
Eidetic AI:

A Framework for Lifelong Machine Pedagogy via
Quantum-Inspired Conceptual Inference and
Co-Evolutionary Curriculum Generation

Your Name

August 8, 2025

*Submitted to the Department of Electrical Engineering and Computer
Science*

Massachusetts Institute of Technology

Abstract

We introduce **Eidetic AI**, a new architectural paradigm for artificial tutors that fundamentally transcends the limitations of autoregressive models. We posit that true machine pedagogy requires not just knowledge retrieval, but conceptual synthesis. Our framework achieves this through three core innovations. First, we formalize the *Quantum Conceptual Engine* (QCE), an inference mechanism that generates explanations by solving for the ground state of a dynamically constructed conceptual Hamiltonian, ensuring maximal coherence. Second, we propose *ASCENDANT-K*, a multi-agent reinforcement learning (MARL) framework where teacher and critic agents co-evolve toward a Nash Equilibrium, producing curricula of unparalleled robustness. This process is fueled by a novel pipeline that distills confusion patterns from large-scale data into compact, adversarial student models. Third, we architect a *Hierarchical Associative Memory Graph* with Hebbian update rules, enabling lifelong learning and adaptation over timescales of years without catastrophic forgetting. We will validate this framework through a rigorous suite of automated simulations and a capstone human randomized controlled trial, demonstrating superior pedagogical utility, transfer learning, and computational efficiency against state-of-the-art baselines.

Contents

1	Introduction	4
2	Core Intellectual Contributions	5
3	Formal Problem Statement	6
4	Mathematical Framework	8
4.1	The Quantum Conceptual Engine (QCE)	8
4.1.1	Continuous relaxation	8
4.2	ASCENDANT-K MARL Framework	9
4.3	Hierarchical Associative Memory	9
5	Theoretical Roadmap	10
6	Methods, Systems, and Implementation Plan	12
6.1	Thrust 1: QCE Core Implementation & Prototype	12
6.2	Thrust 2: ASCENDANT-K Training Environment and Baselines . .	13
6.3	Thrust 3: Hierarchical Associative Memory & Human Validation .	13
7	Evaluation Framework & Experiments	14
7.1	Datasets and Benchmarks	14
7.2	Metrics	14
7.3	Key Experiments	15

List of Figures

List of Tables

8.1 High-level timeline and milestones. 16

Chapter 1

Introduction

The dominant paradigm in artificial intelligence, particularly in large language models (LLMs), is autoregressive prediction [1]. While extraordinarily powerful for tasks involving sequence completion and pattern matching, this approach is fundamentally one of mimicry, not synthesis. In the domain of education, this results in AI tutors that are sophisticated information retrievers [lewis2020retrieval] and paraphrasers, yet lack the genuine pedagogical insight required to invent a truly novel analogy or re-frame a concept from first principles.

This thesis introduces **Eidetic AI**, a framework designed to move beyond mimicry towards true machine pedagogy. Our approach is founded on the premise that a pedagogical explanation is not a sequence of high-probability tokens, but rather a coherent, self-consistent constellation of concepts. We formalize this notion by recasting explanation generation as a problem in quantum-inspired optimization. The core of our system, the *Quantum Conceptual Engine* (QCE), formulates an explanation as the ground state of a conceptual Hamiltonian—an energy function whose minimization corresponds to maximizing conceptual coherence, inspired by Hopfield networks [2] and Ising models [lucas2014ising].

This inference mechanism is situated within a broader ecosystem designed for robust and lifelong learning. The ASCENDANT-K framework leverages co-evolutionary dynamics from multi-agent reinforcement learning [4] to harden curricula against student misconceptions, akin to a pedagogical Generative Adversarial Network [goodfellow2014generative]. Finally, a hierarchical associative memory graph with Hebbian update rules provides a substrate for long-term adaptation and personalization. This document lays out the full theoretical foundation, algorithmic approach, and evaluation plan.

