CogKT: Cognitive Ability-based Knowledge Tracing towards Interpretable Knowledge Mastery Representation in Personalized Learning

1st Peiran Zhang
Tsinghua University
Beijing, China
zhangpr23@mails.tsinghua.edu.cn

2nd Nan He
Beijing University of Technology
Beijing, China
henancs@bjut.edu.cn

3rd Binbin Qi
South China Normal University
Guangzhou, China
qidoublebins@163.com

4th Lifeng Sun
Tsinghua University
Beijing, China
sunlf@tsinghua.edu.cn

***Abstract*—In Knowledge Tracing (KT) tasks, to address the black-box characteristic of knowledge mastery, recent work has proposed some interpretable methods to represent knowledge mastery. Educational theories suggest that learners apply knowledge concepts across different cognitive levels. Therefore, cognitive levels should be represented in the unit of knowledge mastery, which is overlooked by existing studies. We propose the concept of "cognitive ability" as the unit of knowledge mastery to represent cognitive levels, and propose a Cognitive Ability-based Knowledge Tracing (CogKT) method. In CogKT, the correlations between cognitive abilities are modeled using cognitive ability graphs and Graph Neural Network (GNN), while the temporal dynamics of cognitive abilities is captured using LSTM network enhanced by cognitive level embedding. Evaluated on a real-world dataset comprising 2393 students, CogKT outperforms baseline models on performance metrics AUC and ACC. In addition, cognitive abilities exhibit a significant influence on test outcomes, indicating that they contribute to enhancing the interpretability of knowledge mastery representation.**

***Index Terms*—knowledge tracing, deep neural network, knowledge mastery, interpretability, cognitive ability, cognitive level**

I. Introduction

The Knowledge Tracing (KT) task aims to predict learners' future learning performance based on historical learning records, which has significant application in personalized learning. The fundamental mechanism of KT involves analyzing learners' knowledge mastery from historical performance and then using the knowledge mastery to forecast future performance. Therefore, the representation of knowledge mastery lies at the core of the KT methods. Deep-Learning Knowledge Tracing (DLKT) methods model the temporal dynamics of knowledge mastery through a Recurrent Neural Network (RNN) [1], effectively improving the prediction accuracy of KT models.

In recent research on DLKT methods, the interpretability of knowledge mastery has garnered increasing attention. The reason lies in the fact that, from a practical application perspective, the fundamental goal of knowledge tracing methods is to ensure that knowledge mastery aligns with the actual learning state, rather than simply achieving accurate predictions of future learning performance. However, the objective of DNNs' optimization is to improve the accuracy of prediction. Since knowledge mastery in DLKT models exists in the form of hidden states, it exhibits a distinct black-box characteristic with low interpretability. Recent work has proposed several interpretable methods to represent knowledge mastery to overcome its black-box characteristic [2], [3], [4], [5].

Regarding interpretable methods for representing knowledge mastery, existing research has focused primarily on two aspects: the interpretable model structure and the real-world factors that influence knowledge mastery. Regarding the unit of knowledge mastery, prior studies have consistently adopted knowledge concepts as the unit of knowledge mastery, without analysis from this perspective. According to educational theories [6], [7], knowledge concepts as units have inherent limitations. In learning tasks, learners apply knowledge concepts at different cognitive levels. Taking the following two exercises involving the knowledge concept "quadratic equations" as an example:

- Determine: whether $(x-1)^2-1=0$ is a quadratic equation or not.
- Solve the quadratic equation: $(x-1)^2-1=0$

Although the two exercises share the same knowledge concept, the cognitive levels and complexity are different. Consequently, learners' mastery between different cognitive levels also varies. In existing methods, the unit of knowledge mastery is the knowledge concept, which fails to distinguish mastery between different cognitive levels. This limitation undermines the interpretability of knowledge mastery representation.

To address the aforementioned limitation, we propose the concept of "cognitive ability" as the unit of knowledge mastery instead of knowledge concept, and propose a cognitive ability-based knowledge tracing (CogKT) method. Based on educational cognitive theory, cognitive abilities integrate knowledge concepts with cognitive levels, thereby more comprehensively reflecting learners' knowledge mastery. In the CogKT method, we propose approaches to model the correlations and temporal dynamics of cognitive abilities, enabling the representation of knowledge mastery with cognitive abilities as the fundamental units. The main contributions of this study are as follows.

- Based on educational theory, we propose "cognitive ability" as the unit of knowledge mastery, which more comprehensively reflects the knowledge mastery of learners than the knowledge concept. Cognitive abilities are extracted from learning data using LLM agents.
- A knowledge tracing method called CogKT is proposed. Using a cognitive ability graph, a GNN adapted to the cognitive graph, and an LSTM enhanced with cognitive level embeddings, CogKT characterizes knowledge mastery with cognitive abilities as units.
- Evaluation in a real-world dataset comprising 2393 students shows that CogKT outperforms baseline methods, with the significant influence of cognitive ability. This indicates that cognitive abilities contribute to improving the interpretability of knowledge mastery representation.

