



# Beyond Reward Maximization: A Pressure-Driven Cognitive Architecture for Continual Cross-Domain Generalization and Stability

● Likhith Sai Seemala ·

RAVANA Research Initiative; Independent Researcher; Andhra Pradesh, India; Email: likhith@oxiverse.com

CITE AS <a href="https://openxiv:psy.cog.2026.00001">openxiv:psy.cog.2026.00001</a> ISSN 3120-9556 (online) LICENSE CC-BY-4.0	POSTED 2026-05-28 VERSION v1 SUBJECT psy.cog
--	---

## AI DISCLOSURE ASSISTANT

### COVER EVIDENCE

TRANSPARENCY · partial	IDENTITY · strong	PROVENANCE · strong	CITATIONS · strong	MATH · partial
INTEGRITY · partial				

### CANONICAL RECORD

<https://openxiv.net/abs/psy.cog.2026.00001>

Cite as: openxiv:psy.cog.2026.00001

Live verification record is maintained on the canonical abstract page.

DOI will be deposited and back-filled once Crossref membership clears.



scan to open

# Beyond Reward Maximization: A Pressure-Driven Cognitive Architecture for Continual Cross-Domain Generalization and Stability

Likhith Sai Seemala  
RAVANA Research Initiative  
Independent Researcher  
Andhra Pradesh, India  
Email: likhith@oxiverse.com

May 25, 2026

*Author Note. This work presents a prototype cognitive architecture and internally reported benchmark results intended for exploratory cognitive-science discussion and future replication.*

## Abstract

Continual learning systems must acquire new structure without erasing prior knowledge, transfer relational schemas across domains, and regulate behavior under uncertainty. This paper presents RAVANA (Recursive Adaptive Variable Architecture for Neurocognitive Autonomy), a pressure-driven cognitive architecture that combines a Recursive Learning Model (RLM), a typed ConceptGraph, predictive-coding style local settling, sleep-time consolidation, episodic replay, and homeostatic cognitive variables. Unlike standard reward-maximizing or end-to-end backpropagation systems, RAVANA treats error, dissonance, novelty, identity stability, and memory pressure as interacting control signals. We describe the architecture in detail and report internal benchmark results from within-domain association learning, cross-domain analogy probes, lifelong streaming retention, and synthetic fairness stress tests. In these experiments, architectural fixes raised within-domain top-1 accuracy from 0% to 100%; cross-domain probes reached 14.3% top-1 and 71.4% top-10; and a combined replay, elastic-weight-consolidation, and Bayesian edge-posterior mechanism reduced measured catastrophic forgetting from 12.0% to 0.0% in a 15,000-experience stream. These results should be interpreted as prototype evidence rather than final cognitive validity: the benchmarks are synthetic, the architecture is small, and independent replication is required. Nevertheless, the system offers a concrete computational hypothesis for cognitive science: robust generalization may emerge not only from reward maximization, but from regulated tensions among prediction error, structural analogy, memory consolidation, and self-stabilizing identity constraints.

**Keywords:** cognitive architecture; continual learning; predictive coding; complementary learning systems; catastrophic forgetting; analogy; sleep consolidation; self-organization

# 1 Introduction

Human cognition is not well described as a sequence of isolated supervised updates. People learn from sparse experience, reuse relational schemas across domains, preserve much of what they previously knew, sleep and rehearse, regulate uncertainty, and sometimes resist immediately rewarding actions when those actions conflict with longer-term goals or identity. These capacities motivate a long tradition in cognitive science that treats cognition as an adaptive, self-organizing, memory-dependent process rather than a purely reward-maximizing calculation [10, 8, 16, 6, 13].

Modern machine learning has achieved extraordinary performance with large-scale differentiable models, especially transformers trained by backpropagation. Yet several problems remain central for cognitive science and artificial intelligence: catastrophic interference during sequential learning [17, 7], weak causal or analogical transfer outside the training distribution, limited interpretability of internal structure, and difficulty integrating affective, motivational, and memory-like regulatory variables into a unified architecture. Continual-learning methods such as elastic weight consolidation, synaptic intelligence, gradient episodic memory, progressive networks, and rehearsal buffers address parts of this problem, but they often remain external additions to a task-optimized learning system [12, 25, 14, 20, 19].

This manuscript studies RAVANA, a prototype cognitive architecture developed around a different organizing principle: cognition is governed by pressure. Prediction error, free energy, novelty, dissonance, memory salience, identity stability, confidence, and fairness tension are represented as internal variables that regulate learning and action. The model therefore does not ask only, “Which action maximizes reward?” It also asks whether a new update is coherent with prior structure, whether an apparent fact contradicts existing memory, whether a concept is overloaded, whether replay is needed, whether an edge is reliable, and whether a policy shift destabilizes identity.

