TPR: A Training Procedure Representation to Augment XR Simulations with LLMs

Michael Guevarra[1], **Christabel Wayllace**[2],
Srijita Das[3], **Carrie Demmans Epp**[4], **Alan Tay**[1]
[1] Illumia Labs,
[2] New Mexico State University,
[3] University of Michigan-Dearborn,
[4] University of Alberta,
{michael.gue, alan.tay}@illumialabs.ai, cwayllac@nmsu.edu, sridas@umich.edu, demmanse@ualberta.ca

Abstract

Extended reality (XR) is well suited to support the situated learning of technical procedures. At the same time, AI-driven intelligent tutoring systems (ITS) can complement XR by providing adaptive pedagogical support. Many domains would benefit from this combination, especially when trainers, equipment, or team members are limited. We present a domain-agnostic XR-based ITS that integrates a training procedure representation (TPR), XR simulation, and an LLM-driven instructor. We demonstrate the tutor's use for tissue sample handling and engine repair, showing how it delivers adaptive feedback, collaborative roleplay, and dynamic scenario management to create realistic and pedagogically meaningful training experiences.

Introduction

Training for procedural tasks, such as handling human tissues in a lab or equipment repair, requires mastering not only the correct actions, but also their correct sequence. In these scenarios, mistakes can lead to costly or unsafe results. While domain experts can provide instruction, their availability is limited, key learning opportunites may not present themselves, and additional teamates are not always available for collaborative tasks. Extended reality (XR) environments can address the scarcity of instruments, while intelligent tutoring systems (ITS) can provide structured feedback and scaffolding.

Prior work on using XR in situated learning has been relatively under-explored. Applications have been trialed in higher education (Hsu et al. 2023), healthcare (Barteit et al. 2021), and sign language education (Wen et al. 2024). A lack of realistic simulators and pedagogical support for varied educational settings is a key inhibitor in this space. AI-based tutoring systems could help alleviate these pressures (Hajahmadi et al. 2024; Cheng et al. 2024; Liu et al. 2024) through dialogue (Leandro et al. 2024; Guevarra et al. 2025) or by generating personalized feedback using large language models (Jensen, Sankaranarayanan, and Hayes 2024). However, most tutoring systems are restricted to a single task, thus making the setup too specific without possibility for adaptation (learning and supervision) to other types of tasks.

We present an XR-based ITS that unifies XR simulation with a training procedure representation (TPR) and a large language model (LLM)-driven instructor, enabling direct interactions between student, LLM, and virtual environment. The architecture is domain-agnostic: the same framework can be used to train tasks as diverse as surgical procedures, engine repair, or laboratory safety. This flexibility makes it possible to extend training across domains that differ in complexity, risks, and dynamics as well as those that require coordination with others.

This paper highlights how the XR-ITS supports procedural learning through three interactive instructional features: adaptive feedback, dynamic scenario management, and collaborative roleplay. Together, these capabilitites demonstrate how the tutor can extend beyond static simulations to deliver realistic and relevant training experiences.

System Architecture

Figure 1 shows the architecture of our XR-based ITS. The design is inspired by concepts of classical planning, where a task is represented as a sequence of actions governed by preconditions and effects.

A field expert specifies the training scenario, including (1) the task goal (e.g., repairing an engine), (2) the relevant objects and tools, (3) the set of possible actions, and (4) valid action sequences (procedure steps). This information is encoded in two complementary representations. First, the XR training simulation is configured by instantiating objects, tools, and visual interactions, setting the procedure steps, and logging the action history. Second, the TPR formalizes

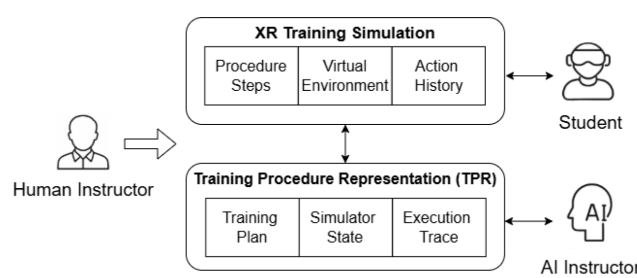

Figure 1: Block Diagram of TPR System.

