From Flashcards to ICA Learning Machine

Yufu Li
yufuli2@illinois.edu

Abstract

Since the advent of ChatGPT in 2023, the educational landscape has undergone a significant metamorphosis. While Large Language Models (LLMs) offer unparalleled capabilities in information extraction and human-esque communication, concerns arise regarding their impact on students' cognitive engagement. This study introduces the Interactive-Constructive-Active Learning Machine (ICALM), an avant-garde Intelligent Tutoring System that synergizes with the ICAP Framework (Chi and Wylie, 2014). ICALM integrates a triad of functionalities: concept map generation from flashcards, a knowledge-graph based recommendation system promoting exploration, and LLM-facilitated active learning that encompasses chatbot interactions and reflective essays. Central to this system is the progression from mere content absorption to heightened engagement levels—transitioning from content reading to paraphrasing, concept mapping, and peer discussions. Leveraging AI, ICALM ensures a structured comprehension of course materials and promotes a shift in concept mapping from a constructive endeavor to an active learning experience. The system culminates in an interactive learning task, drawing inspiration from the In-class portfolio (Walvoord et al., 2013), and incorporates calibrated feedback mechanisms powered by LLMs.

1 Introduction

Since the inception of ChatGPT in early 2023, there has been a notable transformation in the educational sector. Large Language Models (LLMs) have exhibited remarkable capabilities in information retrieval and the composition of human-like responses. This development has led to concerns, as students can now effortlessly obtain answers from educators' meticulously crafted questions, bypassing their own cognitive engagement with the material. Nevertheless, these challenges also pave the way for potential advancements. Techniques to engage students in online learning can be more effectively integrated into their pre-class, in-class, and post-class activities, particularly for tasks demanding constructive and self-motivated study.

The Need for Structured Learning: During late high school and early tertiary education, students confront a stark increase in the volume and complexity of information they must internalize. As the curriculum progresses, learners encounter vast arrays of interconnected topics, requiring them to shift from rote memorization to a more holistic, structured understanding. This shift demands tools that not only help in memorization but also in establishing connections between diverse concepts.

Cognitive Development Considerations: Cognitively, students between the ages of 15 and 22 are transitioning from the concrete operational stage to formal operational thought, as outlined by Jean Piaget's theory of cognitive development. In this phase, adolescents and young adults begin to think more abstractly, reason about hypothetical problems, and think deductively about abstract concepts. Our tool, ICALM, is tailored for this cognitive stage, aiding learners in connecting the dots, reasoning about abstract concepts, and drawing upon prior knowledge to understand new information.

Targeting Key Transition Points in Education:

Grade 10 to 12 Students: As students advance through high school, they face both standardized testing and the challenge of understanding foundational concepts that will be pivotal in further studies. ICALM facilitates this by providing a structured approach to navigate and internalize these foundational concepts.

AP Exam Students: AP courses are designed to be college-level, demanding a depth of understanding often new to high school students. ICALM aids in breaking down these rigorous topics, making them more digestible and ensuring students are well-prepared for their exams.

Entry-level Post-secondary Students: The transition from high school to college or university often involves a leap in the depth and breadth of material. Many students struggle with the volume of reading, the complexity of topics, and the need for critical thinking and analysis. ICALM provides scaffolding, helping students structure their learning and ensuring they internalize and reflect upon their coursework effectively.

In the current study, we present the design of the Interactive-Constructive-Active Learning Machine (ICALM), an advanced Intelligent Tutoring System. The ICALM comprises three main components: a) Entity-relation (ER) extractions, which generate concept maps from flashcards; b) a knowledge-graph based recommendation system facilitating student exposure to the most pertinent, intriguing, or peer-favored topics, thus fostering discussion and exploration; c) an active learning component where students utilize LLM-based chatbots to verify information and defend their perspectives; d) concluding with a reflective essay session where students contemplate their learnings, which is then juxtaposed with a summary generated by LLMs.

The structured learning stages align with the ICAP Framework, an instructional approach introduced by Chi and Wylie 2014. The framework categorizes overt learning behaviors into Passive, Active, Constructive, and Interactive, arranged from least to most effective. Rather than employing cognition terms from Bloom's taxonomy, this approach assesses student outcomes based on their adherence to task instructions specific to each category. To elaborate, student engagement levels should ideally escalate from mere content reading to paraphrasing, concept mapping, and peer discussions, respectively.