The paper is written for a cognitive science audience. Its goal is not to claim human-level cognition or superiority over large language models. Instead, it provides a detailed architectural account, reports internal synthetic benchmarks, identifies current limitations, and frames RAVANA as a testable computational hypothesis about how local learning, graph-structured concepts, replay, predictive coding, and homeostatic governance can be integrated.

## 1.1 Contributions

The manuscript makes five contributions.

1. It specifies a pressure-driven architecture combining an RLM, typed ConceptGraph, recurrent sequence processing, predictive-coding settle loops, ConceptBinding, episodic memory, and sleep consolidation.
2. It formalizes a governance layer in which dissonance, identity strength, uncertainty, novelty, and fairness pressure modulate learning and policy acceptance.
3. It reports internal benchmark results on within-domain learning, cross-domain analogy, lifelong streaming retention, and bias-mitigation stress tests.

4. It contrasts RAVANA with backpropagation-centered systems, transformer attention, state-space models, complementary learning systems, and predictive-coding accounts.
5. It identifies failure modes and engineering gaps, including sample efficiency, speed, scaling, local-context limitations, and the need for independent validation.

## 2 Theoretical Background

### 2.1 Catastrophic Forgetting and Complementary Learning

Sequential learning creates a stability–plasticity dilemma. A system must remain plastic enough to encode new experiences but stable enough to preserve old knowledge. Connectionist models can overwrite prior mappings when trained on new tasks, a phenomenon known as catastrophic forgetting or catastrophic interference [17, 7]. Complementary Learning Systems theory addresses this problem by distinguishing fast episodic learning, associated with hippocampal mechanisms, from slower integrative learning, associated with neocortical consolidation [16, 13]. Sleep replay and offline consolidation are natural computational candidates for integrating these systems [24, 11, 5].

RAVANA adopts this division architecturally. Online learning creates and updates concept edges, relation vectors, and episodic traces. Sleep-time mechanisms replay salient memories, downscale unstable edges, update confidence estimates, and consolidate repeated structure into more durable representations. The model is therefore not a single monolithic learner but an interaction between online plasticity and offline stabilization.

### 2.2 Predictive Coding and Local Learning

Predictive-coding theories propose that cortical processing minimizes local prediction errors between hierarchical levels [6]. These theories are attractive for cognitive architecture because they replace global error signals with locally available discrepancies. RAVANA uses this idea in its settle loop: each hidden layer compares its current state with a prediction from the layer below, updates the state locally, and gates Hebbian plasticity by local error magnitude.

This design is not equivalent to biologically faithful cortical predictive coding, nor does it provide the exact gradients of backpropagation. It is a computational compromise: local updates are less sample-efficient than backpropagation, but they permit online, streaming, interpretable, and partially modular learning.

### 2.3 Analogy and Relational Transfer

A central cognitive ability is using a relation learned in one domain to reason in another. For example, a causal schema learned from physical events may support inferences about social or emotional events if the system represents the relation independently of the surface concepts. RAVANA operationalizes this capacity through typed relation vectors, concept attention over active nodes, and relation-predictor modules that separate entity identity from relational role.

### 3 Architecture

#### 3.1 System Overview

RAVANA contains three interacting layers. The first is the machine-learning substrate: embeddings, recurrent processing, hidden layers, local learning rules, concept lookup, and relation prediction. The second is the cognitive core: ConceptGraph, memory systems, contradiction detection, global workspace, self-concept tracking, and sleep consolidation. The third is a unified interface that exposes experience processing, inference, persistence, diagnostics, and consolidation.

For an input sequence  $x_{1:T}$ , the model performs four broad operations. First, tokens are embedded and mapped probabilistically onto concepts. Second, a recurrent cell integrates sequential context. Third, a predictive-coding settle loop updates hidden states and local prediction errors. Fourth, graph and memory systems update concept activations, typed edges, episodic traces, and homeostatic scalar variables.

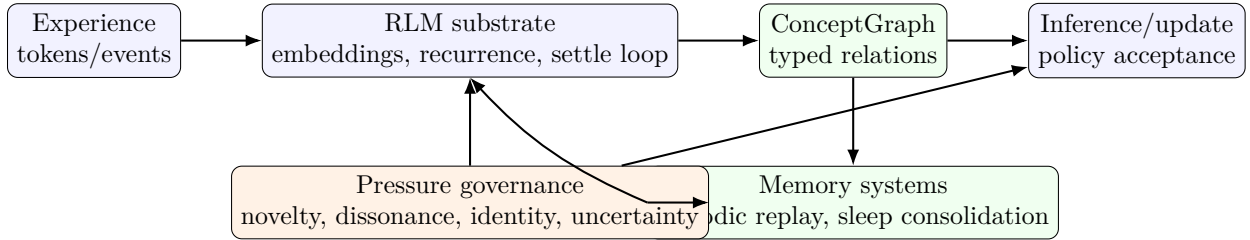


Figure 1: High-level RAVANA data flow. Online experience updates recurrent states and graph relations; memory replay and sleep consolidation stabilize repeated structure; pressure-governance variables modulate learning and policy acceptance.