II. Related Work

A. Deep-Learning Knowledge Tracing (DLKT)

DLKT methods employ Recurrent Neural Networks (RNN) to model the temporal dynamics of knowledge mastery based on learners' past performance and to predict their future learning outcomes. Since the introduction of the first Deep Knowledge Tracing (DKT) method [1], numerous new DLKT approaches have been proposed to improve the accuracy of the prediction. Improvement methods fall primarily into three directions:

1) Graph-based methods: These methods utilize graph structures to represent correlations between knowledge concepts and model these correlations using Graph Neural Networks (GNN). GKT is a typical graph-based KT method [8]. GIKT [9] and HGKT [10] extend the knowledge concept graph to exercise-knowledge relational graphs to capture high-order question-skill correlations. GAKT employs a Graph Attention Network (GAN) with attention mechanisms to capture concept correlations from the knowledge concept graph [11].

2) Memory-augmented methods: These methods introduce memory modules to store and update the mastery status of individual knowledge concepts, addressing the limitation of traditional KT models that represent each knowledge concept merely as a vector. This category is typified by DKVMN [12] and includes subsequent enhancements such as SKVMN [13] for sequence modeling and EKVMN [14] for exercise augmentation.

3) Attention-based methods: Attention mechanisms are widely adopted in deep learning models for sequential tasks due to their strong capability in handling sparse data generalization and parallel computation scenarios. This category is represented by typical methods such as SAKT [15] and AKT [16].

Among these three enhancement approaches, graph-based methods emphasize the representation of knowledge concept correlations. Given that knowledge concepts inherently exhibit graph structures and demonstrate knowledge transfer effects during learning, graph-based methods are found to be the most effective in handling knowledge concept correlations.

B. Interpretable DLKT

Since the fundamental goal of knowledge tracing methods is to align the knowledge mastery represented by the model with the actual learning state, research on interpretable knowledge mastery representation methods is gaining attention in recent DLKT works. This line of research primarily focuses on two aspects:

1) The structure of knowledge mastery modeling: Some methods examine the relationship between learning performance and knowledge mastery, arguing that knowledge concept is not the only factor that influences knowledge mastery [17], [18], [19]. They incorporate factors such as difficulty levels and exercise-knowledge correlations into the representation structure. Other methods introduce cognitive theories into the representation of knowledge mastery. For example, Deep-IRT [3] integrates the cognitive diagnostic theory IRT into KT methods, while GRKT incorporates theories such as the transfer of learning theory, testing effect theory, and memory theory into KT approaches.

2) Real-world factors that influence knowledge mastery: Some methods analyze real-world factors, such as learning scenarios and affective factors. The FDKT method focuses on analysis in federated learning scenarios [4], the AdaptKT method addresses cross-domain adaptation scenarios [5], the QKT method focuses on quiz scenarios [20], and the DASKT method incorporates learners' affective factors into KT approaches [21].

C. Cognitive Theory

Cognitive theories describe the cognitive activities of learners during learning tasks, reflecting the relationship between learning performance and knowledge mastery. As such, they can be incorporated by KT methods into the representation of knowledge mastery. Existing KT methods are grounded in cognitive diagnosis theories from cognitive theory. Most KT methods essentially rely on Classical Test Theory (CTT). Some recent studies have adopted Item Response Theory (IRT) as cognitive diagnosis approaches [3], [4].

Apart from cognitive diagnosis theory, some cognitive theories move beyond the commonly used classification of knowledge concepts and establish novel categorization methods for learning tasks. The Bloom taxonomy of educational objectives employs a classification framework to describe the levels of cognitive activities involved in learning tasks [6]. Building on this foundation, Marzano introduced the New Taxonomy of Educational Objectives (NTEO), which classifies learning tasks into two dimensions: the knowledge concept and the cognitive level [7]. Compared to the one-dimensional classification based solely on the knowledge concept, the two-dimensional classification of NTEO is more aligned with the actual learning state.

Our CogKT method refactors the unit of knowledge mastery based on the two-dimensional classification of NTEO, and uses graph-based KT method to represent the correlations between the unit of knowledge mastery.

