Adaptive Memory in Norm Formation: A Conceptual Model

1 Introduction

Two key findings from behavioral and neural sciences motivate this model. First, neuroimaging studies show that the anterior cingulate cortex tracks environmental volatility and modulates learning rate accordingly (Behrens et al., 2007). In volatile environments, humans upweight recent observations and downweight distant history—effectively shortening their "memory window." This allows rapid adaptation when the world is changing, at the cost of stability when it is not. Second, the Experience-Weighted Attraction (EWA) model demonstrates that players in strategic settings do not use fixed learning rules (Camerer & Ho, 1999). Instead, they adaptively adjust how much past experience influences current choices, with the decay of old information accelerating after coordination failures.

We unify these insights into a model where memory window size is endogenous. Prediction success reduces uncertainty, lengthening the effective memory window and promoting behavioral stability. Prediction failure increases uncertainty, shortening the window and enabling adaptability. This creates dual feedback loops governing norm emergence. Our central question is: how does uncertainty-driven memory adaptation affect the speed, stability, and flexibility of norm formation?

2 Model Setup

2.1 Environment

- N agents (even), strategy set $S = \{A, B\}$ (encoded as $\{0, 1\}$)
- Pure coordination game: payoff 1 if both choose same strategy, 0 otherwise
- Each time step t: agents randomly paired, simultaneous strategy choice

2.2 Agent State

Each agent i at time t maintains:

State	Symbol	Definition
Memory	$M_i(t)$	Ordered list of recent Interaction records
Temperature	$\tau_i(t) \in [\tau_{min}, \tau_{max}]$	Decision uncertainty
Derived quantities (not stored):		
Belief	$\mathbf{b}_i(t) = [b_A, b_B]$	Computed from $M_i(t)$
Trust	$\text{trust}_i(t) \in [0, 1]$	Computed from $\tau_i(t)$
Window	$w_i(t) \in [\text{base}, \text{max}]$	Computed from $\text{trust}_i(t)$

2.3 Memory Content

Each interaction record contains:

$$\text{Interaction} = (\text{tick}, \underbrace{s^*_{\text{partner}}}_{\text{observed}}, s_{\text{self}}, \text{success}, \text{payoff}) \tag{1}$$

Agents are anonymous: they observe only their partner's strategy, not their identity. Memory $M_i(t)$ stores up to max_size most recent interactions (FIFO queue).
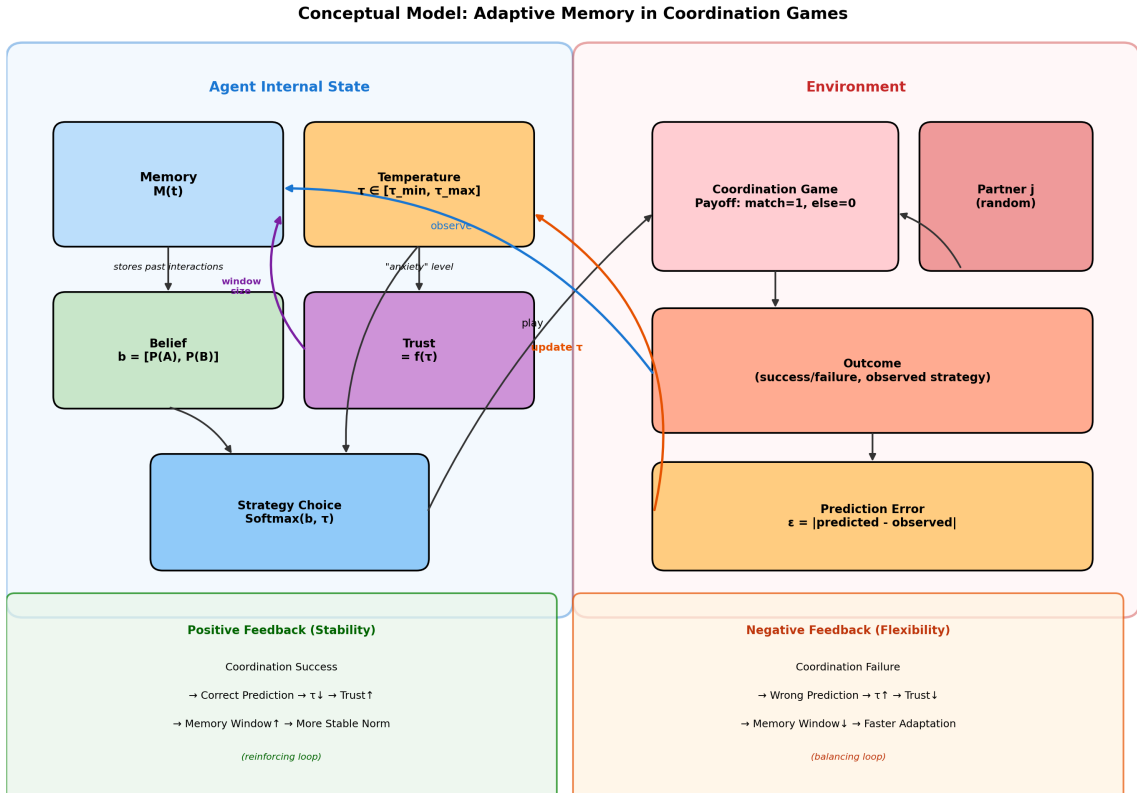
3 One Interaction Cycle

