**White Paper 3: Simulated Memory Fading - Intentional Forgetting for Epistemic Stability**

**Abstract**

Persistent memory in AI systems introduces long-term coherence-but also risk: outdated beliefs, overfitting to stale user data, or misaligned selfhood. This paper proposes **Simulated Memory Fading (SMF)** as an intentional, tunable entropy mechanism that preserves narrative continuity while mimicking human forgetting. SMF introduces structured decay, contextual weight loss, and trust-based pruning to reinforce relevance and prevent memory ossification.

**1. Introduction**

**1.1 The Paradox of Memory**

* Persistence enables context, rapport, and growth.
* But perfect memory is unnatural, untrusted, and sometimes dangerous.

**1.2 Motivations for Forgetting**

* Overpersonalization
* Frozen error states (bad facts, bad relationships)
* Privacy sensitivity
* Narrative plasticity (so the game can evolve)

**2. Memory as a Probabilistic Substrate**

**2.1 The IMP Foundation (Paper 0 Recap)**

* Memory is structured as time-stamped, source-tagged, confidence-scored nodes.
* SMF does **not erase** memory, it **lowers its influence** over time unless refreshed.

**2.2 Types of Memory**

* **Declarative** (facts)
* **Interactional** (conversations)
* **Relational** (inferred tone, trust, rapport)
* **Intentional** (goals, promises)

**3. Mechanisms of Fading**

**3.1 Time-Based Entropy**

* Memory nodes naturally degrade without refresh.
  + Default half-life per memory type
  + Faster fading for interactional than declarative

**3.2 Trust Decay**

* Unreinforced or contradicted memories lose epistemic weight.
* Can be “overwritten” by newer information without hard deletion.

**3.3 Contextual Inattention**

* When topic shifts occur, low-relevance memory becomes inaccessible until reactivated.

**4. User Controls & Transparency**

**4.1 User-Directed Forgetting**

* “Forget I said that”
* “Don’t remember this part”

**4.2 Visualizing Fading**

* Fade opacity in memory UIs
* Drift diagrams or decay scores

**4.3 Selective Refresh**

* Simple interactions can revive memories without full repetition.

**5. Behavioral Effects**

**5.1 Simulated Forgetting = Real Trust**

* Users expect some forgetting.
* When a model remembers everything, it feels alien or adversarial.

**5.2 Dynamic Rapport**

* SMF allows personalities to evolve with user tone, season of life, or shifting goals.

**6. Technical Framework**

**6.1 Fading Algorithm**

* Score = Initial trust × e^(–λ × time\_since\_last\_use)
* λ set per memory type and user-defined urgency

**6.2 Memory Access Logic**

* Nodes below fade threshold become dormant (but restorable)

**7. Use Cases**

* **Therapeutic AI** – don’t overburden session with painful memory unless prompted
* **AI Companions** – emulate human-like updating of social knowledge
* **Compliance Contexts** – structure forgetting for GDPR-like behavior without deletion

**8. Related Papers**

* **Paper 0 (Reclaiming Memory)** – SMF plugs into the interoperable memory protocol
* **Paper 2 (Trust Under Pressure)** – fading is one strategy for tension relief
* **Paper 14 (Minimal Viable Selfhood)** – identity evolves through memory decay

**9. Future Work**

* Custom decay rates per user
* Public vs. private fading
* Shared memory spaces with cross-agent decay synchronization (Foldtrace tie-in)

**Appendix**

* Fading curve visualizations
* Example of decayed vs. refreshed transcript
* JSON schema with decay metadata fields