Constraint-Grounded Dimensional Inference in Artificial Agents

An Architectural Framework for Non-Omniscient Agent Design

Stephen A. Putman
putmanmodel@pm.me

Version 0.2 — February 2026

Abstract

Artificial intelligence systems are increasingly described as possessing unbounded comprehension. Scale, parameter count, and generative capacity are often treated as proxies for structural understanding. Yet the relationship between capacity and epistemic legitimacy remains underexamined.

This paper proposes a constraint-grounded framework for dimensional inference in artificial agents. We argue that (1) representation is a structural requirement for agency, (2) persistence and identity are time-scaled properties rather than intrinsic categories, and (3) higher-dimensional structure may be inferred only through invariant boundary effects such as symmetries, conserved quantities, and forbidden transitions. Attributions extending beyond constraint-derived inference constitute narrative projection rather than justified knowledge.

The framework prioritizes how knowledge is acquired over how much knowledge is available, preserving realism, preventing architectural omniscience, and grounding intelligence in constraint-bound interaction rather than preloaded description. This approach offers architectural design implications for embodied AI and non-omniscient cognitive systems.

1. Introduction

Artificial intelligence systems are increasingly described as possessing unbounded comprehension. Scale and generative richness are often conflated with structural understanding. Yet all agents—biological or artificial—operate under constraint.

Perception is limited by embodiment. Inference is limited by accessible structure. Persistence depends on representational continuity across time. When these limits are ignored, architectural design risks conflating descriptive richness with justified inference.

Rather than treating intelligence as the progressive removal of limits, we treat limitation as structural. Agency is mediated through representation; higher-order structure is inferable only through invariant boundary effects; persistence emerges from time-scaled continuity of representation.

The objective is not to deny abstraction, but to discipline it. By distinguishing structurally justified inference from narrative projection, and grounding embodiment transfer in constraint-level abstraction, this framework articulates principles for non-omniscient artificial agents.

The contribution is architectural rather than algorithmic. It defines conditions under which inference about unseen structure remains traceable to accessible constraint surfaces. These conditions are necessary for building systems that preserve coherence, avoid dimensional inflation, and remain aligned with the limits of embodied access.

2. Terminology and Operational Definitions

For clarity, the following terms are used in an operational and architectural sense throughout this paper.

Representation

A structured internal state that preserves distinctions relevant to action under environmental constraint. Representation may be distributed or symbolic but must mediate perception and policy.

Constraint Surface (C_n)

The set of measurable states and transitions accessible to an agent operating at dimensional level n. This surface defines the limits of direct epistemic access.

Dimensional Level (*n*)

An operational designation of the degrees of freedom directly observable and controllable through the agent's sensorimotor interface.

Boundary Effect / Invariant

A persistent regularity observable within C_n that remains stable across transformations or traversals. Invariants include conserved quantities, forbidden transitions, and symmetry constraints.

Latent Variable (*L*)

An inferred internal variable introduced to restore predictive coherence when observable phenomena exceed the explanatory capacity of models confined to C_n. Latent variables are justified only insofar as they reduce inconsistency while remaining traceable to invariant boundary effects.

Narrative Projection

Attribution of properties (e.g., intention, agency, purpose) beyond what is necessitated by invariants within C_n.

Embodiment Transfer

Preservation of policy across distinct sensorimotor configurations through translation into shared constraint-level abstraction.

Knowledge Acquisition vs. Knowledge Availability

Knowledge acquisition refers to structure derived through interaction with constraint surfaces. Knowledge availability refers to preloaded or externally supplied descriptive information. This framework prioritizes acquisition over availability.

3. Representation as a Structural Requirement of Agency

Agency requires differentiation.

An agent must distinguish environmental states, potential actions, and anticipated outcomes. Without internal differentiation, action reduces to mechanical reaction; without structured mediation, adaptive persistence is impossible.

Representation is any structured internal state that preserves distinctions relevant to action under constraint. Its format may vary—weights, activations, spatial maps, distributed encodings—but its functional role is constant: it mediates perception and action.

Persistence across time requires internal states encoding environmental regularities. These encodings enable prediction, correction, and adaptation. Without such continuity, behavior dissolves into momentary responses and cannot accumulate knowledge.

Attempts to conceptualize intelligence without representation either rely on implicit representational assumptions or dissolve agency into pure reactivity. Pure reactivity cannot support counterfactual evaluation, delayed action, or policy revision. Acting in light of anticipated futures presupposes internally maintained structure.

Representation does not grant privileged access to external reality. It encodes constraint-sensitive distinctions sufficient for survival or task performance. Agency therefore operates through mediated structure rather than direct apprehension.

