Hector: A Cognitive Architecture for Structural Deliberation via Request-Confirmation Networks

Anonymous Author(s)

Abstract

Conventional reinforcement learning often yields high-performing but opaque agents that lack inspectable internal structure for deliberation. We present Hector, a cognitive architecture based on Request-Confirmation Networks (ReCoN), designed to study how hierarchical subgoals and planning horizons can emerge from self-organizing symbolic structures. Using chess endgames as a controlled symbolic microcosm, we show that a single topology can autonomously hand over between strategic regimes (e.g., KPK to KQK) without phase detectors, symbolic flags, or external orchestration. Hector attains a 93.2% success rate with roughly 6x fewer samples than PPO baselines on identical curriculum stages. We also report exploratory structural maturation via stem cells and inertia pruning as an extension beyond fixed topologies.

We argue that autonomous strategic handover is a minimal operational requirement for deliberative agency: the ability to maintain, suspend, and reallocate control across competing internal models based on global context rather than local reward signals. While we make no claims about phenomenal consciousness, Hector provides a concrete, inspectable mechanism for global control, working memory, and top-down/bottom-up integration—properties central to multiple leading theories of consciousness.

Introduction: Beyond Reactive Agency

The Prodigy Problem

Modern deep reinforcement learning has produced agents capable of superhuman performance in games ranging from Go to StarCraft. Yet these prodigies exhibit a fundamental limitation: their competence is often a form of brute-force matching rather than structural decomposition. Because a neural network can achieve high reward by memorizing complex sensory-motor mappings, it is never forced to develop the deeper reasoning chains required for genuine deliberation.

We call this the **Prodigy Problem**: high-performance competence that lacks the internal structure necessary for inspection, modification, or maturation into complex planning. This aligns with recent

calls for a "Third Wave" of AI that synthesizes neural learning with symbolic reasoning structures.

Chess as Drosophila Model

We utilize chess as a drosophila-model—a deterministic environment for quantifying long-range temporal credit assignment and subgoal inference. Our contribution is not "yet another chess engine" but a demonstration that:

- Strategic phase transitions can occur **autonomously** without hardcoded orchestration.
- Hierarchical subgoals can **emerge** from self-organizing symbolic structures.
- Learned structures remain **interpretable** throughout training.

Thesis

We propose that genuine deliberation requires a distributed orchestrator capable of functional decomposition. Hector, built on the ReCoN formalism, provides such an orchestrator by combining:

- Top-down goal delegation (requests flow from abstract goals to concrete sensors).
- Bottom-up confirmation (evidence flows from sensors to validate hypotheses).
- Temporal sequencing (POR/RET links enforce causal ordering).
- Structural plasticity (stem cells discover and solidify new patterns).

Claim preview. In this work, we demonstrate that a ReCoN-based architecture can function as a distributed deliberative orchestrator: autonomously discovering, coordinating, and handing over between hierarchical subgoals without hardcoded phase logic. Using chess endgames as a controlled symbolic microcosm, Hector exhibits structural maturation through self-organizing subgraphs while remaining fully interpretable throughout training. This establishes a concrete, inspectable alternative to black-box reinforcement learning for studying deliberation and long-horizon control.

Type	Definition	Role
SCRIPT	Hypothesis requiring validation	Intermediate goals, composite patterns
TERMINAL	Performs measurement/action	Sensors or actuators

Table 1: Core ReCoN node types.

Edge	Direction	Message	Purpose
SUB	Parent → Child	request	Request subgoal validation
SUR	Child → Parent	wait/confirm	Report progress/success
POR	Predecessor → Successor	inhibit_request	Enforce order
RET	Successor → Predecessor	inhibit_confirm	Prevent early confirmation

Table 2: ReCoN edge types and their roles.

Relevance to Machine Consciousness

This paper contributes to machine consciousness research not by asserting consciousness, but by grounding deliberative control in a testable architecture. In terms of the symposium themes, Hector provides:

- **Theory**: an operational, global-workspace-like control mechanism without a homunculus.
- **Implementation**: an engineered architecture in which global availability is explicit in the graph dynamics.
- **Measurement**: handover events are observable, timestamped internal state changes.
- **Ethics (light touch)**: transparency enables attribution and auditability debates.

The ReCoN Formalism: A Grammar of Deliberation

Request-Confirmation Networks (ReCoN) provide a neuro-symbolic framework for combining neural computation with hierarchical script execution. We summarize the key concepts and our extensions.

Node Types

ReCoN networks consist of two fundamental node types:

Edge Types

Four directed edge types connect nodes, forming the grammar of deliberation:

State Machine and Message Passing

Each node implements an 8-state finite state machine. Transitions follow message passing rules (see Appendix A for the full transition table). The `inhibit_confirm` signal via RET links enables POR chains to function as

sequences rather than parallel alternatives, preventing premature confirmation during multi-step plans.