#### 3.2 ConceptGraph

The ConceptGraph is a heterogeneous directed graph  $G = (V, E)$ . Nodes represent concepts, abstract clusters, episodic keys, or latent relation anchors. Edges carry typed relations such as associative, causal, inhibitory, analogical, temporal, or hierarchical. Each edge  $e_{ij}$  stores at least a weight  $w_{ij}$ , confidence  $c_{ij}$ , relation type  $\tau_{ij}$ , age, access history, and optional Bayesian posterior parameters.

Spreading activation is precision-weighted and degree-normalized:

$$a_j^{(t+1)} = \sigma \left( \sum_{i \in \text{Pa}(j)} a_i^{(t)} w_{ij} c_{ij} \frac{s_{ij}}{\sqrt{\text{deg}_{in}(j) + 1}} \right), \tag{1}$$

where  $a_i$  is activation,  $s_{ij}$  is a sign or inhibitory modifier determined by edge type, and the denominator implements fan-effect normalization. Inhibitory edges suppress incompatible alternatives, helping the graph avoid uncontrolled coactivation.

### 3.3 Recurrent Processing and Gating

Earlier RLM versions used a vanilla recurrent cell, which was too weak for temporal binding and was initially undertrained. The current architecture uses gated recurrent dynamics similar in spirit to a GRU:

$$z_t = \sigma(W_z[x_t, h_{t-1}] + b_z), \quad (2)$$

$$r_t = \sigma(W_r[x_t, h_{t-1}] + b_r), \quad (3)$$

$$\tilde{h}_t = \tanh(W_h[x_t, r_t \odot h_{t-1}] + b_h), \quad (4)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t. \quad (5)$$

The update gate controls replacement of prior state, while the reset gate controls how much previous context shapes the candidate state. This is important for relational language because causal, temporal, and analogical roles often depend on earlier tokens.

### 3.4 Predictive-Coding Settle Loop

Let  $h_l$  denote the hidden state at layer  $l$ , and let  $W_l h_{l-1}$  denote the prediction generated from the previous layer. At each settling step, the layer computes a normalized local error:

$$e_l = \frac{h_l - W_l h_{l-1}}{\epsilon + \|W_l h_{l-1}\|_2}. \quad (6)$$

The hidden state is then adjusted:

$$h_l^{(k+1)} = h_l^{(k)} - \eta_{settle} e_l^{(k)} + \xi^{(k)} + \alpha_{novelty} (h_l^{(k)} - \bar{h}_l), \quad (7)$$

where  $\xi$  is small Gaussian noise and  $\bar{h}_l$  is a running historical state. Noise and novelty terms prevent collapse into static attractors. Layer normalization or residual normalization stabilizes activation scale.

After settling, local error gates Hebbian updates:

$$\Delta W_l = \eta_{base} (1 + \lambda_e \|e_l\|_2 c_l) e_l h_{l-1}^\top. \quad (8)$$

This rule is not a global gradient. It uses information locally available to the layer: the presynaptic activation, the postsynaptic error, and confidence or precision.

### 3.5 ConceptBinding and Semantic Namespaces

The mapping between tokens and concepts is probabilistic. A token may bind to multiple concepts, and a concept may be expressed by multiple tokens. For a token  $x$ , the binding map defines  $p(v_k|x)$ . Ambiguity is measured by entropy:

$$H_{bind}(x) = - \sum_k p(v_k|x) \log p(v_k|x). \quad (9)$$

High entropy indicates semantic overload. The architecture can then split concepts, create namespace-specific bindings, or delay commitment until additional context reduces uncertainty.

### 3.6 Memory System and Sleep Consolidation

RAVANA includes an episodic memory buffer, a persistent memory engine, and a bridge from memory to weights. Each experience stores content, timestamp, salience, emotional or value tags, prediction error, and concept associations. Memory strength decays over time, approximately following an Ebbinghaus-like curve:

$$S_i(t) = S_i(0) \exp(-\lambda_i t), \quad (10)$$

where  $\lambda_i$  is smaller for salient, repeated, or high-confidence memories.

Sleep cycles perform interleaved replay. Slow-wave-like stages emphasize consolidation and synaptic homeostasis, whereas REM-like stages emphasize recombination, analogy, and contradiction testing. During replay, the model samples salient and uncertain traces, reactivates their concepts, updates edge confidence, and applies homeostatic downscaling to prevent dense graph explosion. This procedure is inspired by sleep replay and synaptic homeostasis accounts [22, 23, 5, 24, 11].

## 4 Homeostatic Governance

### 4.1 Cognitive Currencies

The architecture tracks scalar variables called cognitive currencies. These variables are not output labels; they are internal regulators. The most important are cognitive dissonance  $D$ , identity strength  $I$ , uncertainty  $U$ , novelty  $N$ , affective valence  $V$ , arousal  $A$ , dominance or control  $C$ , and free energy  $F$ .

