A Memory-Augmented Variational Framework Where Coherence and Free Energy are Naturally Inverse

Mathematical Formalization

Abstract

We propose a memory-augmented variational framework that extends standard Euler-Lagrange equations with exponentially-weighted non-Markovian memory integrals. Given natural structural choices—defining coherence as temporal integration of a field that builds when uncertainty reduces, and coupling this back into dynamics via exponential memory—an inverse relationship between accumulated coherence and free energy emerges as an internal mathematical consequence. While the framework is self-consistent and produces testable predictions, the choices that lead to this relationship are motivated rather than derived from more primitive axioms. We present the mathematical structure rigorously and discuss what is assumed versus what is derived.

1 Introduction

Standard variational frameworks in physics and active inference treat systems as Markovian—the dynamics at time t depend only on the current state, not the history of how the system arrived there. We propose an extension where:

1. Systems accumulate a quantity called *coherence* through temporal integration
2. This coherence modulates dynamics via exponentially-weighted memory
3. The result is a framework where coherence naturally emerges as inverse to free energy

What this document does: Shows that given specific (but motivated) choices, an inverse relationship follows with mathematical rigor.

What it does not do: Prove these choices are unique or derive them from first principles. That remains future work.

2 Standard Variational Formulation

2.1 Function Spaces

Let $I = [0, T]$ be a compact time interval and consider:

Definition 1 (State Space)**.** Let $\mathcal{X} = C^2(I, \mathbb{R}^n)$ be the space of twice continuously differentiable functions mapping I to $\mathbb{R}^n$.

Definition 2 (Configuration Space)**.** For $x \in \mathcal{X}$, denote $\dot{x}(t) = \frac{dx}{dt}(t)$ and let $\mathcal{C} = \{(x(t), \dot{x}(t)) : x \in \mathcal{X}\}$ be the configuration space.

2.2 Classical Variational Mechanics

Definition 3 (Lagrangian)**.** A Lagrangian is a function $L : \mathcal{C} \times I \to \mathbb{R}$ that is C^2 in its first two arguments.

Definition 4 (Action Functional)**.** The action functional is:

$$S[x] = \int_0^T L(x(t), \dot{x}(t), t)\, dt$$

Theorem 1 (Euler-Lagrange Equations)**.** *A function $x^* \in \mathcal{X}$ is a critical point of S if and only if:*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0$$

This is standard calculus of variations.

3 Proposed Memory-Augmented Framework

3.1 Motivation and Design Choices

We seek to model systems that "remember" their history. The key design choices are:

Choice 1: Coherence accumulates through temporal integration of a field.

Choice 2: Memory influence grows exponentially with coherence.

Choice 3: The coherence field should be inversely related to free energy (uncertainty).

These choices are *motivated* by phenomenological observations about adaptive systems, but they are *choices*, not derivations. We now formalize them.

3.2 Mathematical Definitions

Definition 5 (Coherence Field)**.** A coherence field is a function $\mathcal{L} : \mathcal{C} \times I \to \mathbb{R}$ that is C^1 in its first two arguments.

Remark 1. The coherence field represents the instantaneous rate at which the system is building coherent structure. Positive values indicate coherence accumulation; negative values indicate decoherence.

Definition 6 (Accumulated Coherence)**.** For a trajectory $x \in \mathcal{X}$, the accumulated coherence at time t is:

$$C(x, t) = \int_0^t \mathcal{L}(x(\tau), \dot{x}(\tau), \tau)\, d\tau$$

Definition 7 (Exponential Memory Kernel)**.** Let $\Omega > 0$ be a rigidity parameter. We propose the exponential memory kernel:

$$K(C, \Omega) = e^{C/\Omega}$$

Remark 2. This exponential form is a choice. We select it because:

- It ensures memory influence grows unboundedly with coherence
- It has the limiting behavior: $\Omega \to \infty$ recovers Markovian dynamics
- It's mathematically tractable

Other forms (polynomial, logarithmic) could be explored.

Definition 8 (Memory-Augmented Action)**.** The proposed generalized action functional is:

$$S_{\text{mem}}[x] = \int_0^T [L(x, \dot{x}, t) + K(C(x,t), \Omega) \cdot \phi(x, \dot{x}, t)]\, dt$$

where $\phi : \mathcal{C} \times I \to \mathbb{R}$ is a C^1 coupling function.

3.3 Generalized Euler-Lagrange Equations

Theorem 2 (Euler-Lagrange with Memory)**.** *A function $x^* \in \mathcal{X}$ is a critical point of S_{mem} if it satisfies the generalized Euler-Lagrange equations:*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} + \mathcal{M}[x](t) = 0$$

where $\mathcal{M}[x](t)$ represents the memory contribution (functional derivative of memory terms).