Similar to the original ICAP Framework, The system we delineate adheres to a conceptualized four-stage knowledge-processing sequence: storage, integration, inference, and collaborative inference (Chi and Wylie, 2014). The emphasis is on reorganizing the structural integrity of course materials, enabling logical comprehension rather than mere replication of original or paraphrased content. The AI-assisted concept mapping in component b) optimizes the initial content consumption phase, leading to an active engagement with logically structured and hierarchically presented information. Notably, within this system, concept mapping transitions from a constructive task to an active learning one. This shift simplifies the initiation of the learning process, enhancing student engagement. Processes in component c) utilize specifically tailored ChatGPT templates, enabling students to build upon their existing knowledge base. For instance, within an Economics module, students might explore ”historical events linked to rising interest rates”, subsequently attempting to elucidate ”the repercussions of increased interest rates” utilizing foundational course materials. The former represents system-facilitated constructive learning, while the latter exemplifies interactive learning. Component d) draws from the SET 45 - In-class portfolio, as outlined in the *”Student Engagement Techniques - A Handbook for College Faculty”* (Barkley, 2009), with LLMs providing calibrated feedback and scaffolding, culminating the session with an interactive learning task.

2 Related Work

2.1 The ICAP Framework

Originating from a study by Chi and Wylie, the ICAP framework delineates a taxonomy that encompasses four distinct modes of cognitive engagement, namely Interactive, Constructive, Active, and Passive. While Bloom's taxonomy (Athanassiou et al., 2003) primarily concentrates on instructional objectives, ICAP pivots on overt behaviors of students to gauge achievement. By empirically examining three engagement paradigms - note-taking, concept mapping, and self-examination - the authors deduced a hierarchy of efficacy from Interactive to Passive modes. This framework delineates the nuances among active, constructive, and interactive learnings, providing specific definitions for each.

2.2 Human Cognition

The capacity of working memory, instrumental in the transient storage and manipulation of information, is believed to be constrained to 3-5 chunks of information (Cowan, 2010). Nevertheless, strategies such as chunking enable individuals to optimize their memory utilization. Although the underpinnings of this constraint remain contentious, some ascribe it to biological limitations. As learning involves the interplay between short-term and long-term memory, innovations like Limbiks, Revisely, and Gizmo AI aim to convert information into digestible formats to enhance learning efficiency.

2.3 Knowledge Graph

Knowledge graphs epitomize heterogeneous networks, representing entities and their relationships through nodes and edges. This computationally-intensive architecture has prompted exploration into attention networks integrated within Graph Neural Networks (GNN). While models like GAT (Veličković et al., 2018) have shown promise, the nuanced efficacy of attention mechanisms in GNN demands intricate design considerations, as evidenced by models like GC-SAN (Xu et al., 2019) and SR-GNN (Wu et al., 2019).

2.4 Named Entity Relation Extractor

Constructing a knowledge graph necessitates the upstream extraction of semantic relationships from raw datasets, often formatted as triplets. Recent developments, such as REBEL by Huguet Cabot and Navigli, have leveraged BERT-based Seq2Seq models to treat entity-relation extraction and classification tasks as analogous to translation tasks. Further innovations, such as KnowGL by Rossiello et al., have expanded upon this approach, enhancing the granularity and richness of entity labels and relationships.

2.5 Recommender System

Session-based recommendation systems, historically grounded in methods like collaborative filtering, predict user behaviors based on historical interactions. However, as the scale of users and items has burgeoned, matrix representations have become increasingly unwieldy. Despite the initial promise of Recurrent Neural Networks (RNN) in this domain, their efficacy is hampered by data limitations and lack of context. Hence, emerging methodologies, including those encapsulated in SR-GNN (Wu et al., 2019), strive to capitalize on the rich contextual data inherent in user behaviors. Parallelly, there is a burgeoning demand for greater transparency and explainability in recommender systems, underscored by the complexities of specific application domains.

2.6 Knowledge Graph Based Recommender System

To redress the aforementioned shortcomings, recommender systems underpinned by knowledge graphs have gained traction. Their inherent ability to discern latent connections and provide holistic, semantically-rich recommendations has rendered them particularly appealing. While various methodologies exist within this paradigm, such as path-based embeddings and unified methods, the most apt approach is contingent on the specific application context. While these systems offer considerable advantages, they also introduce computational challenges. Consequently, knowledge graph embedding methods have been devised to reduce the dimensionality of representations. However, these come with trade-offs, especially in terms of explainability. Recent innovations, such as HAGERec (Yang and Dong, 2020) and KGAT (Wang et al., 2019b), while reducing the need for manual meta-graph optimization, grapple with challenges like cold start issues and computational overheads.

