The Imagination Machine III: Prediction, Control, and Representational Closure in Quasi-Periodic Environments

Mark Tracy
Boston University
mrktracy@bu.edu

Abstract

This paper develops a unified treatment of prediction, control, and representational closure for embedded epistemic systems situated in quasi-periodic environments. We proceed in three stages. First, we motivate the quasi-periodic environment as the naturalistic setting in which human temporal metacognition evolved: the Earth–Sun–Moon system presents embedded observers with incommensurate cycles whose relative phases continually drift, selecting for predictive and inductive cognitive machinery. Second, we formalize a minimal computational realization of this setting in which a predictive agent recovers latent dynamical structure from relational observations through prediction error alone. Third, we extend the framework to include action, showing that reinforcement learning arises naturally as a special case of embedded epistemic closure when policy is defined over the compressed representational classes induced by a world model. Across all three stages, the same compression–extension architecture governs representation, prediction, and control. Convergence in reinforcement learning corresponds to a fixed point of a joint model–policy closure operator, unifying representation learning and control under the structural mechanism developed throughout the Imagination Machine series.

Contents

1 Introduction: Temporal Metacognition in Quasi-Periodic Environments

History becomes possible when at least three natural cycles repeat with incommensurate periods, producing configurations whose relative phases continually drift and never exactly repeat.

The Earth–Sun–Moon system presents observers with a particular sort of temporal environment: a set of stable but incommensurate cycles whose relative phases continually drift and admit comparison. Meanwhile, the objects traveling in these cycles interact in complex ways—through tides, energy transfer, and other processes—that become crucial for biological life. Human temporal metacognition develops in response to this structured signal.

Human temporal metacognition emerges from the interaction of four recurrent processes experienced by an Earth-bound observer: the solar day, the lunar phase cycle, the solar year, and the circadian sleep–wake rhythm. The three celestial cycles possess incommensurate periods, generating quasi-periodic patterns—stable motions whose relative phases continually drift. This underlying structure makes the development of counting, memory, and inductive estimation advantageous, since empirical estimates of ratios between characteristic constants of these quasi-periodic processes accumulate through repeated observation rather than diverging without bound or collapsing into exact repetition.

Societies historically come to represent relations between characteristic constants of these cycles through ratios between them (e.g., ≈ 365 solar days per solar year), calendars, and continuous real-valued models of time. The circadian rhythm simultaneously segments subjective experience into discrete episodes through the sleep–wake cycle. The result is a dual conception of time: discrete lived intervals embedded within continuously modeled celestial motion. Each human life becomes entrained into this dynamical scaffold.

Change appears as aperiodic variation within an underlying pattern of stability. Because the relative phases of the celestial cycles continually drift, accounting for such variation benefits from the abstraction of recursively nested temporal demarcations—days within months, months within years, and so on—together with the maintenance of records across cycles.

Epistemically, the ratios governing these cycles are inductive approximations derived from observation and record-keeping. Their numerical values are refined through repeated measurement and expressed as real-valued quantities in continuous temporal models. Human symbolic systems thereby impose numerical structure on a multi-body procession whose precise relative phases are never exactly and fully known. There remains novelty amidst structure.

History becomes the maintenance of physical records of possible continuations of universal relative motion under a particular superimposed continuous and cyclic temporal model: a stochastic process of sampling-through-externalization within the world-process, enacted through acts of demarcation, recursively interpreted and reinterpreted interpersonally through (1) abstraction, which enables compressive projection; (2) analogy, which allows domain transfer of hypotheses; and (3) communication among agents.

The two formal parts of this paper develop a minimal model of the cognitive machinery this environment selects for. Part I formalizes a predictive agent that recovers latent dynamical structure from relational observations. Part II extends the framework to action, showing that control arises naturally within the same representational architecture.

Part I: Prediction

2 The Quasi-Periodic Environment

Let the environment consist of three cyclic variables

$$\theta_1(t),\, \theta_2(t),\, \theta_3(t) \in S^1 \cong \mathbb{R}/2\pi\mathbb{Z}.$$

Their dynamics are

$$\theta_i(t+1) = \theta_i(t) + \omega_i \pmod{2\pi}, \qquad i = 1, 2, 3,$$

or equivalently $\theta(t+1) = \theta(t) + \omega \pmod{2\pi}$ where $\theta(t) = (\theta_1(t), \theta_2(t), \theta_3(t))$ and $\omega = (\omega_1, \omega_2, \omega_3)$.