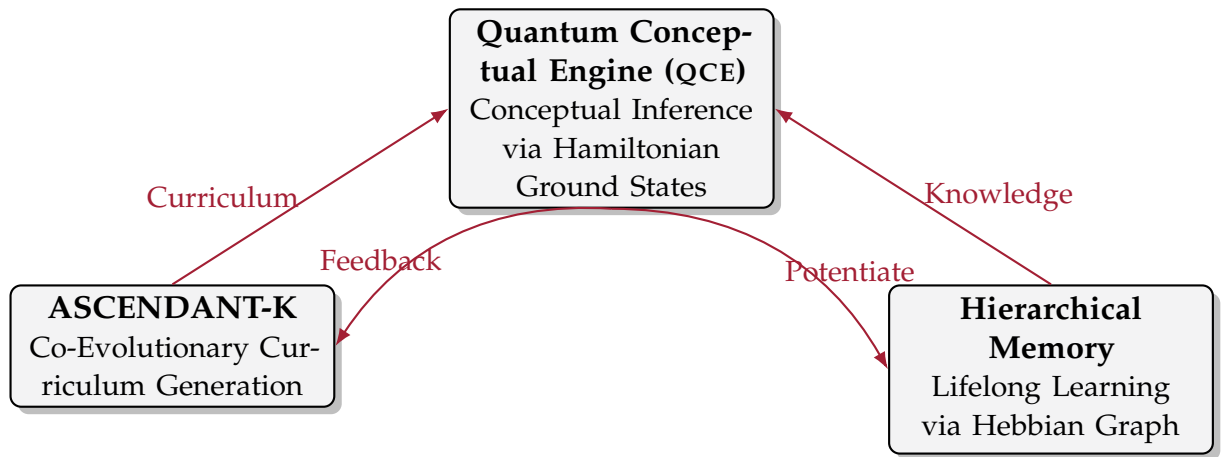


Figure 1.1: The core architectural pillars of the **Eidetic AI** framework.

Chapter 2

Core Intellectual Contributions

This thesis will establish and formally validate the following intellectual contributions:

1. **A Formalism for Pedagogical Coherence.** Define and operationalize pedagogical coherence as the ground state energy of a higher-order *conceptual Hamiltonian*, providing a novel, optimizable objective distinct from token-level log-likelihood.
2. **An Approximation Guarantee for QCE.** Prove that for a chosen class of sparse Hamiltonians, a continuous relaxation plus projected gradient solver achieves a provable ε -approximation guarantee.
3. **Convergence Guarantees for Co-evolutionary Curriculum.** Characterize ASCENDANT-K as a general-sum game and prove conditions under which a population-based training algorithm converges to a correlated equilibrium.
4. **A High-Performance QCE Implementation.** Deliver an open-source C++/CUDA QCE-Core with measurable performance targets (latency and memory budgets).
5. **Hierarchical Associative Memory for Lifelong Pedagogy.** Design a multi-layer memory graph with Hebbian potentiation that supports long-term personalization and prevents catastrophic forgetting.
6. **End-to-End Empirical Validation.** Demonstrate improved pedagogical outcomes in simulations and a human RCT against GPT-class LLMs with RAG and standard intelligent tutoring systems.

Chapter 3

Formal Problem Statement

We state three formal problems corresponding to the three thrusts of the thesis.

Definition 3.1 (Quantum-Inspired Conceptual Inference). Given a student’s latent knowledge state $s \in \mathcal{S}$ and a universe of concepts $\mathcal{C} = \{c_1, \dots, c_N\}$, construct a higher-order energy function (the *conceptual Hamiltonian*) $H(x; s, \mathcal{C})$, where $x \in \{0, 1\}^N$ is a binary indicator vector representing an explanation. The inference task is to find the ground state:

$$x^* = \arg \min_{x \in \{0, 1\}^N} H(x; s, \mathcal{C}).$$

Hypothesis 3.1 (QCE Superiority). For a fixed computational budget, explanations derived from the QCE ground state will yield a $> 20\%$ increase in measured conceptual coherence and a statistically significant improvement in simulated student learning gain (Normalized Learning Gain; Cohen’s $d \geq 0.3$, $p < 0.05$) compared to explanations from state-of-the-art autoregressive LLMs with RAG.

