Singular Action, Complex Cognition: An Intelligent Tutoring System in Riichi Mahjong

Marshall An(✉), Mufei He, Wei Yao, and John Stamper

Carnegie Mellon University, Pittsburgh PA 15213, USA
{haokanga,mufeihe,weiyao2,jstamper}@andrew.cmu.edu

Abstract. This paper investigates the application of Intelligent Tutoring Systems (ITS) in facilitating the acquisition of complex cognitive skills within the context of Riichi Mahjong. Riichi Mahjong, also known as Japanese Mahjong, is a domain where singular actions are the result of complex cognitive processes. Utilizing Cognitive Task Analysis (CTA), we have constructed an expert cognitive model and developed an ITS that breaks down these actions into subgoals and subsequently into Knowledge Components (KCs). A pilot observational study provided preliminary evidence of the ITS's effectiveness in improving skill acquisition. By analyzing learning curves with DataShop and applying the Multimethod Approach to Data-Driven Redesign (MADDRED), we present the proposed ITS redesign to address over-practice. Our future research agenda is to conduct a randomized controlled trial to assess the efficacy of the original versus the redesigned ITS. This study contributes to the broader discourse on the efficacy of ITS in teaching skills where complex cognition manifests in singular actions, paving the way for future advancements in intelligent tutoring across various domains.

Keywords: Intelligent Tutoring Systems · Students Learning · Instructional design · Feedback · Educational Data Mining

1 Introduction

Imagine the climactic finale of a Riichi Mahjong [13] tournament. Our protagonist, in second place, is awaiting one more tile to complete their hand. In an intriguing twist, when a player in fourth place discards a tile that matches, the protagonist opts for a strategic pass. After several turns, as the leading player discards another matching tile, our protagonist declares hand completion without hesitation. This move secures points directly from the leader and satisfies the reversal condition to win the championship. This pivotal moment, though a singular observable move, encapsulates complex cognitive deliberation.

R. Ferreira Mello et al. (Eds.): EC-TEL 2024, LNCS 15160, pp. 51–56, 2024.
https://doi.org/10.1007/978-3-031-72312-4_4

Riichi Mahjong, also known as Japanese Mahjong, presents challenges in developing expertise. Every move requires learners to integrate and apply various Knowledge Components (KCs), which are discrete units of cognitive functions [6]. In particular, the same observable action might be the outcome of different combinations of diverse KCs, depending on the context, much like the varied situations that might prompt drivers to apply the brakes. To explore expertise manifesting through singular actions driven by complex cognition, we have chosen Riichi Mahjong as our investigative domain. With its extensive player base, including a substantial online community, Riichi Mahjong offers a vast research landscape, providing access to a wide range of study participants. The exploration of cognitive intelligence within Riichi Mahjong has attracted considerable interest from researchers, especially in the field of artificial intelligence (AI), resulting in multiple AI models designed to outperform human players [7–9]. However, these models lack the ability to explain their moves in a manner that novices can understand and learn from. To address this gap, we decided to investigate the use of Intelligent Tutoring Systems (ITS) in teaching Riichi Mahjong. Grounded in the ACT-R theory [2], ITSs utilize machine-readable and human-readable production rules to decompose complex problem-solving into smaller, manageable steps. Each step is associated with specific KC(s), thereby offering transparent and explainable cognitive models. This transparency facilitates the replication of expert cognitive processes by novices. Extensive research over the years has underscored the efficacy of ITS in diverse educational settings, establishing best practices for their design and implementation [1,11,12,14]. We started the investigation with the following research question: Is ITS an effective teaching method in domains where each action requires cognitive processes involving integrating and applying multiple Knowledge Components (KCs)?