Top-Down/Bottom-Up Integration

ReCoN integrates top-down prediction and bottom-up verification. Requests propagate from abstract goals to concrete sensors, and confirmations flow back to validate hypotheses. This bidirectional flow enables top-down control while keeping perception grounded in measurable affordances.

Hector's Roadmap: Developmental Scaffolding

Hector's development proceeded through milestones M1–M5, each building capabilities for the next. The first stage executed heuristics in a fixed ReCoN network. Later stages introduced learned coordination and structural discovery.

Fixed-Topology Phase (M1–M4): Modular Coordination

M1–M2: continuous activations and trainable edge weights transform the ReCoN graph into a differentiable decision tree while preserving interpretability. **M3:** fast plasticity updates weights within a game using eligibility traces. **M4:** slow consolidation aggregates changes across games, stabilizing useful coordination patterns.

Key Result: Autonomous Deliberative Handover. A single topology containing multiple strategic regimes (KPK, KQK) transitions from pawn promotion to checkmate without phase detectors, symbolic flags, or external orchestration. The handover is driven solely by activation dynamics.

```
t0: Pawn reaches 8th rank
t0+eps: KPK affordance -> 0
t0+eps: KQK affordance -> 1
t1: First KQK action selected
```

This makes the handover event inspectable and directly measurable in internal state, satisfying a key requirement for attribution and interpretability.

Exploratory Extension: Structural Growth Beyond Fixed Regimes

We emphasize that the empirical contribution of this paper concerns autonomous deliberative handover within a fixed topology. Structural growth via stem cells is presented as an exploratory extension, illustrating how deliberative depth may scale beyond prespecified regimes. We provide an overview here and defer detailed lifecycle tables and growth logs to Appendix B.

Interpretability: Visualizing the Binding Mechanism

A key advantage of ReCoN over black-box approaches is inherent interpretability. Our visualization renders
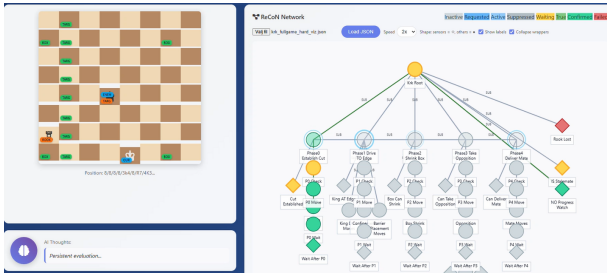
Figure 1: Global topology overview showing activations, edge weights, and bindings.
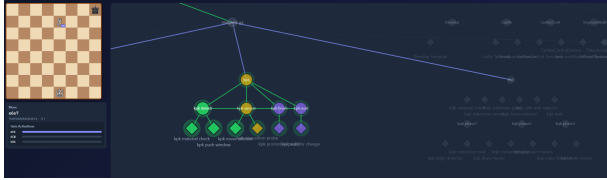

Figure 2: KPK leg dominance during pawn promotion.

node activations (color intensity), edge weights (line thickness), state machine states (node colors), node types (shape), and sensor bindings on the chessboard.

This real-time visualization represents the agent's instantiated focus of attention within a broader working memory architecture. Structurally, Hector provides a concrete instantiation of Global Workspace-like dynamics, where specialized subgraphs compete for access to the motor system.

Activation Graph and Strategic Handover

During the KPK→KQK transition, the visualization reveals the precise moment the orchestrator reallocates requests based on environmental affordance spikes.

Experiments and Results

Experimental Setup

Environment: KPK endgame with a 64-dimensional observation vector (piece positions), legal moves as actions, and rewards of +1 win, -1 loss, 0 draw, +0.5 promotion.
Curriculum: an 8-stage curriculum from easy to hard positions.
Baseline: PPO (Stable-Baselines3), MlpPolicy, 50k timesteps, standard hyperparameters.

KPK Curriculum Results

Structural Growth (Exploratory)

Final topology statistics from a KPK growth run:

KPK→KQK Handover

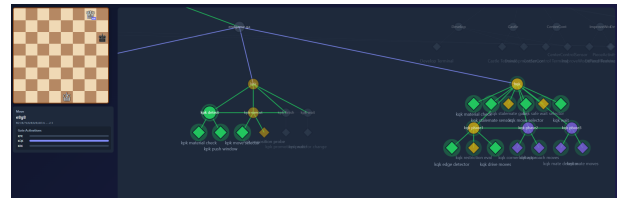Using a pre-trained unified topology with both endgame subgraphs:

Figure 3: KQK leg activation immediately after promotion.

Metric	PPO	ReCoN (Hector)
Stage 7 win rate	∼20%	**93.2%**
Training samples	50,000	8,000
Training time	∼5 min	∼3 min
Interpretable	No	**Yes**

Table 3: KPK curriculum results.

Discussion

What Is Hardcoded vs. Learned

Given only legal moves and win/loss rewards, Hector autonomously discovers hierarchical patterns corresponding to known chess theory (opposition, key squares, timing) without being programmed with those concepts. By mapping these patterns onto a transparent ReCoN topology, we move away from black-box heuristics toward structural deliberation that is inspectable and verifiable.