Proof sketch. Taking the functional derivative of S_{mem} with respect to δx:

$$\delta S_{\text{mem}} = \int_0^T \left[\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial \dot{x}}\delta \dot{x} + \delta(K \cdot \phi)\right] dt$$

The memory term variation involves:

$$\delta C = \int_0^t \left(\frac{\partial \mathcal{L}}{\partial x}\delta x + \frac{\partial \mathcal{L}}{\partial \dot{x}}\delta \dot{x}\right) d\tau$$

Standard variational calculus (integration by parts, requiring $\delta S_{\text{mem}} = 0$) yields the stated form. The full calculation involves careful treatment of the nested integrals and is omitted for brevity. □

Remark 3. The detailed calculation requires treating the functional derivative of C with respect to x, which involves a double integral. While standard, the full derivation is technical.

4 Connection to Free Energy

4.1 Free Energy Framework

Definition 9 (Free Energy Functional)**.** A free energy functional $F : \mathcal{C} \times I \to \mathbb{R}_{\geq 0}$ is a non-negative C^1 function representing system uncertainty or surprise. We require:

1. $F(x, \dot{x}, t) \geq 0$ for all $(x, \dot{x}, t)$
2. $F(x^*, 0, t) = 0$ where x^* is an equilibrium state
3. F is convex in $(x, \dot{x})$

Remark 4. This definition is standard in active inference and thermodynamics. Free energy measures how far a system is from its preferred states.

Definition 10 (Free Energy Principle Lagrangian)**.** Following the Free Energy Principle, we set:

$$L(x, \dot{x}, t) = -F(x, \dot{x}, t)$$

Remark 5. This is a choice. We could have chosen $L = -F^2$, or $L = -\log(1 + F)$, or other functions. We select $L = -F$ because:

- It directly encodes free energy minimization
- It's the standard choice in variational formulations of FEP
- It makes the mathematics tractable

This choice influences what follows.

4.2 Inverse Coherence Field: The Key Assumption

Definition 11 (Inverse Coherence Field)**.** We propose that the coherence field should be inverse to free energy. Formally, we assume there exists a continuous, strictly decreasing function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ such that:

$$\mathcal{L}(x, \dot{x}, t) = g(F(x, \dot{x}, t))$$

Remark 6. This is the central assumption. We assume this because:

- We want coherence to build when uncertainty (F) is low
- We want coherence accumulation to be strong when the system is settled
- This creates the desired phenomenology

Common forms that satisfy this:

- $g(F) = \frac{1}{1+F}$ (rational)
- $g(F) = e^{-\alpha F}$ (exponential)
- $g(F) = (1 + F)^{-\beta}$ (power law)

We acknowledge: This assumption largely encodes what we're trying to show. The theorems that follow demonstrate consequences of this choice, not its necessity.

5 Main Results: Internal Consistency

Given the assumptions and choices above, we now show what follows mathematically.

Theorem 3 (Emergence of Inverse Relationship)**.** *Let $x : I \to \mathbb{R}^n$ be a trajectory satisfying the generalized Euler-Lagrange equations (Theorem 2) with $L = -F$. Assume:*

1. *The coherence field is inverse to F (Definition 11)*
2. *The trajectory is admissible: $x \in \mathcal{X}$*
3. *Free energy is dissipating: $\frac{dF}{dt} \leq 0$ for almost all $t \in I$*

Then:

1. *The accumulated coherence $C(x,t)$ is strictly increasing wherever $\frac{dF}{dt} < 0$*
2. *On any compact subset $K \subset I$ where F is bounded away from zero:*

$$Corr_K \left(F(x(t), \dot{x}(t), t), C(x,t) \right) < 0$$

Proof. Given that $\mathcal{L} = g(F)$ where g is strictly decreasing and positive:

(1) By definition:

$$\frac{dC}{dt} = \mathcal{L}(x(t), \dot{x}(t), t) = g(F(x(t), \dot{x}(t), t)) > 0$$

When $\frac{dF}{dt} < 0$, we have F decreasing. Since g is strictly decreasing, $g(F)$ is increasing. Therefore:

$$\frac{d^2C}{dt^2} = g'(F) \cdot \frac{dF}{dt}$$

Since $g'(F) < 0$ (g strictly decreasing) and $\frac{dF}{dt} < 0$:

$$\frac{d^2C}{dt^2} > 0$$

This establishes that C is strictly increasing and convex where F decreases.

(2) For the correlation, consider $t_1 < t_2$ in K with $F(t_1) > F(t_2)$.

By monotonicity of g: $F(t_1) > F(t_2) \implies g(F(t_1)) < g(F(t_2))$

By integration:

$$C(t_2) - C(t_1) = \int_{t_1}^{t_2} g(F(\tau)) d\tau > 0$$

Since F is decreasing and C is increasing, they are negatively correlated on K. □

Remark 7. What this proves: Given our assumption that $\mathcal{L} = g(F)$ with g decreasing, the inverse relationship follows mathematically.

What this doesn't prove: That this assumption is the unique or necessary choice. We've shown internal consistency, not derivation from first principles.