Definition 3.2 (Co-Evolutionary Curriculum Generation). Let the teacher be an agent T_θ and let the critic (distilled student/adversary) be S_ψ . Define a two-player general-sum game with payoffs derived from student improvement and fragility. The aim is to find an equilibrium (θ^*, ψ^*) that yields curricula robust to adversarially distilled student failure modes.

Hypothesis 3.2 (ASCENDANT-K Efficiency). Curricula generated by ASCENDANT-K will enable simulated student agents to achieve concept mastery with a $> 25\%$ reduction in Time-to-Mastery (TTM) and a $> 15\%$ increase in long-term retention compared to expert-designed syllabi and SOTA adaptive tutoring baselines.

Definition 3.3 (Lifelong Pedagogical Memory). Design a memory architecture \mathcal{M} as a hierarchical associative graph that stores and retrieves personalized

pedagogical strategies and supports updates via biologically-inspired potentiation rules while avoiding catastrophic forgetting.

Hypothesis 3.3 (Memory Graph Robustness). A hierarchical associative memory graph with Hebbian updates will maintain $> 95\%$ efficacy on previously mastered personalized teaching strategies after one simulated year of interaction and demonstrate $> 50\%$ improvement in zero-shot transfer performance compared to standard RAG pipelines.

Chapter 4

Mathematical Framework

4.1 The Quantum Conceptual Engine (QCE)

We model an explanation as a binary vector $x \in \{0, 1\}^N$. The conceptual Hamiltonian is an order- p polynomial:

$$H(x; s) = \sum_i a_i(s) x_i + \sum_{i < j} b_{ij}(s) x_i x_j + \sum_{i < j < k} T_{ijk}(s) x_i x_j x_k + \dots \quad (4.1)$$

The coefficients are outputs of a learned *Hamiltonian Constructor* network $M_\Phi(s, \mathcal{C})$. Inference is the combinatorial optimization:

$$x^* = \arg \min_{x \in \{0, 1\}^N} H(x; s),$$

which is QUBO-like and generally NP-hard [lucas2014ising].

4.2 ASCENDANT-K MARL Framework

The teacher T_θ has policy $\pi_T(a \mid s, c)$ producing actions a , and the critic S_ψ has policy $\pi_S(\delta \mid s, a)$ producing failure modes δ . The teacher's objective is:

$$\max_{\theta} \mathbb{E}_{s, c \sim \mathcal{D}} [\mathcal{R}_{\text{learn}}(s, a) - \lambda \mathcal{R}_{\text{fragility}}(s, a, \delta)], \quad (4.2)$$

where $a \sim \pi_T(\cdot \mid s, c)$ and $\delta \sim \pi_S(\cdot \mid s, a)$. Training uses population-based co-evolution.

4.3 Hierarchical Associative Memory

The memory is a directed weighted graph $\mathcal{G} = (V, E, W)$. After a successful interaction yielding utility $U > \tau$, activated node pairs (i, j) receive a potentiation update:

$$\Delta w_{ij} = \eta \cdot U \cdot \text{act}(c_i) \text{act}(c_j) - \gamma w_{ij}, \quad (4.3)$$

with learning rate η and decay $\gamma > 0$ ensuring boundedness.

Chapter 5

Theoretical Roadmap

Theorem 5.1 (QCE Solver Approximation Guarantee). *Let $H(x)$ be an order- p Hamiltonian whose interaction tensors satisfy a k -sparsity condition and suppose the Hamiltonian exhibits a spectral gap $\Delta > 0$. For the continuous relaxation \tilde{H} with a suitable convexifying regularizer, our projected gradient solver yields a solution \hat{x} satisfying*

$$H(\hat{x}) - H(x^*) \leq \varepsilon(N, k, p, \Delta),$$

where $\varepsilon(\cdot)$ is a function whose explicit form will be derived.