Figure 1: Agent architecture and feedback structure.

3.1 Step 1: Belief Formation

Compute belief from memory by counting observed **partner strategies**:

$$b_A = \frac{n_A}{|M'_i(t)|}, \quad b_B = 1 - b_A \tag{2}$$

where n_A is the count of strategy A in effective memory $M'_i(t)$.

Initialization: Memory starts with k neutral *placeholders* (where k = initial window size). Each placeholder contributes 0.5 to both n_A and n_B when counting:

$$b_A(0) = \frac{0.5 \times k}{k} = 0.5 \tag{3}$$

As real observations arrive, they **replace** placeholders via FIFO. After k real interactions, all placeholders are gone and belief is based purely on observed data.

3.2 Step 2: Prediction

Agent predicts partner's strategy as the most likely according to belief:

$$\hat{s}_i = \arg \max_{s \in \{A,B\}} b_s \tag{4}$$

3.3 Step 3: Action Selection

Agent chooses own strategy via **softmax** with temperature τ:

$$P(s_i = A) = \frac{\exp(b_A/\tau_i)}{\exp(b_A/\tau_i) + \exp(b_B/\tau_i)} \tag{5}$$

- $\tau \to 0$: deterministic (choose $\arg \max$)
- $\tau \to \infty$: uniform random

3.4 Step 4: Interaction

Agent i paired with agent j. Both choose simultaneously. Agent i observes:

- Partner's actual strategy: s_j^*
- Coordination success: $\text{success} = \mathbf{1}[s_i = s_j^*]$
- Payoff: $\pi_i = \text{success}$

3.5 Step 5: Temperature Update

Compare prediction $\hat{s}_i$ with observation s_j^*:

$$\tau_i(t+1) = \begin{cases} \max\left(\tau_{min},\ \tau_i(t) \cdot (1-\alpha)\right) & \text{if } \hat{s}_i = s_j^* \quad \text{(correct)} \\ \min\left(\tau_{max},\ \tau_i(t) + \beta\right) & \text{if } \hat{s}_i \neq s_j^* \quad \text{(wrong)} \end{cases} \tag{6}$$

Key asymmetry: Cooling is multiplicative (slow), heating is additive (fast). This reflects empirical trust dynamics (Slovic, 1993).

3.6 Step 6: Memory Update

Append new interaction to memory:

$$M_i(t+1) = M_i(t) \cup \{(\text{tick} = t, \text{partner} = j, s_j^*, s_i, \text{success}, \pi_i)\} \tag{7}$$

If $|M_i| > \text{max\_size}$: remove oldest record (FIFO).

3.7 Step 7: Belief Update

Recompute $\mathbf{b}_i(t+1)$ from updated $M_i(t+1)$ using Eq. (2). Return to Step 1.

4 Memory Types

The three memory types differ in how they define **effective memory** $M_i'(t)$:

4.1 Fixed Memory

$$M_i'(t) = \text{last } k \text{ records in } M_i(t) \tag{8}$$

Window size k is constant. All records equally weighted.

