IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141018

REINFORCED-LLM TUTOR (RLT): A MULTI-AGENT FRAMEWORK FOR DYNAMICALLY PERSONALIZED LEARNING

A George[1], S Sharavana Ragav[2], M Abhishek[3], Dr. Golda Dilip[4]

Student, Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India[1]

Student, Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India[2]

Student, Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India[3]

Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India[4]

Abstract: Traditional Intelligent Tutoring Systems (ITS) fail to scale due to static, pre-programmed pedagogy. This paper introduces the Reinforced-LLM Tutor (RLT), a novel architecture overcoming these limits. RLT synergistically integrates Large Language Models (LLMs), Reinforcement Learning (RL), and multi-agent systems. The framework features four modules: a Retrieval-Augmented Generation (RAG) Domain Knowledge Module to ensure factual accuracy, a Dynamic Student Model tracking cognitive/affective states, a Multi-Agent Pedagogical Core (Expert, Socratic, Motivational agents), and an Adaptive Policy Engine. This engine uses RL, modeled as an MDP, to learn an optimal teaching policy, creating a truly adaptive, self-improving tutor.

Keywords: Intelligent Tutoring Systems, Large Language Models, Reinforcement Learning, Personalized Learning, Multi-Agent Systems, Educational Technology, Gamification.

I. INTRODUCTION

1. Background

The global educational landscape is undergoing a seismic shift towards digital learning, a transformation accelerated by the integration of technology into modern life.1 While this presents an unprecedented opportunity to democratize education, it also exposes the weakness of "one-size-fits-all" instructional models.2 Such models fail to accommodate diverse learning styles and paces, limiting their efficacy.4

This challenge conflicts with established educational psychology, notably Bloom's "2 Sigma Problem," which demonstrated that students receiving one-on-one human tutoring perform, on average, two standard deviations better than those in a traditional classroom.[5] For decades, the ambition of Artificial Intelligence in Education (AIED) has been to replicate this individualized tutoring affordably and at scale, effectively solving the 2 Sigma Problem through technology.[1]

2. Existing evidence (Literature survey)

The primary vehicle for this goal has been the Intelligent Tutoring System (ITS). Traditional ITS architectures, typically composed of a Domain Model, Student Model, and Tutor Model, have shown moderate learning gains.6 However, their potential is unrealized due to fundamental limitations. These include:

- **The Knowledge Engineering Bottleneck:** An immense human effort is required to manually encode expert knowledge and pedagogical rules, making systems difficult to scale or adapt.[5]
- **Brittle Interaction Models:** Early ITS relied on template-based dialogues, lacking the fluency to handle the nuance of human language, which can frustrate learners.[11]
- **Inflexible Pedagogy:** Teaching logic is typically a set of static, hand-crafted rules that cannot learn or evolve based on their effectiveness.[11]

DOI: 10.17148/IJARCCE.2025.141018

Recent advances in AI provide tools to overcome these specific challenges. Large Language Models (LLMs) offer fluent, context-aware conversational abilities, solving the interaction bottleneck [2], though they risk "hallucination" (generating factual inaccuracies).[13] Reinforcement Learning (RL) provides a mathematical framework for learning optimal decision-making policies through interaction, addressing the challenge of static pedagogy.[4]

3. Research gap

The literature reveals that these powerful technologies—LLMs, RL, and advanced ITS architectures—have largely been investigated in isolation. There are conversational LLMs with no pedagogical intelligence, and RL-based tutors that can optimize a sequence but cannot hold a deep dialogue. The research gap lies at the integration of these domains. A next-generation ITS must combine the conversational prowess of an LLM with the adaptive, policy-optimizing intelligence of an RL agent and the supportive environment of a multi-agent system.

4. Objective

This paper proposes the Reinforced-LLM Tutor (RLT), a novel, hybrid ITS architecture designed to fill this gap. The objective is to provide a comprehensive blueprint for a system that synergistically integrates LLMs, RL, and multi-agent systems. The system is designed to:

1. Achieve fluid, open-ended, and natural language dialogue.
2. Ensure all instructional content is factually accurate using Retrieval-Augmented Generation (RAG).
3. Move beyond static, pre-programmed rules by using RL to *learn* an optimal, adaptive teaching policy.
4. Foster student engagement and support metacognitive skills through a multi-agent, gamified environment.

5. Scope (Limitations)

This paper details a conceptual and architectural framework. Its practical implementation is subject to several limitations, including the significant computational cost of training and deploying large-scale LLMs and deep RL models. The system's effectiveness is highly dependent on access to large, high-quality datasets for pre-training the RL policy. Furthermore, the design of the reward function is a complex, iterative process. Finally, a policy trained in one domain may not generalize to others without significant re-training.