2 ITS Design Informed by Cognitive Task Analysis

Riichi Mahjong is a game of imperfect information that necessitates a broad range of complex cognitive processes, including the application of sometimes conflicting judgment heuristics. To assess the efficacy of ITS in teaching Riichi Mahjong, our initial investigation focused on a scenario that, in spite of its inherent complexity, is comparatively deterministic and involves minimal conflicting judgments: the calculation of a winning hand's score. Mastery of score calculation and reporting after winning is a must-have for in-person games. In addition, score calculation informs not only strategic decisions before achieving a winning hand but also decision-making during clutch moments, such as skipping or declaring a win, as illustrated in the introductory story. Our work started with Cognitive Task Analysis (CTA) to identify the skills required for task execution [10]. Think-aloud studies [3] with experts facilitated the delineation of an expert cognitive model, depicted in Fig. 1. By dividing the complex score calculation process into discrete subgoals, we systematically categorized the numerous KCs using namespaces, as shown in Table 1. The ITS design, informed by CTA and KC categorization, decomposes each task into four subgoals, as demonstrated

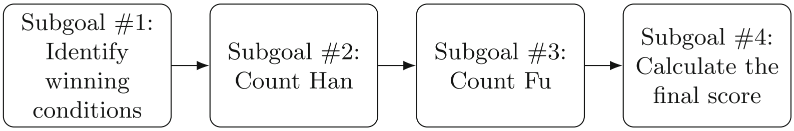

Fig. 1. The expert cognitive model which consists of four subgoals.

Table 1. Namespace Coding of KCs. This table showcases a small subset of the total identified KCs (over 50), illustrating that these KCs, although discrete, can be structured into hierarchical namespaces such as "Fu/Composition/*".

KC Coding w/ Namespaces	Verbose Explanation
Condition/ClosedOrOpen	Identify if a hand is Open or Closed
Condition/WinByRonOrTsumo	Identify if a hand wins by Ron or Tsumo
Fu/Wait/Type	Identify the wait type of a hand
Fu/Composition/Type	Identify the type of each composition in the hand
Fu/Composition/ClosedOrOpen	Identify if each composition is Open or Closed

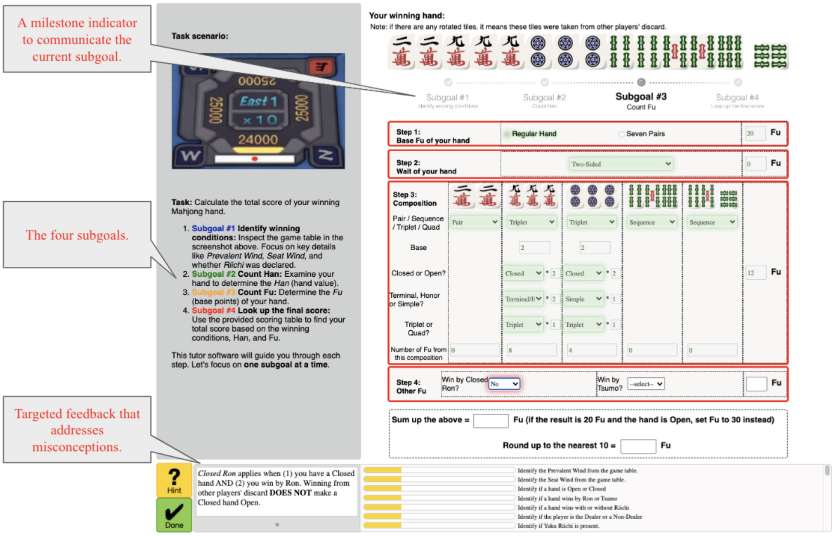

Fig. 2. The ITS allows students to focus on one subgoal at a time.

in Fig. 2. In addition to the features highlighted in Fig. 2, we adhered to established best practices for designing step-loop adaptivity [11,12] by providing hints at multiple levels for each step, as shown in Fig. 3.

