Towards Effective Explanation Generation in Autonomous Coaching Agents with Generative AI

Grace Brazil
Design Intelligence Lab
Georgia Institute of Technology
gbrazil2@gatech.edu

1 INTRODUCTION

This research aims to gain comprehensive understanding of learning theory and explanation generation in autonomous coaching agents in instructional settings. Firstly, it seeks to comprehensively understand the theory of learning and its impact on skill acquisition within educational environments. Secondly, it aims to explore explanation generation methods in AI and how it can incorporate Generative AI methods, particularly Large Language Models, to enhance the effectiveness of autonomous agent systems for learning purposes.

2 RESEARCH OBJECTIVES

In pursuit of addressing the key facets of learning theory and autonomous agents that can explain and reason, the following research objectives are:

- To develop a comprehensive understanding of how active learning and cognitive engagement impacts skill acquisition in educational settings.
- To investigate the concept of "understanding" a skill, guided by the Task, Method, and Knowledge (TMK) theory, developing a framework for integrating bottom-level knowledge into methods for accomplishing goals.
- To explore explanation generation within autonomous agent systems, drawing insights from current research in AI to improve its effectiveness of explanations provided to learners.

3 LITERATURE REVIEW

A comprehensive examination of past studies concerning the following:

- ICAP Theory of Learning: Delve into the ICAP (Interactive, Constructive, Active, Passive) theory of learning, examining its implications for cognitive engagement and skill acquisition in educational contexts.
- TMKL2: Review TMKL2, the language framework employed in the develop-

ment of Intelligent Agents, focusing on its structure and effectiveness in representing problem-solving skills and facilitating active learning.
- Explanation Generation in AI: Review current research on explanation generation in Artificial Intelligence, considering both implementation strategies and evaluation methodologies.

4 METHODOLOGY

In order to evaluate the research goals, the proposed method include:

- Building a Task-Model-Knowledge (TMK) AI model to represent skills that learners need to acquire, which includes testing and data gathering to evaluate the model's effectiveness.
- Experimental evaluation of explanation quality using different generative AI methods, which examines factors such as coherence and relevance of explanations.
- Case studies of how LLMs are used for explanation and reasoning in AI agents.

5 PROJECT TIMELINE

- Week 1-3: Literature Review and Research Design
- Week 4-6: Data Collection
- Week 7-8: Data Analysis
- Week 9-10: Report Writing and Presentation Preparation

6 EXPECTED OUTCOMES

This research aims to achieve two outcomes. First, a comprehensive understanding of the roles played by active learning and cognitive engagement in shaping skill acquisition within educational environments. Second, a deeper insight into explanation generation in autonomous agent systems. Lastly, a practical experience in the implementation of the TMK framework in autonomous agents and instructional systems.

7 EVALUATION

Additional to the expected outcomes mentioned above, a final paper for my summer work will be written and submitted for evaluation.