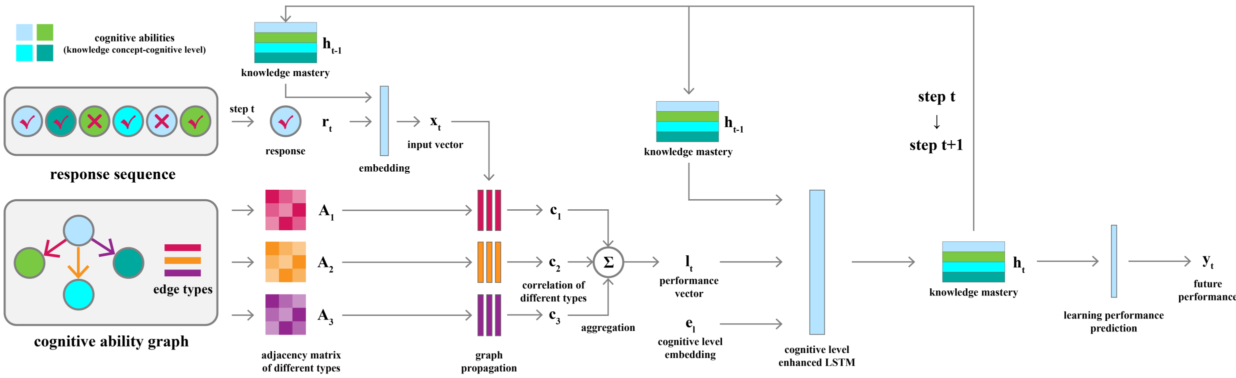

Fig. 1. The framework of CogKT.

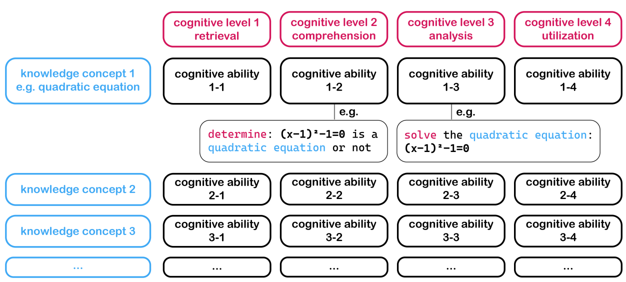

Fig. 2. Definition and implication of cognitive ability. The different level between two example exercises can be distinguished.

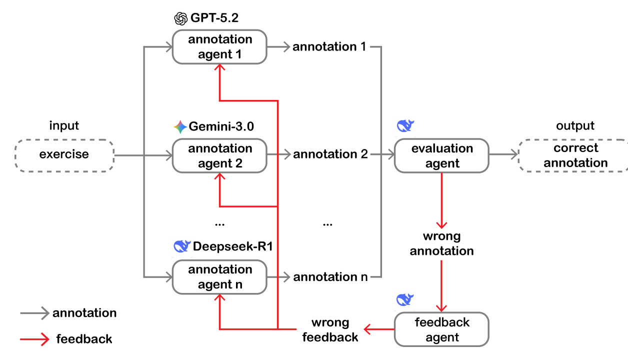

Fig. 3. Procedure of cognitive ability annotation.

III. Methodology

In this section, we present the detailed architecture of our CogKT method. First, we introduce the definition and acquisition approach of cognitive abilities. Then, based on our task formulation, we describe the procedure of the deep learning model in CogKT for modeling the correlations and temporal dynamics of cognitive abilities. The framework of our method is illustrated in Figure 1.

A. *Cognitive Ability*

Cognitive ability serves as the unit of knowledge mastery in the CogKT method. This section introduces its definition and extraction approach.

1) Definition: The concept of cognitive ability is grounded in the NTEO theory described in Section II-C, which classifies learning tasks along two dimensions, knowledge and cognition. We define "cognitive ability" as the joint classification of a learning task (e.g., an exercise) across both the knowledge and the cognitive dimensions. The classification along the knowledge dimension corresponds to the knowledge concept. For classification along the cognitive dimension, NTEO divides it into six levels: decision-making (self-system), planning (metacognitive system), retrieval, comprehension, analysis, and utilization [7]. Specifically, for exercise tasks that this study mainly concerns, since students are not required to make decisions or formulate plans when completing them, we classify exercise tasks into four cognitive levels: retrieval, comprehension, analysis, and utilization. The definition and implication of cognitive ability are illustrated in Figure 2.

2) Extraction approach: Learning data include the knowledge concepts of learning tasks, but do not contain information about their cognitive levels. Therefore, during the preprocessing of learning data, we employ a multi-agent system based on Large Language Models (LLM) to extract the cognitive level from the exercises.

Our multi-agent system includes 3 types of LLM agents: annotation agent, evaluation agent, and feedback agent. Multiple annotation agents individually extract the cognitive level of an exercise, and then an evaluation agent checks the correctness of the annotation agents. Subsequently, a feedback agent provides corrections for wrong annotations back to the annotation agents, thereby preventing similar mistakes in subsequent annotations. The collaboration mechanism in the multi-agent system is shown in Figure 3.