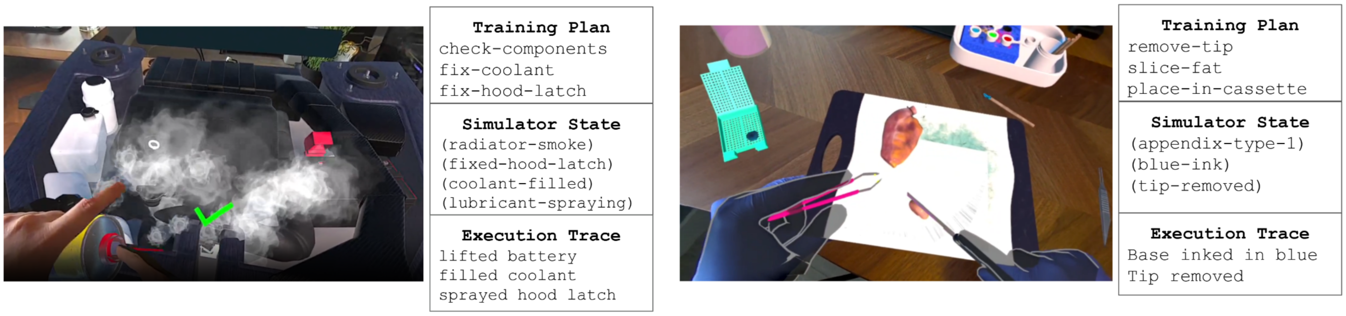

Figure 2: Sample XR-based interactive tutoring sessions. Engine repair (left) and tissue handling (right) are shown with their corresponding training plan, simulator state, and execution trace.

the procedure as a structured plan, specifying the action sequence along with their associated preconditions and effects. TPR also encodes the current environment status as a simulator state defined by state variables and the action history as an execution trace (see examples in Figure 2).

An LLM-driven instructor monitors the student's actions and synchronizes them with the TPR and the XR simulation. Each executed action is evaluated against the training plan: if preconditions are not satisfied, the instructor can either block the action or allow it to proceed, depending on the initial configuration. Therefore, the system can accomodate strict execution (encoding the modeled plan) and explorative execution (allowing deviations to illustrate errors).

Stochastic Outcomes and State Tracking

Unlike classical planning domains with deterministic outcomes, physical training tasks involve stochasticity. For example, a student may drop a tool, producing a state not explicitely modeled in the TPR. The AI instructor accounts for such events by recognizing non-relevant states and realigning the training sequence, preserving the overall coherence of the session. This allows the system to capture realistic contingencies while maintiaining plan-based guidance.

Instructional Features

The system combines the TPR, the XR simulation model, and the learner's action history to implement three instructional features using GPT-4.1:

Scenario Management: The AI instructor can manipulate the simulation model to change training conditions, randomize parameters, or trigger events (e.g., oil leak, tool malfunction). These changes can be seen as analogous to state variables and their values in a planning model. The AI instructor can also control the timing of interventions and the level of scaffolding by enabling or disabling hint and feedback systems. This ensures that training sessions remain adaptive and contextually meaningful.

Roleplay: Using the XR context captured in the simulation model, the AI instructor can roleplay as other team members or supporting staff. For example, in a surgical scenario the AI can play the role of a nurse, while in an engine repair exercise it can act as a more experienced mechanic. These interactions are drawn from the same operations defined in the procedure steps and controlled by the execution trace. In this way, the instructor can maintain procedural flow in multi-actor scenarios. This ensures that team-based training remains pedagogically meaningful.

Feedback. The AI instructor also compares the training plan against the learner's execution trace to identify mistakes and provide feedback. Feedback is primarily formative and constructive, delivered either in real time (by highlighting mistakes, suggesting corrections, or blocking invalid actions) or at the end of the exercise as reports (Hattie and Timperley 2007). Our current prototype relies on LLM-based summaries, which can indicate *what* went wrong; however, representing procedures symbolically through planning structures would enable richer causal explanations of *why* a sequence was invalid (Chakraborti et al. 2017; Vasileiou et al. 2025). To improve the accuracy of feedback, sample errors and corrective actions can be provided by the human instructor based on the training plan.