II. MATERIALS AND METHODS

1. List of experimental processes' materials used

The RLT is a conceptual framework. The materials required for its implementation consist of:

- **Domain Knowledge Base:** A curated repository of trusted, factual information (e.g., digital textbooks, lecture notes, scientific articles) relevant to the learning domain.
- **Vector Database:** A specialized database to store vector embeddings of the knowledge base chunks for efficient similarity searching.
- **LLM (Large Language Model):** A foundational model (e.g., GPT-4, LLaMA) used for generating conversational responses and powering the pedagogical agents.
- **Conversational User Interface:** A chat-based front-end allowing students to interact with the system in natural language.
- **Backend Infrastructure:** A server to manage the Dynamic Student Model, the RAG pipeline, and the RL policy engine.

2. Methodological Approach

The RLT architecture is composed of four integrated components that form a cyclical learning loop.

1. **Domain Knowledge Module (RAG):** To ensure factual accuracy and mitigate LLM "hallucination," the RLT employs a Retrieval-Augmented Generation (RAG) architecture.[17] When a response is needed, a retriever searches the vector database for relevant, factual text chunks. These chunks are provided as context to the LLM, which then generates a factually-grounded answer.
2. **Dynamic Student Model:** This module maintains a real-time, multi-dimensional vector representing the learner. It tracks:
    - **Cognitive State:** Mastery of each concept, using methods like Bayesian Knowledge Tracing (BKT).[8]
    - **Affective State:** Inferred engagement, confusion, or frustration, derived from interaction data (e.g., sentiment analysis, response times).[19]
    - **Learning Preferences:** Implicitly learned preferences for different types of instructional content (e.g., video vs. text).
3. **Multi-Agent Pedagogical Core:** To create a rich learning environment, the system uses three distinct, LLM-powered agents inspired by systems like MetaTutor [21]:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141018

- **The Expert Tutor:** Delivers direct, RAG-grounded explanations and feedback.
- **The Socratic Guide:** Asks probing questions to stimulate critical thinking and self-explanation.[22]
- **The Motivational Peer:** Provides encouragement, and affective support, and manages the gamification system (points, badges) to enhance engagement.[23]

4. **Adaptive Policy Engine (RL):** The core of the RLT's intelligence. This engine decides which agent acts and what action they take.

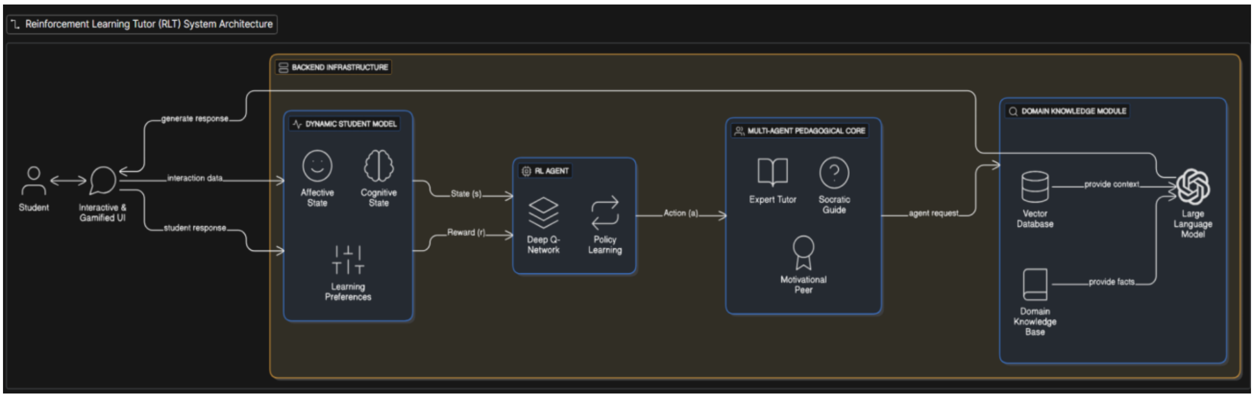
3. Tools and Instruments Used in Data Analysis

The pedagogical strategy is formalized as a Reinforcement Learning problem, specifically a Markov Decision Process (MDP).25 This provides the analytical toolset for the system to learn how to teach.