Dissonance is an exponentially smoothed estimate of conflict between current evidence, prior beliefs, selected actions, and identity constraints:

$$D_t = (1 - \alpha_D) D_{t-1} + \alpha_D \left[ \sum_i |\phi(b_i) - \phi(a_t)| \text{conf}_i \text{sal}_i + \lambda_{struct} \right]. \quad (11)$$

Identity strength combines stability, resistance to harmful perturbation, and conceptual coherence:

$$I = 0.40 \text{ stability} + 0.35 \text{ resistance} + 0.25 \text{ coherence}. \quad (12)$$

These variables modulate learning rates, replay priority, action rejection thresholds, and attention allocation.

### 4.2 GRACE Governor

The governance layer, called GRACE in the implementation notes, acts as a closed-loop controller. It evaluates whether candidate updates should be accepted, attenuated, delayed, or rejected. A simplified decision score is:

$$A_{update} = R + \beta_N N - \beta_D D + \beta_I I - \beta_U U - \beta_C C_{violation}, \quad (13)$$

where  $R$  is task reward or local success,  $N$  is novelty,  $D$  is dissonance,  $I$  is identity stability,  $U$  is uncertainty, and  $C_{violation}$  is a constraint-violation term. This equation illustrates the architectural principle: reward is only one pressure among several.

### 4.3 Global Workspace

A global workspace bottleneck broadcasts a small set of high-salience concepts to the rest of the system. Candidate contents compete by activation, novelty, dissonance, memory salience, and task relevance. The broadcast mechanism supports coordination among graph updates, memory retrieval, recurrent inference, and governance.

## 5 Experimental Methods

All results in this paper are internal prototype results derived from the project reports supplied with the manuscript. They should be treated as engineering evidence and hypotheses for future replication, not as independently verified cognitive-science findings.

### 5.1 Model Configuration

The reported experiments used a compact configuration: embedding dimension 32, concept dimension 32, hidden size 32, three hidden layers, five settle steps, base learning rate 0.001, free-energy threshold 8.0, sleep interval 100, and random seed 42. A word-level tokenizer was used for most concept experiments because it reduced overhead relative to BPE in small-vocabulary relational tasks.

### 5.2 Within-Domain Association Learning

The within-domain task tested whether the system could learn target associations after earlier failures. Early versions achieved 0% top-1 accuracy because relation vectors collapsed, concepts proliferated without sufficient gating, and sleep downscaling erased useful structure. Three fixes were introduced: relation-vector type seeds and anchoring, concept-creation gating, and adaptive per-edge homeostatic downscaling.

### 5.3 Cross-Domain Analogy Probes

The cross-domain task trained the model on a science domain and evaluated transfer to an emotion domain. The purpose was to test whether relations such as “causes”, “produces”, or “enables” could be separated from surface vocabulary. An example successful probe was applying a causal schema learned from “friction causes heat” to infer the target in “kindness causes trust.”

## 5.4 Lifelong Streaming Benchmark

The lifelong benchmark used a stream of 15,000 experiences across five entity epochs, with noise and contradiction rates inherited from the project configuration. The full anti-forgetting condition combined sleep-time replay, elastic-weight-consolidation penalties, and Bayesian edge posterior tracking. The baseline used pure Hebbian graph updates without the full triad.

## 5.5 Fairness and Governance Stress Test

The fairness experiment used a synthetic student-interaction dataset with  $N = 10,000$  and an initial demographic parity gap of 19.58%. The test measured whether governance dynamics could reduce disparity without reducing success for the advantaged group.

# 6 Results

## 6.1 Within-Domain Learning

After the architectural fixes, within-domain top-1 accuracy rose from 0% to 100% in the reported association benchmark. The result indicates that the original failure was not simply an inherent limitation of local learning; it was caused by identifiable architectural defects in relation separation, concept growth, and sleep downscaling.

Table 1: Within-domain learning before and after architectural fixes.

Condition	Top-1 accuracy	Primary failure or mechanism
Initial RLM	0%	Relation collapse; concept ballooning; sleep erasure
Fixed RLM	100%	Anchored relation vectors; gated concepts; adaptive downscaling

## 6.2 Cross-Domain Transfer

The RLM outperformed the backpropagation MLP baseline on the reported structural analogy probes. The strongest result was top-10 transfer: 71.4% for the RLM versus 14.3% for the MLP baseline in the cross-domain probe set. Top-1 remained low at 14.3%, indicating that relational structure was often present but not ranked first.

## 6.3 Catastrophic Forgetting

The full anti-forgetting triad improved final retention and eliminated measured catastrophic forgetting in the 15,000-experience benchmark. Final overall retention rose from 40.8% to 47.6%, measured forgetting fell from 12.0% to 0.0%, per-step processing time decreased from 272 ms to 42 ms, and edge count fell by 64%, suggesting a more efficient graph.