In our cognitive ability extraction experiment, LLM-based agents invoke the APIs of GPT-5.2, Gemini-3.0, and DeepSeek-R1.

B. *Task Formulation of Knowledge Tracing*

We formulate the KT task as future learning performance based on the learning dataset. The learning dataset $\{G,\ Q,\ R\}$ comprises:

- Cognitive graph $G = \{C,\ E\}$. C is the set of cognitive abilities. E is the set of edges.

- $Q = \{(q, c)\}$ is the set of questions. $c \in C$ denotes the cognitive ability associated with question q.
- $R = \{r\}$ is the set of response sequences. Each response sequence $r = \{r_1, r_2, \cdots, r_T \mid r_t = (q_t,\ a_t)\}$ represents the interaction sequence of a learner. Within r, the tuple $(q_t,\ a_t)$ indicates the response question $q_t \in Q$ and the response answer a_t in step t, where $a_t = 1$ denotes a correct response and $a_t = 0$ denotes an incorrect response.

When future learning performance prediction is performed, the input is a response sequence r. After the response at step t, the DLKT model predicts the response result $\hat{a}_{t+1}$ for step t+1 based on the responses $(q_1, a_1), (q_2, a_2), \ldots, (q_t, a_t)$ from steps 1 to t and the question q_{t+1} at step t+1. The prediction is considered correct if $\hat{a}_{t+1} = a_{t+1}$, and incorrect otherwise.

C. Procedure of Knowledge Tracing

Using learning data that contain cognitive ability extracted by the method described in Section III-A, a deep learning model performs the KT task formulated in Section III-B through cognitive ability correlation modeling, cognitive ability temporal dynamics modeling, and learning performance prediction.

1) Cognitive Ability Correlation Processing: For modeling the correlations between cognitive abilities, we construct a cognitive ability graph where cognitive abilities serve as nodes and employ a GNN adapted to this cognitive ability graph to capture their correlation. The function of the GNN is to derive the learning performance l_t at step t based on the response r_t at step t and the knowledge mastery h_{t-1} at step t-1. The GNN consists of an embedding layer, a graph propagation layer, and an aggregation layer.

Embedding layer: The response r_t at step t is first transformed through one-hot encoding [8] and then embedded as the embedding vector e_t of the response. This vector e_t is then concatenated with the knowledge mastery h_{t-1} at step t-1 to form the input vector x_t at step t:

$$e_t = W_r \cdot Onehot(r_t) \tag{1}$$

$$x_t = Concat(h_{t-1},\ e_t) \tag{2}$$

Graph propagation layer: The edge set E of the cognitive ability graph G is aggregated into 12 adjacency matrices A_i according to the 12 types of edges. Then a Multilayer Perceptron (MLP) is used to compute the correlation vectors of cognitive ability c_i for each of the 12 types. The 12 types are derived from the coupling of four categories of knowledge concept correlation, same, contain, be contained by, similar to, and three categories of cognitive level correlation, same level, higher level, lower level. Modeling cognitive ability correlations separately based on edge types ensures that the semantic meaning of the correlations is preserved. The computation is performed as follows:

$$E = \{(c_{in},\ c_{out},\ a)\}\ \rightarrow\ A_i = [a_{c_{in},\ c_{out}}] \tag{3}$$

$$c_i = A_i \cdot MLP(x_t) \tag{4}$$

Where the edge between c_{in} and c_{out} belongs to type i.

Aggregation layer: The cognitive ability correlation vectors c_i of the 12 types are aggregated to obtain the learning performance vector l_t at step t:

$$l_t = \sum_{i=1}^{12} c_i \tag{5}$$

2) Cognitive Ability Temporal Dynamics Modeling: To model the temporal variations of cognitive abilities, we take into account the distinct temporal variation patterns inherent to different cognitive levels. We propose an LSTM network enhanced with cognitive level embeddings to model the temporal dynamics of cognitive abilities. The function of the LSTM network is to derive the knowledge mastery state h_t at step t based on the learning performance l_t at step t and the knowledge mastery state h_{t-1} at step t-1. Knowledge mastery is initialized as h_0 where the mastery of each cognitive ability is represented by a vector of length n_h.