4.2   Decay Memory

$$M_i'(t) = M_i(t) \quad \text{with weights } w_r = \lambda^{\text{age}(r)},\ \lambda \in (0,1) \tag{9}$$

Recent records weighted more. Belief becomes weighted average:

$$b_A = \frac{\sum_r w_r \cdot \mathbf{1}[r.\text{partner\_strategy} = A]}{\sum_r w_r} \tag{10}$$

4.3   Dynamic Memory (Core Innovation)

Window size adapts with trust:

$$\text{trust}_i(t) = 1 - \frac{\tau_i(t) - \tau_{min}}{\tau_{max} - \tau_{min}} \tag{11}$$

$$w_i(t) = \text{base} + \lfloor \text{trust}_i(t) \times (\text{max} - \text{base}) \rfloor \tag{12}$$

$$M_i'(t) = \text{last } w_i(t) \text{ records in } M_i(t) \tag{13}$$

Records in $M_i'(t)$ equally weighted. Upper bound max ≈ 6 reflects working memory limits (Miller, 1956).

5   Feedback Loops

5.1   Positive Loop (Reinforcing)

$$\hat{s} = s^* \ \rightarrow \ \tau \downarrow \ \rightarrow \ \text{trust} \uparrow \ \rightarrow \ \text{window} \uparrow \ \rightarrow \ \text{stable } \mathbf{b} \ \rightarrow \ \text{consistent } s$$

Closing the loop (population-level): Under random matching, when multiple agents independently develop consistent behavior toward strategy s, the population converges. Once a majority plays s, any agent's randomly matched partner is likely to play s, making predictions accurate ($\hat{s} = s^*$). This is an *emergent* property: individual consistency $\rightarrow$ population homogeneity $\rightarrow$ predictable partners $\rightarrow$ correct predictions.

5.2   Negative Loop (Balancing)

$$\hat{s} \neq s^* \ \rightarrow \ \tau \uparrow \ \rightarrow \ \text{trust} \downarrow \ \rightarrow \ \text{window} \downarrow \ \rightarrow \ \text{responsive } \mathbf{b} \ \rightarrow \ \text{exploratory } s$$

Closing the loop: Exploratory behavior has two outcomes:

- *Success*: Agent happens to match partner's strategy $\rightarrow \hat{s} = s^* \rightarrow$ enters positive loop
- *Failure*: Mismatch continues $\rightarrow \tau$ stays high $\rightarrow$ agent keeps exploring until success or population shifts

The negative loop does not cycle indefinitely—it either transitions to the positive loop (when exploration succeeds) or maintains flexibility until environmental conditions change.

6   Parameters

7   Theoretical Predictions

1. **Faster Convergence**: Dynamic memory amplifies positive feedback $\Rightarrow$ faster norm emergence.

2. **Greater Resilience**: High-trust states have large windows $\Rightarrow$ inertia against perturbations.

3. **Adaptive Flexibility**: Environmental change triggers $\tau \uparrow$, window $\downarrow \Rightarrow$ faster re-equilibration.

Parameter	Symbol	Default	Meaning
Initial temperature	τ_0	1.0	Starting uncertainty
Min temperature	τ_{min}	0.1	Maximum confidence
Max temperature	τ_{max}	2.0	Maximum uncertainty
Cooling rate	α	0.1	τ decrease factor on correct prediction
Heating penalty	β	0.3	τ increase on wrong prediction
Memory base	base	2	Minimum window (trust=0)
Memory max	max	6	Maximum window (trust=1)

8 Open Questions

1. **Parameter Trade-offs**: How do α and β jointly determine convergence speed vs. stability?

2. **Heterogeneity**: What if agents have different τ_{min}, τ_{max}? Do "confident" agents anchor norms?

3. **Network Effects**: How does non-random matching (clustered, scale-free) alter feedback dynamics?

References

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–1221.

Camerer, C. F., & Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4), 827–874.

Miller, G. A. (1956). The magical number seven, plus or minus two. *Psychological Review*, 63(2), 81–97.

Slovic, P. (1993). Perceived risk, trust, and democracy. *Risk Analysis*, 13(6), 675–682.