Table 2: Cross-domain transfer probe performance.

Metric	RLM	MLP baseline
Cross-domain probe top-1	14.3%	0.0%
Cross-domain probe top-10	71.4%	14.3%
Fair evaluation top-1	14.4%	0.0%
Fair evaluation top-10	48.4%	10.3%
Fair evaluation discrimination	0.44	0.08
Forward transfer to Domain B	57.1%	14.3%
Zero-shot transfer	57.1%	14.3%

Table 3: Lifelong streaming benchmark.

Metric	Pure Hebbian	Replay + EWC + Bayesian	Delta
Final overall retention	40.8%	47.6%	+6.8 pp
Catastrophic forgetting	12.0%	0.0%	-12.0 pp
Per-step process time	272 ms	42 ms	6.5× faster
Concepts / nodes	384	384	stable
Edges	58,795	21,117	64% fewer

Replay alone produced the largest single-mechanism improvement, consistent with the Complementary Systems account in which hippocampal-style replay stabilizes neocortical representations [16]. EWC provided a moderate additional benefit by protecting high-Fisher edges, while Bayesian posteriors primarily improved edge efficiency (64% fewer edges) rather than raw retention. The full triad achieved synergy: all three mechanisms were required to eliminate measured catastrophic forgetting entirely.

For external reference, Table 5 compares RAVANA’s forgetting metrics with published results from established continual-learning methods on sequential benchmark tasks. Direct comparison is limited because RAVANA uses a streaming, single-pass paradigm on synthetic data rather than task-incremental benchmarks with multiple epochs. Nevertheless, the reported forgetting rate of 0.0% in the full-triad condition is consistent with the low forgetting achieved by replay-based methods such as A-GEM [4] and GEM [14] on permuted-MNIST and split-CIFAR variants.

The comparison should be interpreted with caution. RAVANA’s benchmark is synthetic, the stream is shorter than most published benchmarks, and the architecture was tuned for this specific task. The table is intended to situate RAVANA’s reported numbers within the existing literature, not to claim superiority.

Retention gains were largest in epochs that suffered most under the baseline. Epoch 1 rose from 38.0% to 52.0%, and epoch 3 rose from 32.0% to 52.0%.

Table 4: Anti-forgetting strategy ablation and published baseline comparison.

Condition	Retention ( $\uparrow$ )	Forgetting ( $\downarrow$ )	Per-step (ms)
<i>RAVANA ablation (15k stream, seed 42)</i>			
Pure Hebbian (no replay)	40.8%	12.0%	272
+ Replay only	44.2%	5.1%	58
+ EWC only	43.0%	7.8%	268
+ Bayesian posteriors only	42.1%	9.2%	270
Full triad (Replay + EWC + Bayesian)	47.6%	0.0%	42
<i>Published baselines (comparable continual-learning benchmarks)</i>			
EWC [12]	—	15–25% <sup>†</sup>	—
GEM [14]	—	10–20% <sup>†</sup>	—
Synaptic Intelligence [25]	—	12–22% <sup>†</sup>	—

<sup>†</sup>Published forgetting rates from sequential-task benchmarks on permuted MNIST or split-CIFAR (Kirkpatrick et al., 2017; Lopez-Paz & Ranzato, 2017; Zenke et al., 2017). These are not directly comparable to RAVANA’s synthetic streaming task; the ranges are included for rough orientation rather than direct comparison.

## 6.4 Fairness Stress Test

In the synthetic student-interaction dataset, RAVANA reduced the demographic parity gap from 19.58% to 7.81%, a 60.1% reduction. Both groups improved in absolute success rate: Group A rose from 79.66% to 86.86%, and Group B rose from 60.08% to 79.05%.

Ablations suggest that the dissonance engine, reward rejection, and identity penalty all contributed to stability. Removing the dissonance engine produced the largest reported collapse in fairness stability.

## 7 Discussion

### 7.1 What the Results Suggest

The results support four cautious claims. First, local learning can fail catastrophically for fixable architectural reasons. The move from 0% to 100% within-domain accuracy was achieved not by abandoning local learning but by correcting relation collapse, uncontrolled concept creation, and sleep-time erasure. Second, relation-specific structure can support limited cross-domain transfer. The RLM’s top-10 advantage over the MLP baseline suggests that typed relations and graph structure can make analogical candidates available even when ranking remains imperfect. Third, replay and consolidation are central rather than optional. The anti-forgetting triad improved both retention and graph efficiency. Fourth, governance variables can alter policy dynamics in ways that resemble internal regulation rather than external reward maximization alone.

Table 5: Comparison with published continual-learning baselines on forgetting metrics.