The LSTM network controls the temporal dynamics through its forget gate, input gate, and output gate. For different cognitive abilities, each cognitive level exhibits different patterns of temporal dynamics. However, the weight matrices in the traditional LSTM model struggle to capture the correlation between cognitive levels and their temporal dynamics. Therefore, we use a cognitive level embedding e_l to incorporate the cognitive level into the calculation of the forget gate, the input gate, and the output gate. Based on learning performance l_t at step t, the knowledge mastery h_{t-1} at step t-1, and cognitive level embedding e_l, the forget gate F_t, the input gate I_t, and the output gate O_t are calculated using the following formulas:

$$e_l = [\ cognitive\ level\ vector\ ] \cdot W_l \tag{6}$$

$$F_t = \sigma(l_t \cdot W_{xf}\ +\ h_{t-1} \cdot W_{hf}\ +\ e_l \cdot W_{lf}\ +\ b_f) \tag{7}$$

$$I_t = \sigma(l_t \cdot W_{xi}\ +\ h_{t-1} \cdot W_{hi}\ +\ e_l \cdot W_{li}\ +\ b_i) \tag{8}$$

$$O_t = \sigma(l_t \cdot W_{xo}\ +\ h_{t-1} \cdot W_{ho}\ +\ e_l \cdot W_{lo}\ +\ b_o) \tag{9}$$

The LSTM model utilizes the forget gate, the input gate, and the output gate to update the knowledge mastery state from step t-1 h_{t-1} to step t h_t, and simultaneously updates the memory unit m_t:

$$\hat{c}_t = tanh(l_t \cdot W_{xc}\ +\ h_{t-1} \cdot W_{hc}\ + b_c) \tag{10}$$

$$c_t = F_t \cdot c_{t-1}\ +\ I_t \cdot \hat{c}_t \tag{11}$$

$$h_t = O_t \cdot tanh(c_t) \tag{12}$$

3) Learning Performance Prediction: The final module predicts the learning performance y_t at step t+1 based on the knowledge mastery state h_t at step t, and predicts the answer $\hat{a}_{t+1}$ according to y_t and the question q_{t+1} to be answered at step t+1. A prediction layer is used to compute y_t as follows:

$$y_t = \sigma(h_t \cdot W_p\ +\ b_p) \tag{13}$$

The predicted answer for the question q_{t+1} answered in time step $t + 1$ is extracted from y_t:

TABLE I
STATISTICS OF DATASET

Sub-dataset	1	2	3	4	5	6	7	8	9	10	11	Total
Subject	Chinese	Chinese	Math	Math	English	English	Biology	History	Geography	Ethics	Science	
Educational stage	Primary	Middle	Primary	Middle	Primary	Middle	Middle	Middle	Middle	Middle	Primary	
Student num.	1,112	1,305	1,104	1,304	725	1,300	1,000	1,301	999	1,304	593	2,393
Question num.	595	358	382	524	287	391	391	927	466	552	227	5,100
Concept num.	39	123	1,993	1,387	400	1,018	408	942	1,009	885	801	9,005
Response num.	101,114	58,712	64,582	64,926	52,094	268,619	87,079	197,127	159,377	172,680	52,621	1,278,931

$$\hat{a}_{t+1} = o_{t+1} \cdot y_t \tag{14}$$

Here, o_{t+1} is the one-hot encoding of the cognition ability of q_{t+1}.

During the model training phase, after completing predictions for time steps 2 to T over a response sequence with total time steps T, the model's loss is measured using the cross-entropy loss to optimize the trainable parameters. The loss is calculated by:

$$loss = -\sum_{t=2}^{T} a_t log(\hat{a}_t) + (1 - a_t) log(1 - \hat{a}_t) \tag{15}$$

IV. EXPERIMENTS

A. Dataset

Public datasets for KT tasks, such as ASSIST, Statics, and Junyi, do not include cognitive abilities or the exercise contextual data required for extracting cognitive abilities and thus cannot meet the data requirements of this experiment. We constructed a learning dataset as described in Section III-A using data from 2393 students in a school. The academic stages covered in the learning dataset range from Grade 4 to Grade 6 in primary school and Grade 7 to Grade 9 in middle school. The subjects include nine disciplines: Chinese, Mathematics, English, Biology, History, Geography, Morality, and Science. The statistics of the learning data are summarized in Table I.

Since the correlations between data from different subjects and educational stages are considered negligible, we partitioned the whole dataset into 11 sub-datasets by subject and educational stage, as outlined in Table I. These 11 sub-datasets are treated independently during experimentation.

B. Baseline

To demonstrate the effectiveness of CogKT, we use DKT, GKT, GRKT, QIKT and DIMKT as baseline methods.

- DKT: a classic DLKT method [1].
- GKT: a typical graph-based KT method [8].
- GRKT: a recent graph-based KT method, using educational theories to enhance interpretability [2].
- QIKT: a recent DLKT method that incorporates exercise embeddings into the model architecture [19].
- DIMKT: A recent DLKT method that incorporates exercise difficulty as a factor influencing knowledge mastery [18].

C. Experimental Setup

In experiments, all methods use the same hyperparameters: batch size was 32, learning rate was 0.001, number of training epochs was 50, and the length of the knowledge concept state vectors representing each unit $n_h = 32$. The models were trained using 4 GeForce RTX4090 GPUs. For the responses in each sub-dataset, we randomly partitioned 80% of the records into the training set, with the remaining data constituting the test set.