Definition 1 (Rational Independence)**.** *Real numbers $\omega_1, \omega_2, \omega_3$ are rationally independent if the only integer solution to $k_1\omega_1 + k_2\omega_2 + k_3\omega_3 = 0$ with $k_1, k_2, k_3 \in \mathbb{Z}$ is $k_1 = k_2 = k_3 = 0$.*

Definition 2 (Quasi-Periodic System)**.** *The dynamical system defined above is quasi-periodic if $\omega_1, \omega_2, \omega_3$ are rationally independent. In this case the trajectory is dense on the torus $\mathbb{T}^3 = S^1 \times S^1 \times S^1$.*

2.1 Observation Model

The environment state is $x_t = \theta(t)$. The agent observes only relational quantities $o_t = h(x_t)$, where

$$h(x_t) = (\cos\Delta_{12},\, \sin\Delta_{12},\, \cos\Delta_{13},\, \sin\Delta_{13},\, \cos\Delta_{23},\, \sin\Delta_{23})$$

and $\Delta_{ij}(t) = \theta_i(t) - \theta_j(t) \pmod{2\pi}$. The agent observes only relational phase differences between the cyclic processes.

2.2 Environment Distribution

Frequency vectors are sampled from the normalized simplex

$$\Delta^2 = \{\, \omega \in \mathbb{R}^3 : \omega_i > 0,\ \omega_1 + \omega_2 + \omega_3 = 1 \,\}.$$

Each sampled vector defines a distinct quasi-periodic environment.

3 The Predictive Agent

A predictive agent is defined by three functions:

$$o_t = h(x_t), \qquad s_{t+1} = u(s_t, o_t), \qquad \hat{o}_{t+1} = g(s_{t+1}),$$

where s_t is internal state and $\hat{o}_{t+1}$ is the predicted next observation.

3.1 Neural Parameterization

The state update is parameterized by a neural network:

$$s_{t+1} = \mathrm{MLP}_u([s_t,\, o_t]).$$

The prediction head is a linear readout:

$$\hat{o}_{t+1} = W s_{t+1} + b.$$

The linear prediction head creates a representational bottleneck: the internal state must organize information in a form directly readable through linear transformations.

3.2 Prediction Error and Training

Prediction error is $e_{t+1} = \hat{o}_{t+1} - o_{t+1}$. Training minimizes

$$\mathcal{L}_{t+1} = \|\hat{o}_{t+1} - o_{t+1}\|_2^2.$$

4 Observable Invariants and the Koopman Connection

4.1 Time-Rescaling Symmetry

Proposition 1. *The frequency vector ω is identifiable only up to multiplication by a positive scalar when inferred from relational phase observations alone.*

Proof. Let $k > 0$ and define $\omega' = k\omega$. Then $\theta'(t) = \theta(0) + k\omega t$, which is equivalent to $\theta'(t) = \theta(\tau)$ for $\tau = kt$. The orbit is unchanged; only the parameterization by time differs. Since the observation function h depends only on phase differences Δ_{ij}, which are invariant under uniform rescaling of ω, no relational observation can distinguish ω from $k\omega$. □

Observable invariants are therefore the projective equivalence class $[\omega_1 : \omega_2 : \omega_3]$. Normalizing via $\omega_1 + \omega_2 + \omega_3 = 1$, distinct environments correspond to points in the interior of Δ^2.

4.2 Koopman Representation

Writing the complex observable $z_{ij}(t) = e^{i\Delta_{ij}(t)}$, the relational dynamics imply

$$z_{ij}(t+1) = e^{i(\omega_i - \omega_j)}\, z_{ij}(t).$$

The observable evolves through multiplication by a constant complex phase factor, constituting a linear evolution in observable space. This is precisely a Koopman eigenfunction: the nonlinear state dynamics on $\mathbb{T}^3$ become linear in the space of relational observables. The linear prediction head therefore tests whether the agent has learned an internal representation that approximates this Koopman eigenfunction structure. Accurate prediction through a linear readout implies that the internal state encodes the relevant dynamical invariants in a linearly accessible form.