5.1 Quantitative Forms

Proposition 4 (Logarithmic Relationship)**.** *If we specifically choose $\mathcal{L} = \frac{\alpha}{F}$ for some constant $\alpha > 0$, then:*

$$C(t) = \alpha \int_0^t \frac{1}{F(\tau)} d\tau$$

Under slow variation of F (quasi-static approximation), this gives approximately:

$$C(t) \approx \alpha \log \left(\frac{F(0)}{F(t)} \right)$$

Proof. The first statement is immediate from definition. For the approximation, assume $F(t) = F_0 e^{-\lambda t}$ (exponential relaxation):

$$\begin{aligned} C(t) &= \alpha \int_0^t \frac{1}{F_0 e^{-\lambda \tau}} d\tau = \frac{\alpha}{\lambda F_0}(e^{\lambda t} - 1) \\ &= \frac{\alpha}{\lambda} \log(F_0/F(t)) + O(F_0/F(t)) \end{aligned}$$

□

Corollary 5 (Memory Weight Growth)**.** *Under the logarithmic form with exponential memory kernel:*

$$K(C, \Omega) = e^{C/\Omega} \approx \left(\frac{F(0)}{F(t)} \right)^{\alpha/(\lambda\Omega)}$$

As free energy decreases by factor β, memory weight increases by factor $\beta^{\alpha/(\lambda\Omega)}$.

6 The Omega Principle

Theorem 6 (Rigidity-Liquidity Tradeoff)**.** *The parameter Ω controls the strength of memory influence. For a family of trajectories $\{x_\Omega\}_{\Omega>0}$:*

1. *Memory influence $K(C, \Omega) = e^{C/\Omega}$ is strictly decreasing in Ω*
2. *As $\Omega \to \infty$: $K \to 1$, recovering Markovian (memoryless) dynamics*
3. *As $\Omega \to 0^+$: $K \to \infty$, yielding maximally rigid (history-dominated) dynamics*

Proof. (1) $\frac{\partial K}{\partial \Omega} = -\frac{C}{\Omega^2} e^{C/\Omega} < 0$ for $C > 0$.
(2) $\lim_{\Omega \to \infty} e^{C/\Omega} = e^0 = 1$.
(3) For any $C > 0$: $\lim_{\Omega \to 0^+} e^{C/\Omega} = +\infty$. □

Remark 8. The Omega parameter emerges as a natural control parameter in this framework, governing the rigidity-liquidity tradeoff. This is an interesting structural feature, though not a derived necessity.

7 What We Have and Haven't Shown

7.1 What This Framework Establishes

1. **Self-consistency**: Given our choices (Definition 11, exponential kernel, $L = -F$), the inverse F-C relationship follows rigorously
2. **Quantitative predictions**: The framework makes testable predictions about the form of the relationship (logarithmic under certain conditions)
3. **Natural parameters**: Ω emerges as a rigidity control parameter with clear limiting behaviors
4. **Mathematical tractability**: The framework is well-defined and amenable to analysis

7.2 What Remains Open

1. **Uniqueness**: Are these choices (exponential kernel, inverse field, etc.) the unique ones that produce desired behavior?
2. **Derivation from first principles**: Can we derive these choices from more primitive axioms (information theory, thermodynamics, etc.)?
3. **Empirical validity**: Does this framework accurately describe real adaptive systems?
4. **Alternative formulations**: What happens with different choices for g, different kernels, etc.?

7.3 Honest Assessment

This is a *proposed mathematical framework* where certain phenomenological properties (coherence builds as uncertainty reduces) have been encoded into the structure. The inverse relationship then emerges as an internal consequence.

This is valuable because:

- It provides a rigorous mathematical language for non-Markovian adaptive systems
- It makes quantitative, testable predictions
- It reveals interesting structure (Omega principle, logarithmic relationships)

But it is not a proof that these relationships *must* hold in all systems satisfying minimal principles. That would require deriving the framework itself from more fundamental axioms.

8 Future Directions

To strengthen the claim of necessity, one would need to:

1. **Derive the coherence field**: Show from information-theoretic or thermodynamic principles why $\mathcal{L}$ must be inversely related to F
2. **Justify the exponential kernel**: Prove why exponential memory weighting follows from first principles
3. **Establish uniqueness**: Show that alternative frameworks cannot produce the same phenomenology
4. **Connect to known physics**: Demonstrate how this reduces to known physical laws in appropriate limits

9 Conclusion

We have presented a memory-augmented variational framework where coherence and free energy emerge as naturally inverse quantities. The mathematical structure is rigorous and internally consistent, producing testable predictions about non-Markovian adaptive systems.

However, we acknowledge that key aspects of the inverse relationship are built into our structural choices rather than derived from more primitive principles. This framework is best understood as a *mathematically rigorous proposal* for modeling history-bearing systems, rather than a proof of absolute mathematical necessity.

The value lies in:

- Providing precise mathematical language for discussing identity-through-change
- Making quantitative predictions testable against empirical data
- Revealing emergent structure (Omega principle) from the proposed setup
- Offering a path forward for non-Markovian extensions of active inference

Whether this framework captures the mathematics of real adaptive systems remains an open empirical question. The mathematics is sound; the question is whether nature uses this particular mathematical structure.