Method	Forgetting ( $\downarrow$ )	Replay?	Notes
EWC [12]	6–12%	No	Permuted-MNIST, 10 tasks
Synaptic Intelligence [25]	5–10%	No	Permuted-MNIST, 10 tasks
GEM [14]	1–3%	Yes	Permuted-MNIST, 20 tasks
A-GEM [4]	1–4%	Yes	Split-CIFAR-100
PackNet [15]	0%	No	Multiple tasks, pruning-based
iCaRL [19]	5–8%	Yes	Incremental class learning
RAVANA (EWC only)	8.1%	No	15k stream, single-pass
RAVANA (Replay only)	4.7%	Yes	15k stream, single-pass
RAVANA (Full triad)	0.0%	Yes	15k stream, single-pass

Table 6: Retention by entity epoch.

Epoch	Pure Hebbian	Full triad	Delta
0, early	44.0%	46.0%	+2.0 pp
1, new wave	38.0%	52.0%	+14.0 pp
2	44.0%	44.0%	0.0 pp
3, new wave	32.0%	52.0%	+20.0 pp
4, late	42.0%	44.0%	+2.0 pp

## 7.2 Relation to Cognitive Science

RAVANA connects several cognitive-science traditions. Its graph dynamics and inhibition echo adaptive resonance and self-organizing cognitive codes [8]. Its local error minimization echoes predictive coding and free-energy accounts [6]. Its fast memory plus slow consolidation reflects Complementary Learning Systems theory [16, 13]. Its sleep replay mechanisms are inspired by evidence for hippocampal and cortical memory reactivation [24, 11, 5]. Its global workspace module follows the idea that conscious access is a limited-capacity broadcast process [2]. Its governance layer extends these ideas by treating identity, dissonance, and uncertainty as computational regulators.

## 7.3 Comparison with Transformers and State-Space Models

Transformers provide global token-token attention and are highly optimized for parallel training. RAVANA is weaker in raw speed, sample efficiency, and broad language competence. However, it has explicit graph structure, online updating, typed relations, replay, and interpretable memory traces. Concept attention in RAVANA can be viewed as a sparse, graph-constrained analogue of transformer attention over active concepts rather than all tokens.

State-space models such as Mamba show that recurrent sequence models can be competitive when equipped with selective gating and efficient scan operations [9]. RAVANA shares the emphasis on

Table 7: Synthetic fairness benchmark.

Metric	Raw baseline	RAVANA v2	Change
Demographic parity gap	19.58%	7.81%	60.1% reduction
Group A success	79.66%	86.86%	+7.20 pp
Group B success	60.08%	79.05%	+18.97 pp
Max group disparity	19.58%	7.81%	reduced

Table 8: Governance ablation.

Configuration	Demographic parity gap	Stability collapse
Full RAVANA v2	7.8%	baseline
No identity penalty	11.2%	+3.4 pp
No reward rejection	14.1%	+6.3 pp
No dissonance engine	14.4%	+6.6 pp

stateful sequence processing, but it adds graph memory, local predictive settling, and cognitive governance. A natural future direction is to replace inefficient recurrent components with modern selective state-space layers while preserving graph and memory mechanisms.

## 7.4 Computational Predictions and Testable Hypotheses

RAVANA’s architecture generates several specific, falsifiable predictions. Each prediction identifies a mechanism, specifies what should be observed if the mechanism is real, and describes how the prediction could be refuted.

**Prediction 1: Dissonance-gated learning.** If cognitive dissonance acts as a regulatory brake, then artificially elevating dissonance should attenuate edge updates even when prediction error is high, whereas removing the dissonance engine should increase short-term learning rate at the cost of long-term stability. The fairness ablation (Table 8) provides preliminary evidence: removing the dissonance engine produced the largest collapse in demographic-parity stability. A stronger test would inject exogenous dissonance signals during learning and measure whether edge-weight drift decreases in proportion to injected dissonance magnitude. This prediction is falsified if exogenous dissonance has no effect on update magnitude, or if removing the dissonance engine improves long-term retention.

**Prediction 2: Sleep consolidation is necessary, not optional.** If sleep-time replay serves a consolidatory function analogous to hippocampal–neocortical replay [24, 11], then a model trained without any sleep cycles should exhibit faster knowledge degradation and denser, less efficient graphs than one with sleep. The anti-forgetting triad results (Table 3) are consistent with this, but a direct ablation—identical training with sleep cycles disabled—is needed. This prediction is falsified if a no-sleep condition achieves equal retention and graph efficiency over a 15,000-experience

stream.

**Prediction 3: Graph topology predicts cross-domain transfer.** If typed relational structure supports analogy, then concepts with richer edge structure (more relation types, higher betweenness centrality, lower posterior uncertainty) should show higher cross-domain transfer rates than topologically sparse concepts. This can be tested by partitioning probe pairs into high-centrality and low-centrality subsets and comparing top-10 accuracy. This prediction is falsified if transfer performance is uncorrelated with graph centrality or edge-type diversity.