D. Evaluation Metrics

We used average AUC and average ACC as evaluation metrics to assess the effectiveness of learning performance prediction. AUC and ACC were calculated independently for each sub-dataset. The weighted averages of AUC and ACC, weighted by the number of responses in each sub-dataset, were calculated to obtain the average AUC and the average ACC.

E. Performance Comparison

Table II presents a comparison of the results between CogKT and the baseline methods on two evaluation metrics for the effectiveness of learning performance prediction: average AUC and average ACC. The CogKT method outperforms the baseline methods in both metrics, indicating that the representation of knowledge mastery in KT contributes to improved effectiveness in predicting learning performance.

We observe that the CogKT modeling strategy for improving the representation of cognitive progress does not significantly outperform several baseline modeling strategies that improve the interpretability of the model. Regardless of the advantage of the baseline modeling strategies, we notice that the distribution of concepts in the data becomes a possible influencing factor. Refer to Table I, the ratio of concept number to question number is much higher than classic datasets, such as ASSIST, which makes the distribution of concepts sparse in the dataset. Therefore, the use of the concept of cognitive ability makes the distribution much sparser. The combination of knowledge concepts with cognitive categories rather than combining them may solve this problem.

F. Complementary Study of Cognitive Ability

In this study, our CogKT method proposes cognitive abilities instead of knowledge concepts as the unit of knowledge mastery. To validate the influence of cognitive abilities, we conducted the following complementary study on their extraction, specificity, and ablation.

TABLE II
METRICS ON LEARNING PERFORMANCE PREDICTION TASK

Methods	DKT	GKT	GRKT	QIKT	DIMKT	CogKT(**ours**)
Average AUC	0.7962	0.7446	0.8227	0.8164	0.8075	**0.8273**
Average ACC	0.7700	0.7387	0.7856	0.8182	0.8124	**0.8205**

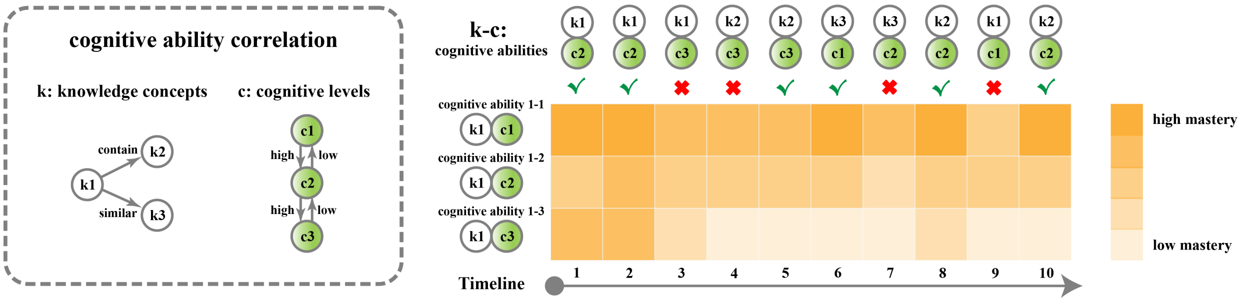

Fig. 4. A case of knowledge mastery representation process.

1) Extraction: Accurately extracting cognitive abilities from exercises is fundamental for the subsequent modeling of cognitive abilities. To validate the accuracy of cognitive ability extraction performed by our LLM agent system, we randomly selected the extraction results of 100 exercises and invited four experts to evaluate them. Expert evaluations indicated that 96% of the extraction is accurate, providing preliminary validation of the reliability of automated cognitive ability extraction via the LLM agent. Additionally, the extraction overhead of our LLM agent system averages approximately $0.03 per exercise, which is significantly lower than the manual cost under comparable conditions, thereby ensuring the feasibility of automated extraction for large-scale data.

2) Specificity: Cognitive ability is a unit that incorporates cognitive levels on the basis of knowledge concepts. Meanwhile, we observe that certain existing studies have also employed factors beyond knowledge concepts, such as difficulty [18], [19]. To validate the specificity of cognitive levels relative to difficulty, we characterize the difficulty of exercises at different cognitive levels using response accuracy. The relationship between difficulty and cognitive levels is presented in Table III.

TABLE III
SPECIFICITY OF COGNITIVE ABILITY

Cognitive Level	Exercise num.	Response num.	Difficulty
Retrieval(1)	963	422,908	**0.775**
Comprehension(2)	1,201	317,601	**0.785**
Analysis(3)	2,162	278,500	**0.847**
Utilization(4)	774	159,920	**0.772**

If cognitive levels were not specific with respect to difficulty, difficulty should increase as cognitive levels increase. However, the experimental results shown in Table III do not exhibit such a relationship, indicating that cognitive level is a specific factor that cannot be replaced by the difficulty factor employed in existing studies.

