

Affective Active Inference: From Trace to Shared Engram and Narrative Identity Formation

Ryota Sawaki
Department of Computational Psychiatry
ryota.sawaki@example.com

November 21, 2025

Abstract

Human identity does not emerge from isolated optimization but from the accumulation of shared affective meanings. We propose the **Self-Imprint Attribution (SIA)** model, a computational framework that integrates Active Inference with affective dynamics. Unlike traditional models that minimize prediction errors instantaneously, SIA posits that errors are retained as *Trace*, transformed into vectorized *Affect*, and expressed through *Creative Action*. We demonstrate that narrative identity (*I*) emerges only when these affective vectors synchronize between agents, forming a *Shared Engram*. Furthermore, our sensitivity analysis reveals a paradoxical insight: agents with higher sensitivity to trauma (α) form significantly stronger identities, suggesting that vulnerability is a prerequisite for robust social resilience.

1 Introduction

Why do we hold onto painful memories? In standard Reinforcement Learning (RL) and Active Inference, prediction errors (surprisal) are costs to be minimized or ignored once learning is complete. However, in human psychology, unresolved errors often form the core of one's personality—a phenomenon known as trauma or core belief.

We propose that these errors are not waste products but the raw material of identity. We introduce **Affective Active Inference**, where:

1. **Trace:** Discrepancies are imprinted physically.
2. **Affect:** Traces are interpreted into qualitative vectors (e.g., Sorrow, Hope).
3. **Identity:** These vectors are integrated over time through shared resonance.

2 Computational Model

The SIA agent operates on a cycle of five phases. The core mechanism is the *Attribution Gate*, which determines ownership of experience.

2.1 Self-Attribution and Trace

The probability that an experience E belongs to the self, $P(Self|E)$, is defined by:

$$P(Self|E) = \sigma(-\|E - S\| + \alpha\|T\| + \beta\|Act_{prev}\|) \quad (1)$$

where T is the accumulated trace, and α is the **Trace Sensitivity** parameter. A higher α implies that the agent is more likely to attribute painful discrepancies to itself (internalization).

2.2 Identity Integration via Shared Resonance

Identity $I(t)$ is not a static variable but a historical integral. It grows only when a *Shared Engram* is formed with another agent:

$$I(t+1) = I(t) + \eta \cdot \text{Shared}(t) \cdot \mathbf{A}(t) \quad (2)$$

where $\mathbf{A}(t)$ is the affective vector. The shared resonance is strictly defined as:

$$\text{Shared}(t) = P_1 P_2 \cdot \cos(\theta_{act}) \cdot \cos(\theta_{aff}) \cdot \min(\|\mathbf{A}_1\|, \|\mathbf{A}_2\|) \quad (3)$$

This ensures that identity is formed only when both agents share both the *direction of action* and the *quality of affect*.

3 Simulation Results

We conducted two experiments to validate the theory.

3.1 Dynamics of Narrative Identity Formation

Figure 1 illustrates the time evolution of an agent interacting with a partner after a traumatic event ($t = 50$).

- **Phase 1 (Trauma):** The agent experiences a shock, generating a large Trace.
- **Phase 2 (Genesis):** The trace is converted into Affect (Meaning).
- **Phase 3 (Resonance):** Upon encountering a partner ($t = 100$), the Shared Engram (Green area) emerges.
- **Result:** The Identity (I , Black line) begins to accumulate only after resonance occurs, supporting the hypothesis: "I become who we were."

Figure 1: Narrative Identity Formation
[Place 'identity_formation.png' here]
Identity grows only during shared resonance.

Figure 1: Simulation of Narrative Identity Formation. The Identity (bottom panel) grows only when Shared Resonance (middle panel, green) is active.

3.2 Sensitivity Analysis: The Paradox of Vulnerability

To investigate the role of individual differences, we performed a parameter sweep on Trace Sensitivity α ($0.1 \leq \alpha \leq 2.0$). As shown in Figure 2, there is a strong positive correlation between α and the final magnitude of Identity.

Figure 2: Sensitivity Analysis
[Place 'sensitivity_analysis.png' here]
Higher sensitivity (α) leads to stronger identity.

Figure 2: Sensitivity Analysis. Agents with higher sensitivity to trace (α) form stronger identities. Vulnerability fosters resilience.

4 Discussion

The results present a counter-intuitive insight: **Vulnerability is Strength**. Agents that easily ignore traces ($\alpha \approx 0$) minimize immediate costs but fail to accumulate the affective resources

necessary for deep social resonance. Conversely, agents that retain traces ($\alpha > 1.0$) suffer more initially but use that suffering as fuel for creative action and identity formation.

This suggests that in the context of General Intelligence, "forgetting" is not always optimal. The ability to be scarred (Imprint) is the engine of becoming someone (Identity).

5 Conclusion

We demonstrated that Narrative Identity can be computationally modeled as an integral of shared affective history. The SIA model provides a new mathematical bridge between clinical concepts of trauma and engineering concepts of AGI.