

EDMIT: An End-to-End Agentic Framework for Enhanced Decision-Making in Interactive Motion Tutoring

Jinhua Du
dujh22@mails.tsinghua.edu.cn
Tsinghua University
Beijing, China

Mufeng Xing
Ziheng Zhou
Ruilin Zhang
Zhongshi Liu
Banlan Technology
Suzhou, China

Zexun Jiang
School of Data Science and
Intelligent Media Communication
University of China,
Beijing, China



Figure 1: Seattle Mariners at Spring Training, 2010.

Abstract

One-on-one tutoring has proven effective, yet current LLM-based educational systems still struggle to achieve individualized decision-making in multi-turn, human-AI co-performance settings. We focus on three open challenges: (i) how to construct usable data from scarce and unstructured coaching experience at low cost, (ii) how to design adaptive feedback mechanisms that accommodate individual differences during interaction, and (iii) whether AI-based coaching can match—or even surpass—human tutors. To address these, we propose EDMIT, an end-to-end agentic framework for interactive motion tutoring. EDMIT builds a seed-to-augmentation data pipeline from coaching trajectories, introduces a standardized decision ontology and executable decision chain, and leverages closed-loop feedback to dynamically adjust prompt granularity, exercise difficulty, and error-correction strategies during sessions. We implemented a prototype and conducted a double-blind controlled study with 10 users and 10 human coaches, comparing EDMIT against a prompt-based baseline under source-agnostic conditions. Results show that EDMIT achieves higher decision consistency with human coaches, significantly improves user satisfaction and acceptance, and reaches state-of-the-art performance relative

to existing methods. These findings demonstrate that agent-driven human-AI co-performance can effectively bridge the gap of individualized decision-making in motion tutoring, offering empirical evidence and methodological insights for the design of future AI coaching systems.

CCS Concepts

• **Human-centered computing** → **User models.**

Keywords

Interactive Motion Tutoring; Agentic Framework, Personalized Decision-Making, Large Language Models (LLMs), Data Augmentation, Feedback Loop, Decision Chain, Double-Blind Study, Empirical Evaluation

ACM Reference Format:

Jinhua Du, Mufeng Xing, Ziheng Zhou, Ruilin Zhang, Zhongshi Liu, and Zexun Jiang. 2018. EDMIT: An End-to-End Agentic Framework for Enhanced Decision-Making in Interactive Motion Tutoring. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/XXXXXX.XXXXXXX>

1 Introduction

One-on-one human tutoring has long been recognized as a highly effective form of instruction [2]. With the rapid advancement of artificial intelligence—particularly the maturation of large language model (LLM)-based technologies—numerous intelligent systems have emerged to support end users in domains such as education [10], healthcare [6], and sports science [3]. More recently, the advent of deep reasoning models, exemplified by o1 [9] and DeepSeek [4],

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, Woodstock, NY

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-XXXX-X/2018/06

<https://doi.org/XXXXXX.XXXXXXX>

has significantly strengthened the reasoning capabilities of intelligent systems. These models have already demonstrated performance at the level of human doctoral students in domains such as mathematics [1] and programming [7].

Despite these advances, significant challenges remain when applying LLM-based systems to real-world tutoring scenarios—most notably the lack of personalized decision-making. This limitation arises because many existing intelligent tutoring systems are primarily built upon AI models coupled with fixed prompt-based query-response mechanisms [5, 8]. Relying on a single prompt and the model itself restricts the system’s ability to adapt to learners’ individualized needs during human–AI interaction. Moreover, the one-shot nature of current generation methods hampers their effectiveness in multi-turn interactions and continuous decision-making. As a result, conventional approaches struggle to address the complexity of authentic tutoring contexts.

Taking interactive motion tutoring as an example, human coaches are able to make personalized decisions tailored to individual learners’ conditions during training. In contrast, virtual coaching systems often fail to replicate such capabilities. This is largely due to the fact that human coaching expertise rarely exists in structured, textual form suitable for model training; the available data are both limited in scale and insufficient to cover the diversity of real-world scenarios. Consequently, there remains no widely adopted interactive motion tutoring system that delivers a satisfactory user experience.

To address this gap and tackle the challenge of personalized decision-making in motion tutoring, we propose EDMIT, an end-to-end agentic framework designed to enhance the decision-making capacity of AI systems and closely approximate the decision behaviors of human coaches in interactive exercise guidance. Specifically, through our user-centered investigation, we seek to answer three key questions:

- **RQ1:** How can we design a fully automated AI system that constructs the necessary data inputs at low cost from highly limited and unstructured human coaching experience?
- **RQ2:** How can we design an effective decision-augmentation mechanism that adaptively addresses diverse and personalized tutoring needs in interactive motion scenarios?
- **RQ3:** To what extent can AI-based motion tutoring systems effectively simulate the guidance of human coaches—and is there potential for them to surpass human performance? In other words, what is the upper bound of intelligence for such systems?