Figure 2 illustrates two example XR tutoring sessions. Each scenario (tissue handling and engine repaifir) is paired with its training plan, simulator state, and execution trace. This provides the LLM with simulator context to generate real-time tutoring.

Conclusion and Future Work

We introduced an XR-based ITS that integrates a TPR, XR simulation, and LLM-driven instructor to simulate different interactive tutoring sessions. By defining the plan, simulator states, and execution trace, the system supports scenario management, collaborative roleplay, and adaptive feedback, demosntrating its potential to deliver realistic and flexible procedural training.

Future Work: Education using XR remains underexplored, particularly in domains where muscle memory and procedural accuracy are critcal. Future work includes using symbolic planning representations to automate procedure generation, adapt tasks across domains, and provide richer causal feedback. Roleplay can also be extended to train metacognitive skills by promting learners to plan, explain, and reflect on their actions. Finally, the system provides a platform to study human-AI interaction, especially how human strategies and AI scaffolds complement each other in procedural training.

Acknowledgements

This work has been completed for Illumia Labs Inc. as part of their advanced training platform in collaboration with SingHealth Duke-NUS Institute of Medical Simulation (SIMS). Part of this work has taken place in the EdTeKLA Lab at the University of Alberta, which are supported in part by grants from Amii; a Canada CIFAR AI Chair, Amii; DRAC; Mitacs; and NSERC.

References

Barteit, S.; Lanfermann, L.; Bärnighausen, T.; Neuhann, F.; and Beiersmann, C. 2021. Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. *JMIR serious games*, 9(3): e29080.

Chakraborti, T.; Sreedharan, S.; Zhang, Y.; and Kambhampati, S. 2017. Plan explanations as model reconciliation: moving beyond explanation as soliloquy. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 156–163.

Cheng, A. Y.; Guo, M.; Ran, M.; Ranasaria, A.; Sharma, A.; Xie, A.; Le, K. N.; Vinaithirthan, B.; Luan, S.; Wright, D. T. H.; et al. 2024. Scientific and fantastical: Creating immersive, culturally relevant learning experiences with augmented reality and large language models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–23.

Guevarra, M.; Bhattacharjee, I.; Das, S.; Wayllace, C.; Epp, C. D.; Taylor, M. E.; and Tay, A. 2025. An LLM-Guided Tutoring System for Social Skills Training. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 29643–29645.

Hajahmadi, S.; Clementi, L.; López, M. D. J.; and Marfia, G. 2024. Arele-bot: Inclusive learning of spanish as a foreign language through a mobile app integrating augmented reality and chatgpt. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 335–340. IEEE.

Hattie, J.; and Timperley, H. 2007. The power of feedback. *Health Psychology Review*, 77(1): 81–112.

Hsu, C.; Chen, Y.; Liu, Y.-J.; Chang, Y.-C.; and Lee, M.-J. 2023. Spelland: Situated Language Learning with a Mixed-Reality Spelling Game through Everyday Objects. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–6.

Jensen, E.; Sankaranarayanan, S.; and Hayes, B. 2024. Automated Assessment and Adaptive Multimodal Formative Feedback Improves Psychomotor Skills Training Outcomes in Quadrotor Teleoperation. arXiv:2405.15982.

Leandro, J.; Rao, S.; Xu, M.; Xu, W.; Jojic, N.; Brockett, C.; and Dolan, B. 2024. GENEVA: GENErating and Visualizing branching narratives using LLMs. In *2024 IEEE Conference on Games (CoG)*, 1–5.

Liu, Z.; Zhu, Z.; Zhu, L.; Jiang, E.; Hu, X.; Peppler, K. A.; and Ramani, K. 2024. Classmeta: Designing interactive virtual classmate to promote vr classroom participation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–17.

Vasileiou, S. L.; Yeoh, W.; Previti, A.; and Son, T. C. 2025. On generating monolithic and model reconciling explanations in probabilistic scenarios. *Journal of Artificial Intelligence Research*, 84.

Wen, H.; Xu, Y.; Li, L.; Ru, X.; Wu, Z.; Fu, Y.; Zheng, X.; and Wang, X. 2024. Enhancing Sign Language Teaching: A Mixed Reality Approach for Immersive Learning and Multi-Dimensional Feedback. In *2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 740–745. IEEE.