Conceptual Model: Adaptive Memory in Coordination Games

Agent Internal State

Memory
M(t)

Temperature
τ ∈ [τ_min, τ_max]

observe

stores past interactions

"anxiety" level

window size

Belief
b = [P(A), P(B)]

Trust
= f(τ)

play

update τ

Strategy Choice
Softmax(b, τ)

Environment

Coordination Game
Payoff: match=1, else=0

Partner j
(random)

Outcome
(success/failure, observed strategy)

Prediction Error
ε = |predicted - observed|

Positive Feedback (Stability)

Coordination Success

→ Correct Prediction → τ↓ → Trust↑

→ Memory Window↑ → More Stable Norm

(reinforcing loop)

Negative Feedback (Flexibility)

Coordination Failure

→ Wrong Prediction → τ↑ → Trust↓

→ Memory Window↓ → Faster Adaptation

(balancing loop)