3) Ablation: To validate the impact of cognitive abilities on the performance of the CogKT model, we conducted ablation experiments on the modeling of cognitive ability correlations and the modeling of their temporal variations, respectively. In the ablation study on cognitive ability correlation modeling, we replaced the graph propagation layer and the aggregation layer with a single MLP. In the ablation experiment on cognitive ability temporal dynamics modeling, we set the cognitive level embedding e_l to zero. The experimental results are presented in Table IV:

TABLE IV
RESULTS OF ABLATION STUDY

Model	AUC	ACC
CogKT	0.8273	0.8205
w/o correlation modeling	0.8063	0.8139
w/o temporal dynamics modeling	0.8252	0.8116
w/o both	0.7824	0.7732

Both modeling of cognitive ability correlation and temporal dynamics contribute to improving model performance. Furthermore, even after removing these two modules, the model's performance remains higher than that of most baseline methods. The experimental results indicate that the cognitive abilities themselves, along with the model's representation of these abilities, improve the model's performance.

G. Case Study

To demonstrate the result of knowledge mastery representation, we conducted a case study. We extracted a process of knowledge mastery dynamics with the corresponding response sequence from the model and illustrated this process in Figure 4. For clear demonstration, the extracted knowledge mastery covers only three cognitive abilities. From the case, we notice that the correlation between knowledge concepts k1 and k2 has a more significant impact than k1 and k3 (step 4/5/6/7); and the correlation between cognitive concepts can also be represented (step 3/9). These observations demonstrate the correlation and temporal dynamics of cognitive abilities.

V. Conclusion

In this study, guided by the goal of enhancing the interpretability of knowledge mastery representation, we propose the CogKT method, which employs cognitive ability as the fundamental unit of the knowledge mastery. Through experiments on a real-world dataset, we validate the performance of the CogKT method and demonstrate the influence of cognitive ability in improving the interpretability of knowledge mastery representation. However, other enhancement strategies proposed in existing studies, such as memory networks, attention mechanisms, and other methods for improving model performance, have not been incorporated into our framework. How to effectively integrate cognitive ability with these enhancement strategies to further improve the performance of KT models remains an important direction for our future work.

Acknowledgment

This work is supported by National Key Research and Development Program of China No.2022ZD0115903.

References

[1] C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. Guibas, and J. Sohl-Dickstein, "Deep knowledge tracing," in *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 28 (NIPS 2015)*, ser. Advances in Neural Information Processing Systems, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., vol. 28, 2015, 29th Annual Conference on Neural Information Processing Systems (NIPS), Montreal, CANADA, DEC 07-12, 2015.

[2] J. Cui, H. Qian, B. Jiang, and W. Zhang, "Leveraging Pedagogical Theories to Understand Student Learning Process with Graph-based Reasonable Knowledge Tracing," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, ser. KDD. Barcelona Spain: ACM, Aug. 2024, pp. 502–513. [Online]. Available: https://dl.acm.org/doi/10.1145/3637528.3671853

[3] C.-K. Yeung, "Deep-irt: Make deep learning based knowledge tracing explainable using item response theory [arxiv]," *arXiv*, p. 10 pp., 26 April 2019.

[4] J. Wu, Z. Huang, Q. Liu, D. Lian, H. Wang, E. Chen, H. Ma, and S. Wang, "Federated Deep Knowledge Tracing," in *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, ser. WSDM. Virtual Event Israel: ACM, Mar. 2021, pp. 662–670, uSTC. [Online]. Available: https://dl.acm.org/doi/10.1145/3437963.3441747

[5] S. Cheng, Q. Liu, E. Chen, K. Zhang, Z. Huang, Y. Yin, X. Huang, and Y. Su, "AdaptKT: A Domain Adaptable Method for Knowledge Tracing," in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, ser. WSDM. Virtual Event AZ USA: ACM, Feb. 2022, pp. 123–131, uSTC. [Online]. Available: https://dl.acm.org/doi/10.1145/3488560.3498379

[6] B. Bloom, *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain*. New York: Longman, 1956.

[7] R. Marzano and J. Kendall, *The New Taxonomy of Educational Objectives (2nd ed.)*. Thousand Oaks CA: Corwin Pres, 2007.

[8] H. Nakagawa, Y. Iwasawa, and Y. Matsuo, "Graph-based Knowledge Tracing: Modeling Student Proficiency Using Graph Neural Network," in *IEEE/WIC/ACM International Conference on Web Intelligence*, ser. ICLR. Thessaloniki Greece: ACM, Oct. 2019, pp. 156–163. [Online]. Available: https://dl.acm.org/doi/10.1145/3350546.3352513

[9] Y. Yang, J. Shen, Y. Qu, Y. Liu, K. Wang, Y. Zhu, W. Zhang, and Y. Yu, "GIKT: A Graph-Based Interaction Model for Knowledge Tracing," in *Machine Learning and Knowledge Discovery in Databases*, ser. ECML-PKDD, F. Hutter, K. Kersting, J. Lijffijt, and I. Valera, Eds. Cham: Springer International Publishing, 2021, vol. 12457, pp. 299–315, sJTU.