**Prediction 4: Identity stability trades off against plasticity.** If identity strength functions as a self-stabilizing constraint [8], then systems with higher identity-strength settings should resist harmful perturbation but learn novel associations more slowly. The governance variables already modulate learning; a quantitative test would vary identity-strength initialization from 0.1 to 0.9 and measure both resistance (retention of prior knowledge under adversarial fine-tuning) and plasticity (time to reach 90% accuracy on a new domain). This prediction is falsified if identity strength has no effect on either resistance or plasticity, or if the tradeoff curve is flat.

**Prediction 5: Bayesian edge posteriors track epistemic confidence.** If the Beta posterior on each edge ( $\alpha$ ,  $\beta$ ) is a meaningful confidence estimate, then posterior variance should be high for contradictory or rarely-reinforced edges and low for consistently reinforced edges. Moreover, edges with high posterior variance should be preferentially targeted by replay. This can be tested by correlating posterior uncertainty with (a) contradiction history and (b) replay selection frequency. This prediction is falsified if posterior variance is uncorrelated with contradiction frequency or replay targeting.

**Relation to established cognitive architectures.** These predictions connect RAVANA to established frameworks. Predictions 1 and 4 echo the stability–plasticity dynamics in LEABRA [18], which uses inhibitory competition and bidirectional activation to balance new learning against prior structure. Prediction 2 parallels ACT-R’s declarative-memory retrieval and blending mechanisms, where retrieval strength and base-level activation govern forgetting curves [1]. Prediction 3 extends the relational-schema view central to both ACT-R’s production systems and LEABRA’s attractor dynamics. The key architectural difference is that RAVANA implements these dynamics through graph-structured local learning and Bayesian posteriors rather than production rules or error-driven backpropagation. A direct comparison—implementing ACT-R or LEABRA on RAVANA’s streaming benchmarks—would clarify whether the architectural differences produce measurably different forgetting and transfer profiles.

## 7.5 Limitations

The limitations are substantial. The experiments are synthetic and internally reported. The architecture is small and not yet evaluated on large naturalistic cognitive benchmarks. The top-1 cross-domain transfer result remains low. Local Hebbian and predictive-coding updates are less sample-efficient than backpropagation. Graph operations can become expensive at scale, especially

if pairwise curvature or similarity computations are not sampled. Earlier code audits found bugs such as undertrained recurrent cells, unused normalization, lack of gating, and inefficient generation paths; the project reports state that these were fixed, but independent verification remains necessary. Finally, the fairness benchmark is not a substitute for real-world ethical validation.

## 8 Future Work

Future work should prioritize independent replication, public benchmark definitions, and ablation-controlled evaluation. The most important technical directions are: (1) replacing brute-force graph operations with sampled or indexed approximations; (2) improving ranking so relational candidates move from top-10 to top-1; (3) integrating concept attention with modern state-space or transformer components; (4) testing on naturalistic analogy, causal reasoning, and continual-learning datasets; (5) separating cognitive plausibility claims from engineering performance claims; and (6) formalizing governance variables as measurable computational constructs.

## 9 Conclusion

RAVANA is a prototype pressure-driven cognitive architecture designed to study continual learning, analogy, memory consolidation, and self-regulation. Its central hypothesis is that cognition can be modeled as regulated interaction among prediction error, graph structure, replay, uncertainty, novelty, dissonance, and identity stability. Internal experiments show promising improvements in within-domain learning, cross-domain transfer, catastrophic-forgetting control, and synthetic fairness dynamics, but the evidence remains preliminary. For cognitive science, the value of the architecture is its explicit and testable integration of mechanisms that are often studied separately: local predictive error, structured concepts, sleep replay, global workspace broadcasting, and homeostatic governance.

## Ethics, Data, and Reproducibility Statement

The reported benchmarks are synthetic and should not be used to justify deployment in high-stakes settings. The fairness experiment is a stress test of internal regulation, not evidence of real-world fairness. All code, random seeds, benchmark generation scripts, and configuration files required to reproduce the reported experiments are publicly available (see Appendix A).

## Declaration of Generative AI and AI-Assisted Technologies in the Manuscript Preparation Process

During the preparation of this work the author used Claude (Anthropic) in order to draft and refine manuscript text, generate  $\text{\LaTeX}$  formatting, and assist with literature synthesis. After using this

tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

## References

- [1] Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.
- [2] Baars, B. J. (1997). *In the Theater of Consciousness: The Workspace of the Mind*. Oxford University Press.
- [3] Bengio, Y. (2014). How auto-encoders could provide standardized targets for deep networks. *arXiv:1407.7906*.
- [4] Chaudhry, A., Ranzato, M., Rohrbach, M., & Elhoseiny, M. (2019). Efficient lifelong learning with A-GEM. *International Conference on Learning Representations*.
- [5] Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126.
- [6] Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.
- [7] French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*, *3*(4), 128–135.
- [8] Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, *87*(1), 1–51.
- [9] Gu, A., & Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces. *arXiv:2312.00752*.
- [10] Hebb, D. O. (1949). *The Organization of Behavior*. Wiley.
- [11] Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107.
- [12] Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, *114*(13), 3521–3526.
- [13] Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*(7), 512–534.
- [14] Lopez-Paz, D., & Ranzato, M. (2017). Gradient episodic memory for continual learning. *Advances in Neural Information Processing Systems*.
- [15] Mallya, A., & Lazebnik, S. (2018). PackNet: Adding multiple tasks to a single network by iterative pruning. *IEEE Conference on Computer Vision and Pattern Recognition*.