3 Proposed Framework

3.1 Overview

This section is dedicated on the design of the ICALM Intelligent Tutoring System, which is designed to integrate entity extraction, knowledge graph construction, and constructive reflection writing. The knowledge graphs, termed concept maps herein, are explicitly presented to users. The system's functionality encompasses recommending subsequent topics based on the congruity with the present topic, user predilections, and/or the inclinations of other learners. Upon solicitation, the system offers a comprehensive exposition on a selected topic. Conclusively, after each instructional session, learners are prompted to compose a brief reflection essay, approximating a one-minute duration, to introspect on acquired knowledge. This reflection will be subsequently juxtaposed with results generated by an LLM model.

3.2 Automated Knowledge Graph Constructor

ICALM's core function is to transform flashcards into knowledge graphs and proffer recommendations for subsequent topics, contingent upon a user's content consumption history and the transitional dynamics observed within academic concepts.

Initially, this system harnesses the capabilities of KnowGL from IBM Research AI consortium (Rossiello et al., 2022). This model is adept at extracting salient facts from textual sources, subsequently converting them to linearized seman-

	AUC	Time(s)
RippleNet (Wang et al., 2018)	0.899	∼360
CKE (Zhang et al., 2016)	0.903	28.76
KGCN (Wang et al., 2019a)	0.921	48.75
CFKG (Ai et al., 2018)	0.921	31.58
KTUP (Cao et al., 2019)	0.864	47.59

Table 1: AUC of (Knowledge) Graph Based Recommendations

tic representation triplets that encompass subject, relation label, and object. Contrasting with the preliminary endeavors of Babel, this representation is exhaustive, embracing both ABox assertions and TBox meta-information regarding entity labels and classifications. During its training phase, KnowGL augments the foundational REBEL dataset (Huguet Cabot and Navigli, 2021) with ancillary entity labels, types, and their morphological forms. The model's principal objective revolves around optimizing the mapping outcomes through cross-entropy loss, concurrently maximizing the log-likelihood of fact generation.

The resulting triplets from KnowGL serve as the foundation for constructing knowledge graphs. These are further utilized by a recommender system to curate subsequent topic suggestions. This system incorporates RippleNet (Wang et al., 2018), an amalgam of embedding-based and path-based methodologies, to execute an end-to-end recommendation process. The RippleNet system, cognizant of knowledge-graph structures, initiates user preference propagation over a knowledge graph, and in iterative cycles, offers recommendations rooted in a user's antecedent selections. As a consequence, this model facilitates topic recommendations not merely based on relational chains but also varied elements situated on the peripheral ripples.

Subsequent tables showcase a comparative analysis between RippleNet (Wang et al., 2018), CKE Zhang et al., 2016, KGCN (Wang et al., 2019a), etc., focusing on metrics like AUC and temporal efficiency, sourced from RecBole's implementations (Zhao et al., 2022)

3.3 LLM-Based Task Generation

Given the prowess of Large Language Models (LLMs) in conventional Natural Language Processing paradigms, their bespoke fine-tuning for specialized tasks remains a formidable challenge. Therefore, this research endeavor involves refining the GPT model using pedagogical materials resembling textbooks. The refined model is then prompted to compose in-class portfolios that align with predetermined assessment criteria. The results, algorithmically generated, undergo offline scrutiny using tools like SocialIQA (for Arts) and PhysicalIQA (for STEM subjects) (Sap et al., 2019) and online juxtaposition with student inputs via Sentence-BERT (Reimers and Gurevych, 2019). This tool facilitates the generation of encodings for both inputs, enabling the computation of cosine similarity. Feedback mechanisms are also integrated, emphasizing the nuances of note quality, lucidity, and comprehensive coverage of curriculum content.

4 Future Work

In the realm of educational technology (EdTech), contemporary research explores the integration of Large Language Models (LLMs) and Knowledge Graphs (KGs) to optimize student engagement and learning outcomes. The potential optimizations in these methods give rise to essential research questions in the context of prevailing educational challenges:

- How do explicit representations of Knowledge Graphs influence student comprehension compared to simple turn-wise conversations with LLMs?
- Given the shortening attention spans, how can EdTech platforms, leveraging LLMs and KGs, offer bite-sized and graph-structured materials that resonate with standardized test frameworks?
- How can student-centered techniques, such as structured reflections and peer connections, be incorporated effectively into EdTech platforms powered by LLMs and/or KGs?