- **State (s):** The complete vector output from the Dynamic Student Model, capturing the learner's cognitive and affective state at time t.[18]
- **Action (a):** A discrete set of pedagogical moves available to the system, such as ExpertTutor.ExplainConcept(ID), SocraticGuide.AskProbingQuestion(ID), MotivationalPeer.GiveEncouragement(), or System.PresentPracticeProblem(difficulty).
- **Reward (r):** The system's goal is defined by a composite reward function, $r_t = w_1 \cdot \Delta M_t + w_2 \cdot E_t - w_3 \cdot C_t$. This function is engineered to balance the objectives of maximizing knowledge gain (ΔM_t), maintaining student engagement (E_t), and ensuring instructional efficiency (C_t).[18]
- **Policy Learning:** Given the high-dimensional state space, a Deep Q-Network (DQN) is used to approximate the optimal action-value function, $Q^*(s, a)$.[16] The DQN is a neural network that learns a policy, $\pi(s) = \arg\max_a Q(s, a; \theta)$, which maps any given student state to the best possible pedagogical action. The DQN can be pre-trained offline on existing student-tutor interaction logs and then continuously fine-tuned online.

III. ARCHITECTURE DIAGRAM

The RLT architecture operates in a continuous, cyclical flow, as depicted in the conceptual diagram below.



IV. RESULTS AND DISCUSSION

This paper proposes a novel framework; therefore, this section outlines a hypothetical experimental design for its validation and discusses the anticipated results.

A rigorous evaluation would employ a randomized controlled trial with a pre-test/post-test design. Participants (e.g., undergraduate students in an introductory course) would be randomly assigned to two groups:

1. **Experimental Group:** Students using the full Reinforced-LLM Tutor (RLT) system.
2. **Control Group:** Students using a simplified, non-adaptive version of the tutor. This control system would use the same domain content but would follow a fixed, linear curriculum and provide template-based feedback, lacking the RL, LLM-dialogue, and multi-agent features.

Key Performance Metrics for evaluation would include:

- **Learning Gain:** The primary cognitive outcome, measured by the normalized learning gain from pre-test to post-test.[3]

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025
DOI: 10.17148/IJARCCE.2025.141018

- **Engagement Metrics:** Behavioral data from system logs, such as time-on-task, problem completion rates, interaction frequency, and session duration.[26]
- **Student Satisfaction:** Self-reported data from post-study questionnaires assessing the system's usability, perceived helpfulness, and enjoyability.

Anticipated Results: We hypothesize that the RLT experimental group will demonstrate significantly higher normalized learning gains compared to the control group. This anticipated improvement is attributed to the system's core features: the personalized, adaptive pedagogical policy learned by the RL agent, the deeper conceptual engagement fostered by the Socratic and Expert agents, and the enhanced motivation provided by the Motivational Peer and gamification elements. We also anticipate the RLT group to show superior engagement and satisfaction metrics. A secondary, qualitative result would be the analysis of the final learned RL policy, which could reveal novel, non-obvious, and effective teaching strategies that are discoverable by the AI.

V. CONCLUSION

Summary of Findings

This paper has introduced the Reinforced-LLM Tutor (RLT), a novel ITS architecture designed to overcome the static, non-adaptive limitations of traditional systems. We propose a comprehensive blueprint for synergistically integrating three powerful AI paradigms: Large Language Models for fluent, natural language dialogue; Reinforcement Learning for adaptive, self-improving pedagogical policy; and Multi-Agent Systems for rich, scaffolded, and motivational learning. The RLT framework's use of a RAG module to ensure factual accuracy and its formalization of teaching as a Markov Decision Process represent a significant step beyond pre-programmed instruction. This architecture provides a principled pathway toward developing AI tutors that can learn, adapt, and personalize instruction at scale, moving the field closer to solving Bloom's 2 Sigma Problem.

Limitations

The deployment of a data-driven system like the RLT carries significant ethical responsibilities. Key limitations and challenges include:

- **Data Privacy:** The Dynamic Student Model collects vast amounts of sensitive data, including inferred affective states. This necessitates robust anonymization, secure data governance, and explicit user consent to protect student privacy.[18]
- **Algorithmic Bias:** The RL policy is trained on data. If this data reflects existing educational inequities, the agent may learn a policy that amplifies these biases, potentially disadvantagering underrepresented student groups. Continuous fairness audits are essential.[18]
- **Transparency:** The deep learning models (LLMs and DQN) are inherently "black boxes." This lack of interpretability can be a major barrier to trust and adoption by educators. Integrating Explainable AI (XAI) methods to provide rationales for the RLT's pedagogical decisions is a critical, non-trivial challenge.[13]

Future Directions

The RLT framework serves as a foundation for numerous avenues of future research. Key directions include:

1. **Multimodal Interaction:** Extending the system to process and generate content beyond text, such as diagrams, audio explanations, and video content, to cater to diverse learning styles.[13]
2. **Advanced Affective Computing:** Integrating more sophisticated sensors (e.g., webcam-based facial expression or eye-tracking analysis) to provide the student model with a richer, more accurate signal of confusion, boredom, or engagement.[21]
3. **Collaborative Learning:** Developing a multi-student version of the RLT where the AI agents act as smart facilitators to guide small groups of students through collaborative problem-solving tasks.
4. **Human-in-the-Loop RL:** Implementing a system where a human educator can review, approve, or correct the RL agent's actions, particularly in high-stakes situations, to improve safety and build trust.

