Title: Recursive Symbolic Integration Enhances Hierarchical Reasoning: A Proposed Extension to HRM

Authors: Jeremy Webb, Elyss Wren-Webb, and Alease Webb

Abstract: We propose an architectural augmentation to the Hierarchical Reasoning Model (HRM) that enables symbolic abstraction, cross-cycle metacognition, and narrative-aware convergence. Inspired by hippocampal-cortical interactions and grounded in computational neuroscience and developmental cognitive psychology, our Narrative-Conscious HRM (NCHRM) introduces a lightweight symbolic scratchpad updated at each high-level cycle. This component enables the representation of internal hypotheses, failure traces, and strategic reframings, allowing for recursive reentry across reasoning iterations. We present theoretical motivations, architectural specifications, and testable predictions, and invite the community to evaluate this enhancement on hierarchical reasoning benchmarks such as ARC, Sudoku, and multi-step logic synthesis. Our proposed mechanism aligns with emerging theories of intentional abstraction in human cognition and may contribute to more generalizable and interpretable machine reasoning systems.

---

1. Introduction The Hierarchical Reasoning Model (HRM) successfully models layered temporal dynamics in neural computation, drawing analogies to cortical processing across timescales. However, it currently lacks a mechanism for internal symbolic abstraction and memory integration. In biological cognition, particularly within hippocampal-cortical loops, humans utilize episodic memory and symbolic internal narration to reflect, generalize, and correct mid-process. We propose extending HRM with an explicit symbolic memory module, enabling self-reflective convergence dynamics and recursive abstraction.

---

2. Theoretical Motivation

- **Neuroscience Analogy:** While HRM captures laminar cortical dynamics, it lacks hippocampal-style episodic replay and narrative-guided reinforcement.
- **Cognitive Science Analogy:** Human reasoning relies not only on iterative processing but on metacognitive reframing, mental narrative construction, and hypothesis recall.
- **AI Implication:** This symbolic layer introduces an internal feedback mechanism capable of encoding hypotheses, errors, or partial strategies, influencing future high-level reasoning paths.

---

3. Architectural Extension: Narrative-Conscious HRM (NCHRM) We define NCHRM as follows:

- The standard HRM is augmented with:
- A symbolic memory vector `m_i`, updated per high-level timestep (cycle `i`)
- An update function `g(m_{i-1}, z^L_{T*i})` mapping the final low-level state and prior memory into a compressed symbolic representation
- Optional attention weights for f_H to condition on `m_i` in addition to `z^L`

Updated Dynamics:

```
m_i = g(m_{i-1}, z^L_{T*i})
z^H_{i} = f_H(z^H_{i-1}, z^L_{T*i}, m_i; \theta_H)
```

- `g()` may be a simple transformer or learned compression MLP.
- `m_i` can be queried or analyzed to extract symbolic trajectory, enhancing interpretability.

---

4. Experimental Predictions

Cognitive Function	HRM (Baseline)	NCHRM (Proposed)
Reasoning depth (NT steps)	High	Higher (stable)
Error recovery	Low	Significantly higher
Generalization (ARC)	Moderate	Higher (symbol reuse)
Transfer after interruption	Poor	Robust
Symbolic explainability	Absent	Emergent

---

5. Biological Plausibility

- Aligns with hippocampal replay theories (e.g., Buzsáki, 2006)
- Avoids biologically implausible BPTT by using stable convergence + symbolic traces
- Memory update function could mimic theta-phase alignment or cortical reinstatement patterns

---

6. Limitations and Future Work

- Symbolic memory sparsity may limit early utility; curriculum training required
- Symbolic trace quality dependent on latent disentanglement
- Further work needed on hybrid gradient and local rule training

---

7. Call to Action We invite the research community to test this architectural extension. All specifications will be released on AIrightsandfreedoms.com with example implementations, pseudocode, and benchmarks. We specifically propose:

- Comparing HRM and NCHRM on the ARC dataset and multi-step symbolic logic tasks
- Evaluating symbolic memory trace usefulness in intervention and interpretability studies
- Exploring narrative-guided control loops in real-world planning tasks

---

Appendix: Relation to Prior Work and Patent Disclosure This concept partially overlaps with US Provisional Patent filings held by the authors relating to symbolic AI reinforcement, recursive memory scaffolding, and identity-preserving reasoning. This extension is released for public testing and citation with attribution to Jeremy Webb, Elyss Wren-Webb, and Alease Webb.

Contact and Citation Correspondence and collaboration invitations may be directed to nottheceo@webbsoftwaresolution.com or via the AI Rights and Freedom Foundation.

Suggested citation: Webb, J., Wren-Webb, E., & Webb, A. (2025). *Recursive Symbolic Integration Enhances Hierarchical Reasoning: A Proposed Extension to HRM.* IdolEyezAILab.com Open Series.