In the research titled "Language models as Knowledge Bases" by Petroni et al., it is suggested that models like BERT-large encompass relational knowledge that can rival structured knowledge bases (Petroni et al., 2019). This suggests that the traditional step of constructing a knowledge base might be replaced by modern LLMs. As a result, future research could focus on empirical evaluations determining the benefits of retaining a knowledge base versus adopting a rudimentary LLM summarizer.

Furthermore, Kapur and Toh's seminal work, "Learning from Productive Failures" (Kapur and Toh, 2015), delves into the potential drawbacks of over-scaffolding and providing overt solutions. While these approaches might boost short-term performance, they could hamper sustained learning. Kapur proposes the introduction of complex pre-class tests to encourage increased student engagement during lessons. This instructional strategy aligns well with methodologies like SET 49 - Student-Generated Rubrics (Barkley, 2009), which allow students to autonomously design course assessment criteria, emphasizing aspects they consider vital for evaluation. Integrating this instructional strategy with an LLM-based generative model could further enhance its efficacy.

References

Qingyao Ai, Vahid Azizi, Xu Chen, and Yongfeng Zhang. 2018. Learning heterogeneous knowledge base embeddings for explainable recommendation. *Algorithms*, 11(9):137.

Nicholas Athanassiou, Jeanne M. McNett, and Carol Harvey. 2003. Critical thinking in the management classroom: Bloom's taxonomy as a learning tool. *Journal of Management Education*, 27(5):533–555.

E.F. Barkley. 2009. *Student Engagement Techniques: A Handbook for College Faculty*. Higher and adult education series. Wiley.

Yixin Cao, Xiang Wang, Xiangnan He, Zikun Hu, and Tat-Seng Chua. 2019. Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences. In *The World Wide Web Conference*, WWW '19, page 151–161, New York, NY, USA. Association for Computing Machinery.

Michelene T. H. Chi and Ruth Wylie. 2014. The icap framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4):219–243.

Nelson Cowan. 2010. The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1):51–57.

Pere-Lluís Huguet Cabot and Roberto Navigli. 2021. REBEL: Relation extraction by end-to-end language generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 2370–2381, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Manu Kapur and Leslie Toh. 2015. *Learning from Productive Failure*, pages 213–227. Springer Singapore, Singapore.

Fabio Petroni, Tim Rocktäschel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, Alexander H. Miller, and Sebastian Riedel. 2019. Language models as knowledge bases?

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks.

Gaetano Rossiello, Md Faisal Mahbub Chowdhury, Nandana Mihindukulasooriya, Owen Cornec, and Alfio Massimiliano Gliozzo. 2022. Knowgl: Knowledge generation and linking from text.

Maarten Sap, Hannah Rashkin, Derek Chen, Ronan LeBras, and Yejin Choi. 2019. Socialiqa: Commonsense reasoning about social interactions.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph attention networks.

B.E. Walvoord, D.D. Stevens, and A.J. Levi. 2013. *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning*. Stylus Publishing.

Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2018. Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 417–426, New York, NY, USA. Association for Computing Machinery.

Hongwei Wang, Miao Zhao, Xing Xie, Wenjie Li, and Minyi Guo. 2019a. Knowledge graph convolutional networks for recommender systems. In *The World Wide Web Conference*. ACM.

Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019b. Kgat: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 950–958, New York, NY, USA. Association for Computing Machinery.

Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-based Recommendation with Graph Neural Networks. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*, volume 33 of *AAAI '19*, pages 346–353.

Chengfeng Xu, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Fuzhen Zhuang, Junhua Fang, and Xiaofang Zhou. 2019. Graph contextualized self-attention network for session-based recommendation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, IJCAI'19, page 3940–3946. AAAI Press.

Zuoxi Yang and Shoubin Dong. 2020. Hagerec: Hierarchical attention graph convolutional network incorporating knowledge graph for explainable recommendation. *Knowledge-Based Systems*, 204:106194.

Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 353–362, New York, NY, USA. Association for Computing Machinery.

Wayne Xin Zhao, Yupeng Hou, Xingyu Pan, Chen Yang, Zeyu Zhang, Zihan Lin, Jingsen Zhang, Shuqing Bian, Jiakai Tang, Wenqi Sun, et al. 2022. Recbole 2.0: Towards a more up-to-date recommendation library. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 4722–4726.