Proposition 5.1 (ASCENDANT-K Convergence). *Under assumptions of bounded rewards and Lipschitz continuity of the expected payoff, the empirical distribution over policies in our population-based co-evolutionary training converges to a correlated equilibrium of the underlying general-sum game.*

Lemma 5.1 (Memory Graph Stability). *The Hebbian update rule (4.3) with decay $\gamma > 0$ and bounded utilities ensures the total weight of the memory graph remains bounded for all time, preventing unbounded potentiation.*

Chapter 6

Methods, Systems, and Implementation Plan

6.1 Thrust 1: QCE Core Implementation & Prototype

High-Performance Solvers (QCE-Core). A dedicated C++/CUDA library implementing parallel simulated annealing and sparse-tensor iterative methods. PyBind11 bindings provide a Python interface. **Performance Targets.** Median inference latency ≤ 150 ms for $N = 256$ concepts on a single consumer-grade GPU.

6.2 Thrust 2: ASCENDANT-K Training Environment and Baselines

Student Confusion Distillation. Process large anonymized datasets (e.g., ASSISTments) to construct compact student models via contrastive learning. **Baselines.** GPT-4-class LLM + RAG, vector-store retrieval, and standard ITS algorithms.

6.3 Thrust 3: Hierarchical Associative Memory & Human Validation

Memory implementation. Multi-layer graph in a high-performance graph database. **Human RCT plan.** An IRB-approved, three-arm (Eidetic, GPT-4 tutor, control)

randomized controlled trial ($N \approx 200$) will measure NLG and retention.

Chapter 7

Evaluation Framework & Experiments

7.1 Datasets and Benchmarks

ASSISTments, EdNet, and synthetic curricula for Newtonian physics and linear algebra.

7.2 Metrics

Normalized Learning Gain (NLG), Time-to-Mastery (TTM), Retention Index, Conceptual Novelty Score, and system performance metrics (latency, memory).

7.3 Key Experiments

1. **Offline Head-to-Head:** QCE vs GPT-4+RAG on a 600-prompt benchmark.
2. **Simulated Curriculum Training:** ASCENDANT-K vs baselines on TTM and retention.
3. **Human RCT:** The capstone experiment measuring real-world learning gains.
4. **Longitudinal Simulation:** A year-long simulated run to measure catastrophic forgetting.

Chapter 8

Timeline and Milestones

Table 8.1: High-level timeline and milestones.

Year	Milestones
Year 1	Formalize theorems; implement Python prototype; initial C++ kernel; workshop paper (NeurIPS/ICLR); IRB submission.
Year 2	Full C++/CUDA QCE-Core release; ASCENDANT-K implementation and large-scale simulated experiments; major conference paper (ICML/NeurIPS).
Year 3	Memory Graph integration; conduct powered RCT; longitudinal simulations; journal submission (JMLR/AIED).
Year 4	Final benchmarking; thesis composition; public release of all code, models, and (IRB-permitted) anonymized datasets.

Chapter 9

Deliverables

By the end of the PhD, the project will produce:

- **QCE-Core Library:** An open-source C++/CUDA library.
- **ASCENDANT-K Framework:** An open-source MARL training pipeline.
- **Benchmarks & Datasets:** Reproducible benchmark harnesses and anonymized datasets.
- **Three Peer-reviewed publications.**
- **The PhD Thesis.**

Appendix A

References

Bibliography

- [1] Ashish Vaswani et al. “Attention is all you need”. In: *Advances in neural information processing systems*. 2017, pp. 5998–6008.
- [2] John J Hopfield. “Neural networks and physical systems with emergent collective computational abilities”. In: *Proceedings of the national academy of sciences* 79.8 (1982), pp. 2554–2558.
- [3] Richard P Feynman. “Simulating physics with computers”. In: *International journal of theoretical physics* 21.6-7 (1982), pp. 467–488.
- [4] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.