Based on modeling and data collection of real coaching trajectories, we first constructed seed training and evaluation datasets. To address RQ1, we developed a data augmentation framework that expands the seed data to arbitrary scales, thereby meeting the data requirements for building intelligent AI systems. To address RQ2, inspired by the feedback loop principle in human motor learning, we standardized the modeling of all possible decision nodes in interactive motion tutoring. We further systematized these into a complete decision chain, whose execution is driven by the agent, enabling adaptive tutoring services that accommodate diverse and personalized user needs.

To evaluate whether the EDMIT framework can genuinely enhance users’ motor performance, we introduce—for the first time in AI-assisted motion tutoring—a double-blind experimental design. We developed a prototype system implementing EDMIT to serve real users, who were unaware of whether the guidance they received came from a human coach or the AI system. After each tutoring session, participants rated the different conditions. We conducted a controlled user study involving 10 real users and 10 professional coaches, comparing EDMIT against both human coaching and traditional baseline methods (RQ3). Results demonstrate that EDMIT delivers decision-making performance on par with human coaches, achieves state-of-the-art results relative to existing approaches, and significantly improves users’ acceptance and satisfaction with AI-assisted motion tutoring.

In summary, this work explores an end-to-end agentic framework for enhancing decision-making in interactive motion tutoring, bridging the gap between human coaching expertise and AI intelligence to improve users’ learning outcomes. Our contributions are fourfold: 1. A methodological framework that introduces a systematic approach to enhancing decision-making in interactive motion tutoring. 2. An open-source system that provides both code implementations and a working prototype to facilitate future research and development. 3. A controlled user study that evaluates the effectiveness of our AI system design in real-world tutoring contexts. 4. Design insights for future AI-driven interactive motion tutoring systems, highlighting pathways toward effective personalization and human–AI interaction.

References

- [1] Xiaoyang Chen, Xinan Dai, Yu Du, Qian Feng, and Naixu Guo. 2025. *DeepMath-Creative: A Benchmark for Evaluating Mathematical Creativity of Large Language Models*. <http://arxiv.org/abs/2505.08744>
- [2] Michelene T.H. Chi, Stephanie A. Siler, Heisawn Jeong, Takashi Yamauchi, and Robert G. Hausmann. 2001. Learning from Human Tutoring. *Cognitive Science* 25, 4 (2001), 471–533. https://onlinelibrary.wiley.com/doi/abs/10.1207/s15516709cog2504_1
- [3] Mark Connor and Michael O’Neill. 2023. *Large Language Models in Sport Science & Medicine: Opportunities, Risks and Considerations*. <http://arxiv.org/abs/2305.03851>
- [4] DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, and Junxiao Song. 2025. *DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning*. <http://arxiv.org/abs/2501.12948>
- [5] Ioana R. Goldbach and Felix G. Hamza-Lup. 2020. *Intelligent Tutoring Systems for Generation Z’s Addiction*. <http://arxiv.org/abs/2005.05024>
- [6] Kai He, Rui Mao, Qika Lin, Yucheng Ruan, Xiang Lan, Mengling Feng, and Erik Cambria. 2025. A Survey of Large Language Models for Healthcare: From Data, Technology, and Applications to Accountability and Ethics. *Information Fusion* 118 (2025), 102963. <https://www.sciencedirect.com/science/article/pii/S1566253525000363>
- [7] Naman Jain, King Han, Alex Gu, Wen-Ding Li, and Fanjia Yan. 2024. *LiveCodeBench: Holistic and Contamination Free Evaluation of Large Language Models for Code*. <http://arxiv.org/abs/2403.07974>
- [8] Soonwoo Kwon, Sojung Kim, Minju Park, Seunghyun Lee, and Kyuseok Kim. 2024. *BIPED: Pedagogically Informed Tutoring System for ESL Education*. <http://arxiv.org/abs/2406.03486>
- [9] OpenAI, Aaron Jaech, Adam Kalai, Adam Lerer, and Adam Richardson. 2024. *OpenAI O1 System Card*. <http://arxiv.org/abs/2412.16720>
- [10] Shen Wang, Tianlong Xu, Hang Li, Chaoli Zhang, and Joleen Liang. 2024. *Large Language Models for Education: A Survey and Outlook*. <http://arxiv.org/abs/2403.18105>

A Research Methods