5 Empirical Protocol

5.1 Latent Structure Recovery via Linear Probing

After training on an environment with frequency vector $\omega^{(k)}$, the agent produces a final internal state $s_T^{(k)}$. A linear probe

$$\hat{y} = Ws + b$$

is trained to predict

$$y^{(k)} = \begin{pmatrix} \omega_1^{(k)}/\omega_3^{(k)} \\ \omega_2^{(k)}/\omega_3^{(k)} \end{pmatrix}$$

using mean squared error. Probe performance measures whether latent dynamical structure is represented in a linearly accessible form in the agent's internal state.

5.2 Generalization and Robustness

Training the probe on a subset of environments and evaluating on held-out environments tests whether the representation captures general dynamical structure rather than environment-specific features. The environment may be extended to weakly nonstationary dynamics by allowing slow frequency drift:

$$\omega(t+1) = \omega(t) + \epsilon_t,$$

where ϵ_t is a small perturbation. This tests the robustness of the learned representation to distributional shift.

Part II: Control

6 Action and Embedded Systems

Part I studied an agent that observes and predicts but does not intervene. Part II extends the same architecture to an agent that additionally selects actions. The formal setting follows The Imagination Machine II, in which an embedded agent and environment form a coupled dynamical system through reciprocal input–output channels:

$$u_A(t) = y_E(t), \qquad u_E(t) = y_A(t),$$

where u_A, u_E denote inputs and y_A, y_E denote outputs to the agent and environment respectively. The observations available to the agent constitute a subset

$$D \subseteq \Omega$$

of the total relational structure Ω. As in The Imagination Machine I, the agent constructs world models $w \in W$ by compressing observational profiles through an inference map $F : \Gamma \to W$, while an implication map $g : W \to \Gamma$ generates predicted observational profiles from those models.

7 Policy as Will Over Compressed Observations

A world model w induces a classifier

$$\pi_w : D \to Z_w$$

partitioning observations into representational classes via the equivalence relation

$$d \sim_w d' \iff \pi_w(d) = \pi_w(d'),$$

with induced quotient space $Q_w = D/\!\sim_w$.

Definition 3 (Policy)**.** *A policy is a stochastic map*

$$\pi : Q_w \to \Delta(A)$$

from representational classes to distributions over an action space A.

Because an embedded agent cannot act on the full observational space—it has access only to the compressed representation Q_w—policy must be defined over representational classes rather than raw observations. Policy is therefore the operational expression of will relative to the world model: the agent's selective pressure over actions, compressed to the resolution its model affords.

8 Evaluative Compression

Standard reinforcement learning treats reward as a primitive signal supplied by an external oracle. In an embedded epistemic framework this is unavailable: the agent has no access to an external vantage point from which to receive unmediated evaluative verdicts. Reward must instead arise as a compression of evaluative observations.

Let D^* denote the set of finite observation trajectories.

Definition 4 (Evaluative Compression)**.** *An evaluative compression is a map*

$$R : D^* \to \mathbb{R}$$

assigning scalar value to observational trajectories.

The reward signal therefore reflects the agent's own evaluative structure, compressed over trajectories in the same way that world models compress instantaneous observations. This is consistent with the inclusion $C \subseteq D$ established in The Imagination Machine I: classifiers—including evaluative classifiers—are themselves observations, subject to the same representational compression as any other element of D.

9 The Reinforcement Learning Closure Operator

When action is introduced, the implication map becomes policy-conditioned:

$$g : W \times \Pi \to \Gamma,$$

where Π denotes the space of policies. Given a world model and a policy, this map generates the predicted observational profile resulting from the coupled agent-environment dynamics under that policy. Inference remains

$$F : \Gamma \to W.$$

Definition 5 (RL Closure Operator)**.** *Let $\mathcal{A} : W \times R \to \Pi$ be an action-selection operator that produces a policy from a world model and an evaluative compression. The reinforcement learning closure operator is*

$$T_{\mathrm{RL}}(w, \pi) = \big(F(g(w, \pi)),\ \mathcal{A}(F(g(w, \pi)),\ R)\big).$$

Definition 6 (RL Closure)**.** *A pair (w^*, π^*) is a reinforcement learning closure if*

$$T_{\mathrm{RL}}(w^*, \pi^*) = (w^*, \pi^*).$$

At such a fixed point the world model accurately predicts the observational consequences of the policy, and the policy is optimal relative to the model and evaluative compression. The pair is jointly self-consistent in the same sense that a world model alone is self-consistent under the epistemic closure operator $T = F \circ g$.