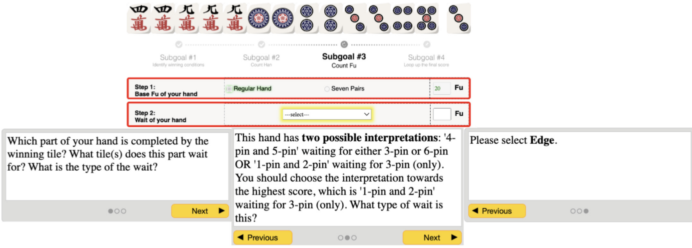

Fig. 3. Hints at multiple levels (facilitative, directive, and bottom-out) in a step.

3 Pilot Observational Study

We conducted an in-vivo experiment in a private research university course on Riichi Mahjong, with 11 students who consented to participate and completed all required activities. The study included a pre-test, an intervention with ITS, and a post-test. Two distinct yet equivalent 10-question quizzes were used for the pre- and post-tests, and each participant was randomly assigned one of the quizzes as their pre-test, with the other serving as their post-test. No feedback, including correctness, was provided during the pre-test. This study design aimed to mitigate potential biases, attributing differences in learning outcomes to the ITS intervention rather than to variations in assessment difficulty or practice effects from the pre-test. Figure 4 illustrates the distribution of pre- and post-test scores. To assess the impact of the intervention on test scores, we first conducted a Shapiro-Wilk test, which indicated the non-normal distribution of differences between paired pre- and post-test scores *($W = 0.83$, $p < .05$)*, necessitating the use of non-parametric methods for further analysis. Consequently, we opted for the non-parametric Wilcoxon signed-rank test to evaluate the intervention's effect. This test revealed a statistically significant difference in median scores post-intervention *($Z = 3.00$, $p < .05$)*.

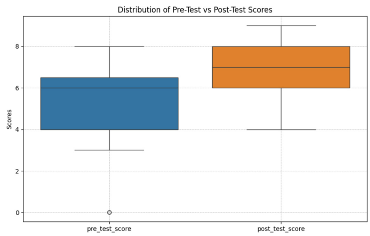

Fig. 4. Distribution of Pre- and Post-Test Performance

To investigate the relationship between practice and performance improvement, we conducted *learning curve* analysis using DataShop to identify trends of improvement or stagnation in learners' performance as they were exposed to an increasing number of practice opportunities [5]. Figure 5 illustrates our observations, including KCs that exhibited good learning curves, as well as instances of over-practice. Upon examining the KCs associated with over-practice, we observed that these steps were less associated with domain-specific Mahjong skills, such as pattern recognition within the hand, and more concerned with basic mathematical tasks, such as summing totals. Consequently, we are considering implementing a feature in the ITS that would automatically perform arithmetic calculations for the students.

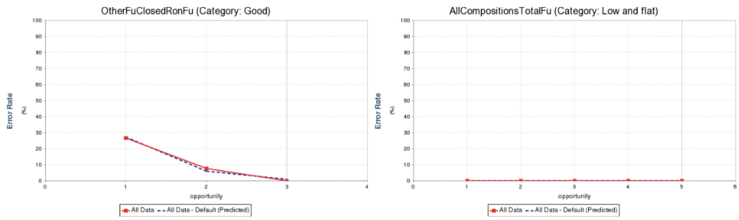

Fig. 5. DataShop learning curve analysis which revealed KCs with good learning curves (left), as well as instances of over-practice (right).

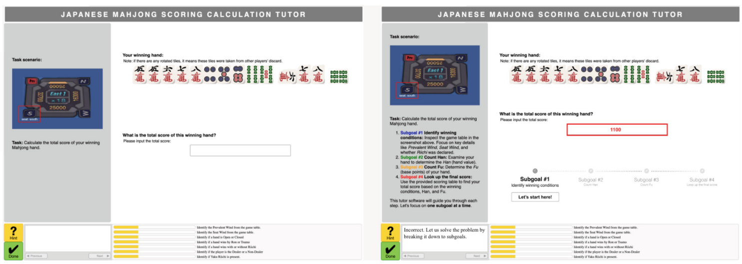

Fig. 6. Planned ITS redesign prototypes. Initially, the tutor requests only the total score, revealing detailed steps only if a learner's answer is incorrect. Problems are broken into subgoals, and further detailed steps are provided for unresolved subgoals.