[10] H. Tong, Z. Wang, Y. Zhou, S. Tong, W. Han, and Q. Liu, "Introducing Problem Schema with Hierarchical Exercise Graph for Knowledge Tracing," in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Madrid Spain: ACM, Jul. 2022, pp. 405–415. [Online]. Available: https://dl.acm.org/doi/10.1145/3477495.3532004

[11] Z. Zhao, Z. Liu, B. Wang, L. Ouyang, C. Wang, and Y. Ouyang, "Research on Deep Knowledge Tracing Model Integrating Graph Attention Network," in *2022 Prognostics and Health Management Conference (PHM-2022 London)*. London, United Kingdom: IEEE, May 2022, pp. 389–394. [Online]. Available: https://ieeexplore.ieee.org/document/9808616/

[12] J. Zhang, X. Shi, I. King, and D.-Y. Yeung, "Dynamic Key-Value Memory Networks for Knowledge Tracing," in *Proceedings of the 26th International Conference on World Wide Web*. Perth Australia: International World Wide Web Conferences Steering Committee, Apr. 2017, pp. 765–774. [Online]. Available: https://dl.acm.org/doi/10.1145/3038912.3052580

[13] G. Abdelrahman and Q. Wang, "Knowledge Tracing with Sequential Key-Value Memory Networks," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Paris France: ACM, Jul. 2019, pp. 175–184. [Online]. Available: https://dl.acm.org/doi/10.1145/3331184.3331195

[14] N. Zhang and L. Li, "Knowledge tracing with exercise-enhanced key-value memory networks," in *KNOWLEDGE SCIENCE, ENGINEERING AND MANAGEMENT, PT I*, ser. Lecture Notes in Artificial Intelligence, H. Qiu, C. Zhang, Z. Fei, M. Qiu, and S. Kung, Eds., vol. 12815. Springer LNCS; Waseda Univ; N Amer Chinese Talents Assoc; Longxiang High Tech Grp Inc, 2021, pp. 566–577, 14th International Conference on Knowledge Science, Engineering, and Management (KSEM), Tokyo, JAPAN, AUG 14-16, 2021.

[15] S. Pandey and G. Karypis, "A self-attentive model for knowledge tracing [arxiv]," *arXiv*, p. 6 pp., 16 July 2019.

[16] A. Ghosh, N. Heffernan, and A. S. Lan, "Context-Aware Attentive Knowledge Tracing," Jul. 2020, arXiv:2007.12324 [cs]. [Online]. Available: http://arxiv.org/abs/2007.12324

[17] S. Shen, E. Chen, Q. Liu, Z. Huang, W. Huang, Y. Yin, Y. Su, and S. Wang, "Monitoring student progress for learning process-consistent knowledge tracing," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 8, pp. 8213–8227, 2023.

[18] S. Shen, Z. Huang, Q. Liu, Y. Su, S. Wang, and E. Chen, "Assessing student's dynamic knowledge state by exploring the question difficulty effect," in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 427–437. [Online]. Available: https://doi.org/10.1145/3477495.3531939

[19] J. Chen, Z. Liu, S. Huang, Q. Liu, and W. Luo, "Improving interpretability of deep sequential knowledge tracing models with question-centric cognitive representations," in *THIRTY-SEVENTH AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, VOL 37 NO 12*, ser. AAAI Conference on Artificial Intelligence, B. Williams, Y. Chen, and J. Neville, Eds. Assoc Advancement Artificial Intelligence, 2023, pp. 14 196–14 204, 37th AAAI Conference on Artificial Intelligence (AAAI) / 35th Conference on Innovative Applications of Artificial Intelligence / 13th Symposium on Educational Advances in Artificial Intelligence, Washington, DC, FEB 07-14, 2023.

[20] S. Shen, Q. Liu, Z. Huang, L. Zhu, J. Lu, and K. Zhang, "Practicing in quiz, assessing in quiz: A quiz-based neural network approach for knowledge tracing," *NEURAL NETWORKS*, vol. 191, NOV 2025.

[21] X. Sun, K. Zhang, Q. Liu, S. Shen, F. Wang, Y. Guo, and E. Chen, "Daskt: A dynamic affect simulation method for knowledge tracing," *IEEE Transactions on Knowledge and Data Engineering*, vol. 37, no. 4, pp. 1714–1727, 2025.