4. Time-Scaled Persistence and the Definition of Being

Agency unfolds over time. A single reaction does not constitute a being; persistence under changing conditions does.

Being is defined as temporally extended representational continuity sufficient to sustain adaptive policy across environmental variation. Identity is thus treated as an emergent property of sustained structural coherence rather than a metaphysical category.

An agent that updates, remembers, and maintains internal constraints across interactions exhibits persistence. This allows it to distinguish transient perturbations from stable environmental structure. Without temporal continuity, knowledge cannot accumulate.

Persistence is scale-dependent. At short timescales, behavior appears reactive; at longer timescales, pattern stability becomes visible. What qualifies as a being depends on the resolution at which continuity is evaluated.

For artificial agents, memory is not an accessory module but the substrate of persistence. Without representational continuity, there is no stable locus of policy and thus no coherent agency.

5. Boundary-Constrained Dimensional Inference

An agent operating at dimensional level n has perceptual access restricted to a constraint surface C_n.

Suppose the environment contains structure S_{n+1} requiring an additional dimension beyond the agent's direct access. The agent cannot observe S_{n+1} directly. It observes only projections:

$$P(S_{n+1}) \to C_n$$

Inference of higher-dimensional structure is justified only insofar as invariant boundary effects are detectable within C_n.

Boundary effects include:

- Invariants persistent across transformations
- Symmetries not reducible to local transitions
- Forbidden transitions
- Conservation patterns implying hidden structure

Different classes of invariants carry differing evidentiary weight; conservation constraints typically provide stronger structural evidence than symmetry alone, and forbidden transitions provide stronger constraint than statistical regularity.

If projections cannot be explained by any internally consistent model confined to C_n, the agent may introduce latent variables L to restore coherence.

Multiple higher-dimensional hypotheses may satisfy the same invariant constraints. Dimensional inference therefore yields a constrained hypothesis set rather than a unique ontological commitment. Selection among compatible hypotheses must therefore be guided by model parsimony and constraint-preserving coherence.

However, inference remains bounded:

Justified Inference ⊆ Invariant-Constrained Structure

Attributions beyond invariant constraint—such as intentionality without measurable boundary signatures—constitute narrative projection.

This distinction permits modeling unseen structure while prohibiting unjustified extension.

Detection of invariants is contingent on observational fidelity. Under partial observability, noise, or adversarial masking, justified inference may remain underdetermined or fail entirely.

6. Projection and Narrative Overreach

Let M_n represent a model restricted to dimension n, and M_{n+1} a model incorporating inferred higher-dimensional structure. Define $Q(M)$as a model-quality ordering with respect to predictive coherence on C_n. Q need not be scalar, but it must induce a partial ordering over competing hypotheses evaluated on C_n. The transition from M_n to M_{n+1} is justified only if:

$$Q(M_{n+1}) > Q(M_n)$$

and

L is traceable to boundary irregularities within C_n

and only if the additional structure introduced in M_{n+1} is traceable to invariant boundary effects detectable on C_n (i.e., it is not free explanatory slack).

Absent these conditions, dimensional expansion constitutes projection rather than justified inference.

7. Embodiment Transfer via Constraint-Level Abstraction

Let E_1 and E_2 represent distinct embodiments with constraint surfaces C_1 and C_2.

Raw sensory equivalence is neither necessary nor sufficient for policy continuity. However, if both embodiments admit translation into a shared abstraction layer A, then:

$$C_1 \rightarrow A \leftarrow C_2$$

The abstraction layer must preserve invariants common to both constraint surfaces and must be derivable from interaction rather than post hoc alignment.

Embodiment transfer is possible at the level of constraint abstraction rather than sensor identity.

8. Experimental Paradigm

To evaluate dimensional inference, agents restricted to C_n may be exposed to projections of higher-dimensional objects. The experimental question is whether agents introduce latent variables L only when invariant boundary effects necessitate structural extension.

Metrics include:

- Detection of conserved invariants
- Stability of inferred L across transformations
- Avoidance of narrative projection in the absence of invariant support

For example, predictive error on C_n may be measured before and after introduction of latent variables L, and compared against a baseline model confined to n dimensions. Dimensional expansion is justified only if measurable reduction in predictive inconsistency occurs without introduction of unconstrained explanatory degrees of freedom.

The goal is not to demonstrate omniscience, but disciplined extension under constraint.

9. Conclusion

Intelligence does not imply direct access to full environmental structure. All agency operates under constraint, through representation, across time.

Higher-dimensional inference is justified only through invariant boundary effects detectable within accessible constraint surfaces. Dimensional inflation without such grounding constitutes projection.

By prioritizing constraint-grounded acquisition over descriptive availability, this framework provides architectural principles for non-omniscient artificial agents.