4 Data-Driven Tutor Redesign and Future Work

Due to the small sample size and the necessity of a normal approximation for p-value calculation, we interpreted the initial results with caution. Nonetheless, the results were encouraging preliminary evidence of the efficacy of ITS. Given that each action necessitates the integration and application of multiple knowledge components (KCs), it is unsurprising that learners may require additional practice for certain KCs while having mastered others. Therefore, obligating learners to complete all detailed steps from start to finish repeatedly can lead to over-practice. To strike a balance between providing comprehensive end-to-end practice opportunities for novices and preventing unnecessary repetition for subgoals and KCs that learners have already mastered, a redesign capable of adapting to student mastery levels is needed. The adaptivity should include both the problem level, through task-loop adaptivity that automatically selects subsequent problems [11,12], and the step level, by determining which specific steps could be skipped upon mastery. We were inspired by the MADDRED (Multi-method Approach to Data-Driven REDesign) framework, a systematic combination of methods for the redesign of tutoring systems informed by data from prior iterations [4], particularly regarding adaptive scaffolding as demonstrated in Fig. 6. Our future research agenda is to conduct a randomized controlled trial to compare the redesigned tutor against the original iteration. We hypothesize that this redesign will effectively reduce over-practice and improve learning outcomes. Our long-term goal is to develop a suite of ITS for a variety of Riichi Mahjong skills.

References

1. Aleven, V., et al.: Example-tracing tutors: intelligent tutor development for non-programmers. Int. J. Artif. Intell. Educ. **26**, 224–269 (2016)
2. Anderson, J.R., Matessa, M., Lebiere, C.: Act-R: a theory of higher level cognition and its relation to visual attention. Hum. Comput. Interact. **12**(4), 439–462 (1997)
3. Ericsson, K.A., Simon, H.A.: How to study thinking in everyday life: contrasting think-aloud protocols with descriptions and explanations of thinking. Mind Cult. Act. **5**(3), 178–186 (1998)
4. Huang, Y., et al.: A general multi-method approach to data-driven redesign of tutoring systems. In: LAK21: 11th International Learning Analytics and Knowledge Conference, pp. 161–172 (2021)
5. Koedinger, K.R., Baker, R.S., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J.: A data repository for the EDM community: the PSLC datashop. Handb. Educ. Data Min. **43**, 43–56 (2010)
6. Koedinger, K.R., Corbett, A.T., Perfetti, C.: The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. Cogn. Sci. **36**(5), 757–798 (2012)
7. Kurita, M., Hoki, K.: Method for constructing artificial intelligence player with abstractions to markov decision processes in multiplayer game of mahjong. IEEE Trans. Games **13**(1), 99–110 (2020)
8. Li, J., et al.: Suphx: mastering mahjong with deep reinforcement learning. arXiv preprint arXiv:2003.13590 (2020)
9. Mizukami, N., Tsuruoka, Y.: Building a computer mahjong player based on monte carlo simulation and opponent models. In: 2015 IEEE Conference on Computational Intelligence and Games (CIG), pp. 275–283. IEEE (2015)
10. Schraagen, J.M., Chipman, S.F., Shalin, V.L.: Cognitive Task Analysis. Psychology Press (2000)
11. VanLehn, K.: The behavior of tutoring systems. Int. J. Artif. Intell. Educ. **16**(3), 227–265 (2006)
12. VanLehn, K.: Regulative loops, step loops and task loops. Int. J. Artif. Intell. Educ. **26**, 107–112 (2016)
13. Wikipedia contributors: Japanese mahjong — Wikipedia, the free encyclopedia (2024), https://en.wikipedia.org/w/index.php?title=Japanese_mahjong&oldid=1210039622. Accessed 1 March 2024
14. Woolf, B.P.: Building Intelligent Interactive Tutors: Student-centered Strategies For Revolutionizing E-learning. Morgan Kaufmann (2010)