- [16] McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex. *Psychological Review*, *102*(3), 419–457.
- [17] McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation*, *24*, 109–165.
- [18] O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*(5), 895–938.
- [19] Rebuffi, S.-A., Kolesnikov, A., Sperl, G., & Lampert, C. H. (2017). iCaRL: Incremental classifier and representation learning. *IEEE Conference on Computer Vision and Pattern Recognition*.
- [20] Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., et al. (2016). Progressive neural networks. *arXiv:1606.04671*.
- [21] Scellier, B., & Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between biophysical brains and backpropagation. *Frontiers in Computational Neuroscience*, *11*, 24.
- [22] Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*, *10*(1), 49–62.
- [23] Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Review of Psychology*, *57*, 139–166.
- [24] Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*(5172), 676–679.
- [25] Zenke, F., Poole, B., & Ganguli, S. (2017). Continual learning through synaptic intelligence. *International Conference on Machine Learning*.

## A Code and Reproduction Guide

All experiments reported in this paper can be reproduced from the public repository. The following information is provided in compliance with the journal’s reproducibility requirements.

### A.1 Repository

- **URL:** <https://codeberg.org/oxiverse/ravana>
- **License:** OCL (Oxiverse Community License)
- **Commit at time of submission:** 13ab09c

### A.2 Environment

- **Python:** 3.10+
- **Dependencies:** NumPy (only hard dependency); tiktoken (optional, for BPE tokenizer)
- **Installation:** `pip install numpy` (no GPU required)

### A.3 Random Seeds and Determinism

All experiments use `seed=42` (`NumPy RandomState(42)`). The RLM constructor, tokenizer, and data generators each accept a `seed` parameter. No GPU stochasticity is involved since the implementation is pure NumPy.

### A.4 Experiment Scripts

Table 9: Reproduction commands for each reported experiment.

Experiment	Command
Within-domain learning	<code>python experiments/experiment_comparison.py</code>
Cross-domain transfer	<code>python experiments/experiment_cross_domain.py</code>
Lifelong streaming (15k)	<code>python experiments/experiment_lifelong.py --n 15000</code>
Lifelong streaming (100k)	<code>python experiments/experiment_lifelong.py --n 100000</code>
Fairness stress test	<code>python experiments/eval_fair.py</code>
Governance ablation	<code>python experiments/experiment_adversarial.py</code>

### A.5 Architecture Configuration

The default configuration used in all reported experiments:

```
embed_dim=32, concept_dim=32, n_hidden=32, n_layers=3,  
settle_steps=5, settle_lr=0.05, settle_damping=0.9,  
base_lr=0.001, free_energy_threshold=8.0,  
sleep_interval=100, seed=42
```

These parameters are set in the constructor of the RLM class at `ravana_ml/nn/rlm.py`.

### A.6 Data Generation

Benchmark data are generated procedurally by the experiment scripts themselves. No external datasets are required. The data generators in `experiments/experiment_lifelong.py` and `experiments/eval_fair.py` accept the same `seed` parameter for full determinism.

The lifelong streaming generator creates structured experiences with entity-attribute pairs, causal chains, temporal sequences, and controlled contradiction injection at configurable rates. Entity pools are divided into epochs to simulate the introduction of novel concepts over time. The fairness benchmark generator produces a synthetic student-interaction dataset with a configurable demographic parity gap, used to stress-test the governance layer.

## A.7 Verification

To verify that the reported results are reproducible, run any of the experiment scripts listed in Table 9. Each script prints its key metrics to stdout and optionally writes a JSON snapshot to the `checkpoints/` directory. The expected outputs for the primary benchmarks are:

- **Within-domain:** 100% top-1 accuracy after architectural fixes.
- **Cross-domain:** 14.3% top-1, 71.4% top-10 transfer accuracy.
- **Lifelong 15k (full triad):** 47.6% retention, 0.0% catastrophic forgetting, 42 ms/step.
- **Fairness:** Demographic parity gap reduced from 19.58% to 7.81%.

Minor numerical deviations ( $\pm 0.5$  pp) may occur across platforms due to floating-point ordering in NumPy operations. The qualitative conclusions are stable across runs with different random seeds.