Remark 1. *The action-selection operator $\mathcal{A}$ is left general here. Specific instantiations correspond to known algorithms: Q-learning, policy gradient methods, and actor-critic architectures each realize particular choices of $\mathcal{A}$ within this framework. Existence of a fixed point (w^*, π^*) requires conditions analogous to those governing the epistemic fixed points of The Imagination Machine I—compactness and continuity assumptions sufficient to warrant a Schauder-type argument.*

10 Exploration as Refinement

Exploration arises when the representational partition induced by the world model is too coarse to support reliable prediction or control.

Definition 7 (Refinement)**.** *A model w_2 refines w_1 if $[d]_{w_2} \subseteq [d]_{w_1}$ for all $d \in D$.*

Refinement corresponds to splitting equivalence classes in Q_w when observations within a class exhibit divergent consequences under action. An agent whose model assigns the same representational class to states with different value cannot distinguish among them in its policy. Exploration is the mechanism by which such distinctions become available.

Remark 2. *Exploration is an epistemic operator rather than random behavior: it seeks observations that maximize the probability of representational refinement. The exploration–exploitation tradeoff is therefore a special case of the knowledge–dogma distinction developed in The Imagination Machine I. An agent that ceases to explore has adopted a dogmatic closure: it holds its representational partition fixed against the pressure of new observations. The cost of this closure is not merely suboptimal reward but the structural foreclosure of refinement.*

11 Value Functions on the Quotient Space

Because policy operates on representational classes, value functions must be defined on the same space.

Definition 8 (Value Function)**.** *For a fixed policy π and world model w, the value function is*

$$V_w^\pi : Q_w \to \mathbb{R},$$

assigning expected evaluative compression to each representational class under π.

Action-value functions are defined analogously:

$$Q_w^\pi : Q_w \times A \to \mathbb{R}.$$

These functions evaluate the expected return of taking action a from representational class $[d]_w$ and thereafter following π.

12 The Koopman Connection in the Control Setting

The Koopman structure established in Part I has a direct consequence for Part II. Because the relational observables $z_{ij}(t) = e^{i\Delta_{ij}(t)}$ evolve linearly in the space of preserved invariants, value functions defined over Q_w inherit this linear structure when the world model has recovered the Koopman representation. A model that encodes the dynamical invariants in a linearly accessible internal state supports value estimation that is linear in the compressed state—which is precisely the structure that makes reinforcement learning tractable in practice.

The quasi-periodic environment is therefore not an arbitrary testbed. It is the minimal environment in which the connection between predictive representation and tractable control is explicit and provable. The Koopman eigenfunctions provide the natural basis for both prediction and value estimation, and the linear prediction head of Part I is the architectural condition that forces the agent to learn them.

13 Conclusion

This paper has developed a unified treatment of prediction, control, and representational closure for embedded epistemic systems in quasi-periodic environments.

The introduction established the naturalistic motivation. The Earth–Sun–Moon system presents embedded observers with incommensurate cycles that select for predictive and inductive cognitive machinery. Novelty amidst structure is not a special feature of this environment—it is its defining characteristic, and it is why the inference–implication loop can never fully close. The cognitive machinery the series formalizes is the machinery this environment selected for.

Part I formalized a minimal predictive agent and showed that latent dynamical structure—specifically, the Koopman eigenfunction representation of the relational observables—becomes linearly recoverable through prediction error alone. The linear prediction head is not an arbitrary architectural choice; it is the condition that forces the internal state to encode dynamical invariants in a form that makes the Koopman connection testable.

Part II extended the framework to action. Reinforcement learning arises naturally when policy is defined over the compressed representational classes induced by a world model, reward is treated as evaluative compression over trajectories, and learning seeks fixed points of a joint model–policy closure operator. Exploration is the behavioral expression of refinement pressure: the agent acts in order to find where its partition is too coarse. An agent that stops exploring has, in the precise sense of The Imagination Machine I, gone dogmatic.

Across all three stages, the same architecture governs representation, prediction, and control. Prediction, control, and valuation are not separate problems. They are different aspects of a single embedded representational structure in which an agent, unable to access the world from outside, must construct, refine, and act from within the only closure available to it.