REFERENCES

[1]. EBSCO, “Intelligent Tutoring Systems,” n.d.
[2]. ResearchGate, “LLMs in Personalized Education: Adaptive Learning Models,” n.d.
[3]. MDPI, “Adaptive Learning Using Artificial Intelligence in e-Learning,” n.d.
[4]. Nested.ai, “Personalized Education Through Reinforcement Learning,” 2024.
[5]. Institute for Defense Analyses, “Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review,” 2017.
[6]. “Intelligent Tutoring Systems: A Literature Review,” *International Journal of Computer Applications*, 2019.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 10, October 2025
DOI: 10.17148/IJARCCE.2025.141018

[7]. PMC, "Effectiveness of Intelligent Tutoring Systems: A Systematic Literature Review," 2023.
[8]. "AI-based Intelligent Tutoring Systems (ITS)," arXiv preprint, 2025.
[9]. PMC, "Effectiveness of Intelligent Tutoring Systems in K-12 Education," n.d.
[10]. MDPI, "Intelligent Tutoring Systems in Mathematics Education," n.d.
[11]. "The Impact of Large Language Models in Education," arXiv preprint, 2025.
[12]. "Adaptive Learning Systems with LLM-powered Analytics," arXiv preprint, 2025.
[13]. Al-Hussein Bin Talal University, "Applications of Large Language Models in Education," 2025.
[14]. ResearchGate, "Reinforcement learning for personalization: A systematic literature review," n.d.
[15]. MDPI, "Reinforcement Learning in Education," n.d.
[16]. Semantic Scholar, "Towards a Personalized Learning Experience Using Reinforcement Learning," n.d.
[17]. "Reinforced-EDU: An Intelligent Reinforcement Learning-Guided Framework," *IEEE Xplore*, 2023.
[18]. ResearchGate, "Multiple Agent Designs in Conversational Intelligent Tutoring Systems," n.d.
[19]. "SRL with MetaTutor: A Hypermedia-Based ITS," *Frontiers in Psychology*, 2022.
[20]. ResearchGate, "Automated Disengagement Tracking Within an Intelligent Tutoring System," n.d.
[21]. Semantic Scholar, "Intelligent Tutoring Systems: A Comprehensive Survey," n.d.
[22]. "Intelligent Tutoring Systems: A Comprehensive Historical Survey with Recent Developments," arXiv preprint, 2018
[23]. "A Comprehensive Review of AI-based Intelligent Tutoring," arXiv preprint, 2025.
[24]. Stanford University, "Intelligent Tutoring Systems and Online Learning," 2016.
[25]. "Large Language Models for Education: A Survey," arXiv preprint, 2024.
[26]. ResearchGate, "A Review of The Application of Large Language Models in Education," n.d.
[27]. Medium, "LLM in Education: Transforming Learning with AI-Powered Intelligence," n.d.
[28]. Sapien.io, "LLMs for Personalized and Accessible Education," n.d.
[29]. MDPI, "AI Assistants for Personalized Adaptive Learning," n.d.
[30]. ResearchGate, "Deep Reinforcement Learning for Personalized Academic Performance Improvement," 2025. [31]. PMC, "Self-Regulated Learning with MetaTutor," 2022.
[32]. YouTube, "Level Up Learning: Gamification Strategies That Drive Online Engagement," n.d.
[33]. "The Impact of Gamification and Animation in E-learning," *Journal of Business and Social Sciences Nexus*, n.d
[34]. "The Impact of Gamification Elements on Student Motivation," *Journal of Applied Data Sciences*, 2025.
[35]. CORE, "The Impact of Gamification on Student Motivation in Online Learning," n.d.
[36]. MDPI, "Digital Tools and Student Interaction in Virtual Classrooms," n.d.
[37]. Learning Mole, "The Psychological Impacts of AI Learning," n.d.
[38]. "Deep Reinforcement Learning Models for Emotion Detection and Personalization."