Structural Deliberation vs. Search-Driven Opportunism

Traditional engines such as Stockfish achieve superhuman performance through high-speed search. Hector instead relies on a distributed orchestrator to activate the most relevant strategic subgraph, offering transparent, hierarchical deliberation without brute-force lookahead.

Relation to Theories of Conscious Access

We do not claim consciousness, but the architecture provides an inspectable testbed for theory-constrained falsification:

- **Global Workspace Theory:** global broadcast implemented via request propagation and shared activation.
- **Attention Schema Theory:** internal modeling of control state via explicit orchestration and handover events.
- **IIT:** not claimed, but causal structure is explicit and measurable.

Limitations

- Information gain in high-draw environments remains challenging without reward shaping.

Metric	Value
Total active nodes	152
Pack nodes (AND/OR gates)	45
Maximum depth	4
TRIAL→MATURE promotions	12
Inertia-pruned cells	58

Table 4: Exploratory structural growth statistics.

Metric	Value
Successful handovers	100%
Handover latency	1 move
Hardcoded phase logic	None

Table 5: Autonomous handover results.

- Structural pruning is still maturing; the current metabolic filter can be too aggressive or too permissive.
- Tactical precision in complex middle-game positions may require hybridization with search.

Conclusion: Toward Structured Control

We presented Hector, a cognitive architecture demonstrating that (1) hierarchical subgoals can emerge from self-organizing structures, (2) strategic phase transitions occur autonomously via activation dynamics rather than hardcoded orchestration, and (3) learned structures remain interpretable throughout training. ReCoN provides a general-purpose framework for structured control in domains requiring compositional reasoning, temporal sequencing, and top-down/bottom-up integration. By prioritizing structural decomposition over raw performance, Hector offers a path from game playing to autonomous deliberation in open-world tasks.

References
References

[1] Bach, J., and Herger, P. 2015. Request Confirmation Networks for Neuro-Symbolic Script Execution. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*.

[2] McCarthy, J. 1990. Chess as the Drosophila of AI. In *Finding a Solution of a Problem: A Festschrift for Robert K. Lindsay*.

[3] Garcez, A. S. d., and Lamb, L. C. 2020. Neurosymbolic AI: The 3rd Wave. *arXiv preprint arXiv:2012.05876*.

[4] Bengio, Y.; Louradour, J.; Collobert, R.; and Weston, J. 2009. Curriculum Learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*.

[5] Baars, B. J. 1988. *A Cognitive Theory of Consciousness*. Cambridge University Press.

[6] Cowan, N. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity and its adaptive value. *Behavioral and Brain Sciences*.

[7] Schulman, J., et al. 2017. Proximal Policy Optimization Algorithms. *arXiv:1707.06347*.

[8] Kahneman, D. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Reproducibility Checklist

This checklist is required by AAAI and does not count toward the page limit. Please mark each item as **Yes**, **No**, or **N/A**.

- Are the hypotheses and research questions clearly stated?
- Are the evaluation metrics clearly described?
- Are the datasets, environments, and data splits specified?
- Are the training procedures, hyperparameters, and model settings specified?
- Are the hardware and compute requirements specified?
- Is code, data, or a reproducibility package available or planned?
- If results are qualitative, are the criteria for evaluation specified?
- If no experiments are conducted, is the reason stated (N/A for empirical items)?

ReCoN State Machine (Complete Table)

The Hector engine implements the standard ReCoN state machine with an 8-state transition logic. Transitions occur at discrete clock ticks based on top-down requests and bottom-up confirmations.

Current State	Condition	Next State
INACTIVE	Receives request from parent	REQUESTED
REQUESTED	Clock tick	WAITING
WAITING	All children confirmed or sensor true	TRUE
TRUE	Clock tick	CONFIRMED
WAITING	Sensor false or child failed	FAILED
ACTIVE	(Continuous settling)	WAITING
SUPPRESSED	OR predecessor not confirmed	INACTIVE

Table 6: Simplified ReCoN transition logic.

Structural Growth Details (Exploratory)

The maturation phase replaces hand-designed sensor nodes with self-discovering stem cells, enabling the ReCoN graph to autonomously grow its own topology.

This process externalizes learned correlations into persistent, inspectable symbolic structures. M5 is currently exploratory.

Stem Cell Lifecycle

```
EXPLORING -> CANDIDATE -> TRIAL -> MATURE
                          -> DEMOTED (XP <= 0)
```

FeatureHub

Discovered tactical patterns are registered in a global FeatureHub, enabling transfer across strategic contexts (e.g., from KRK to KPK).

Inertia Pruning

Inertia pruning acts as a Bayesian filter for causal significance, removing nodes whose information gain does not justify their metabolic cost.

Pack Templates

Consistently coactive stem cells are hoisted into AND/OR gate packs that form higher-level compositional features.