of mechanics, kinematics, and dynamics.

Formally, for an action candidate $a$ and a belief representation $b_t$, we define a physical affordance evaluation as a predicate or score

$$A_{\text{phys}}(a \mid b_t), \tag{1}$$

which encodes constraints such as reachability, stability, collision avoidance, and energy limits. Importantly, $b_t$ is not interpreted as a full system state, but as a belief or world model updated from observations.

## 2.2 Social Affordance

In human-robot interaction, physical feasibility alone is insufficient. Actions must also be socially acceptable, interpretable, and non-threatening.

We therefore introduce *social affordance*, denoted $A_{\text{soc}}$, which captures constraints arising from social norms, human expectations, and affective responses.

Social affordance is evaluated as

$$A_{\text{soc}}(a \mid r_t, \theta_t), \tag{2}$$

where $r_t$ is context retrieved from interaction history, and $\theta_t$ is an affection-related parameter encoding models of human affect, trust, or interaction risk.

**Terminological note.** Although we adopt the term *social affordance* to emphasize the relational and normative nature of these constraints, the parameterization uses affection-related variables. This does not imply that the robot possesses emotions; rather, affection serves as a latent parameterization of social interaction effects, including emotional reactions implicitly exchanged during interaction.

# 3 Memory-Based Semantics and CognitiveRAG

## 3.1 Interaction History and Retrieval

Let the interaction history be represented as a memory store containing tuples $(o_t, a_t, y_t)$, where $o_t$ denotes observations, $a_t$ executed actions, and $y_t$ observed outcomes, including human responses.

From this memory, a retrieval operation produces a context vector

$$r_t = \text{retrieve}(q_t, \theta_t), \tag{3}$$

where $q_t$ is a query derived from the current belief $b_t$ and possibly recent observations, and $\theta_t$ biases retrieval according to social or affection-related considerations.

We refer to this retrieval mechanism as *CognitiveRAG*, not as a database component, but as a semantic operator that reconstructs context relevant to action viability.

## 3.2 State as Reconstructed Constraint Bundle

Crucially, this model does not introduce a state variable $s_t$ nor a transition equation of the form $s_{t+1} = f(s_t, \cdot)$. Instead, what would traditionally be called "state" is reconstructed implicitly through:

- belief representations $b_t$,
- retrieved context $r_t$,
- affordance evaluations $A_{\text{phys}}$ and $A_{\text{soc}}$,
- and viability constraints defined below.

In this sense, state is not an entity that evolves over time, but a bundle of conditions, memories, and constraints assembled at decision time to evaluate action viability.

# 4 Viable Action without State Transitions

## 4.1 Viability Predicate

We define a viability predicate $\mathcal{V}$ over action candidates:

$$\mathcal{V}(a \mid b_t, r_t) = \mathcal{V}_{\text{phys}}(A_{\text{phys}}(a \mid b_t)) \wedge \mathcal{V}_{\text{soc}}(A_{\text{soc}}(a \mid r_t, \theta_t)). \tag{4}$$

An action $a_t$ is selected if and only if $\mathcal{V}(a_t \mid b_t, r_t)$ holds. No state transition is computed; instead, the execution of $a_t$ produces a new interaction outcome $y_t$, which is logged to memory and may influence future retrieval.

## 4.2 Temporal Structure without Transitions

Temporal dependence in this model arises solely from:

- memory accumulation,
- retrieval bias via $\theta_t$,
- and belief updates driven by observations.

There is no requirement for a Markovian state, nor for a predefined transition structure. Action viability is evaluated relationally and historically, rather than through predictive state evolution.

# 5 Discussion

The proposed model reframes the notion of state in robotic action. Rather than treating state as a privileged variable governing transitions, we treat it as an emergent construct reconstructed from memory and constraints.

This perspective aligns with contemporary trends in robotics and AI, where long-horizon interaction, social context, and adaptive behavior cannot be reduced to finite or even continuous state spaces.

By embedding CognitiveRAG as a semantic operator within the action selection process, the model bridges formal affordance-based reasoning and practical, memory-driven implementations. Social affordance, parameterized through affection-related variables, allows the robot to account for emotional and normative aspects of interaction without attributing internal emotions to the robot itself.

# 6 Conclusion

We presented a formal model of viable robotic action that abandons state transitions in favor of affordance-centered, memory-based semantics. By jointly evaluating physical and social affordances, and reconstructing action-relevant context through CognitiveRAG, the model captures essential aspects of human-robot interaction that elude classical state-based approaches. This framework offers a principled foundation for implementing socially viable robotic behavior in complex, history-dependent environments.

# A Implementation Notes: Probabilistic Viability and Learning of $\theta_t$

This appendix provides implementation-oriented refinements of the formal model presented in the main text. The purpose is not to prescribe a specific algorithm, but to clarify how the proposed framework can be instantiated in practical robotic systems while preserving its semantic commitments.

## A.1 Probabilistic Extension of the Viability Predicate

In the main formulation, viability was expressed as a Boolean predicate

$$\mathcal{V}(a \mid b_t, r_t) = \mathcal{V}_{\text{phys}} \wedge \mathcal{V}_{\text{soc}}. \tag{5}$$

For implementation in uncertain and noisy environments, it is often preferable to replace this hard predicate with a probabilistic viability score.

We therefore introduce a probabilistic viability function

$$P_{\text{viable}}(a \mid b_t, r_t) \in [0, 1], \tag{6}$$

defined as a composition of physical and social components:

$$P_{\text{viable}}(a \mid b_t, r_t) = P_{\text{phys}}(a \mid b_t) \cdot P_{\text{soc}}(a \mid r_t, \theta_t). \tag{7}$$

Here,

- $P_{\text{phys}}(a \mid b_t)$ encodes the likelihood that action $a$ is physically feasible and safe under the current belief $b_t$,
- $P_{\text{soc}}(a \mid r_t, \theta_t)$ encodes the likelihood that the same action is socially acceptable given retrieved context $r_t$ and affection-related parameter $\theta_t$.

Action selection can then be formulated as

$$a_t \in \arg\max_{a \in \mathcal{A}} P_{\text{viable}}(a \mid b_t, r_t), \tag{8}$$

optionally subject to a minimum viability threshold

$$P_{\text{viable}}(a_t \mid b_t, r_t) \geq \tau. \tag{9}$$

Importantly, this probabilistic formulation does not reintroduce a state transition model. Temporal dependence remains mediated solely through memory accumulation, retrieval, and belief updates.

### A.2 Learning and Adaptation of the Affection-Related Parameter $\theta_t$

The parameter $\theta_t$ biases both retrieval and social affordance evaluation. It represents a latent model of affective and social interaction factors, such as perceived trust, discomfort, or interaction risk.

Rather than treating $\theta_t$ as part of an internal state, we model it as an adaptive parameter updated from interaction outcomes. Let $y_t$ denote the observed outcome of executing action $a_t$, including explicit feedback or inferred human response.

We define an update rule of the general form

$$\theta_{t+1} = \theta_t + \eta\, \Delta(y_t, a_t, r_t), \tag{10}$$

where $\eta$ is a learning rate, and $\Delta(\cdot)$ is an update signal derived from interaction outcomes.

A concrete instantiation may use a prediction-error-based update:

$$\Delta(y_t, a_t, r_t) = \nabla_\theta \log P_{\text{soc}}(a_t \mid r_t, \theta_t) \cdot \delta_t, \tag{11}$$

where $\delta_t$ measures the discrepancy between predicted and observed social outcomes.

Crucially, this update does not define a transition of a state variable. Instead, $\theta_t$ functions as a slowly adapting bias that reshapes future retrieval and social affordance evaluation. Its temporal evolution is therefore indirect and history-driven, rather than governed by an explicit dynamical system.

### A.3 Semantic Interpretation

From a semantic perspective, the probabilistic viability score and the adaptive parameter $\theta_t$ do not constitute a hidden state of the robot. They are better understood as components of an *operational semantics* for action selection:

- Probability replaces binary admissibility to reflect uncertainty,
- Learning reshapes affordance evaluation rather than internal state,
- Memory and retrieval remain the sole carriers of long-term temporal structure.

Thus, even in its implementation-oriented form, the model preserves the core claim of this paper: viable action emerges from reconstructed constraints and memory-based semantics, not from state transitions.