

氏 名 Wenbin Gan

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 2331 号

学位授与の日付 2022 年 3 月 24 日

学位授与の要件 複合科学研究科 情報学専攻  
学位規則第6条第1項該当

学位論文題目 Dynamic Learner's Knowledge Assessment by Incorporating  
Learner and Domain Modeling in Intelligent Tutoring  
Systems

論文審査委員 主 査 孫 媛  
情報学専攻 准教授  
相澤 彰子  
情報学専攻 教授  
西澤 正己  
情報学専攻 准教授  
市瀬 龍太郎  
情報学専攻 准教授  
Yu Yi  
情報学専攻 助教  
緒方 広明  
京都大学 学術情報メディアセンター 教授

(Form 3)

## Summary of Doctoral Thesis

Name in full      Wenbin Gan

Title      Dynamic Learner's Knowledge Assessment by Incorporating Learner and Domain Modeling in Intelligent Tutoring Systems

The popularity of online learning has increased in recent years, with increasingly many intelligent tutoring systems (ITSs) becoming available to learners. In general, these platforms enable learners to acquire knowledge in the process of a series of individualized learning activities (e.g., exercising solving) that accommodate learners with different needs and knowledge proficiencies. A key technique underlying these adaptive tutoring services is learner knowledge assessment, which aims to model learner performance to discover their latent knowledge proficiency in mastering knowledge concepts in a domain.

This task can greatly facilitate the optimization of human learning because the assessment information is fundamental for the further adaptive services in many real-world ITSs. For example, adaptive remedial learning materials that filling the learners' knowledge gaps can be automatically provided based on learners' individual needs, and content that is predicted to be not in conformity with the learners' current knowledge states can be skipped or delayed, thereby effectively improving their learning efficiency and avoiding any decrement in their engagement. Meanwhile, given the popularity of a growing number of online educational platforms, a large number of learning logs can be collected, which provides the potential to build advanced models for accurate learner knowledge assessment. This has been a popular interdisciplinary research topic across education, psychology, computer science, and cognitive science.

Nevertheless, discovering learners' latent knowledge states from the long-term learning logs in an ITS is a rather challenging task, as human knowledge construction is a dynamic procedure and their knowledge is constantly evolving since learners dynamically learn and forget over time. Moreover, the knowledge attainment during learning is an extremely complex procedure and can be affected by many factors from both the learners and also the learning domains. The existing studies have explored this task and proposed various effective approaches mainly in two directions: cognitive diagnostic assessment (CDA) and knowledge tracing (KT). However, there are still numerous methodological issues, which restrict their practical applications. In this work, we address three important ones, namely, insufficient learning factor modeling, data sparseness and information loss, and fine-grained assessment and interpretability.

To tackle these issues, we proposed a general framework for dynamic learner knowledge assessment by integrating both learner and domain modeling. Based on this

framework, we proposed three approaches, each addressing one specific issue in the existing studies. Specifically, on the first issue of insufficient learning factor modeling, we investigated the learner factors (learning and forgetting) and domain factor (item difficulty) by utilizing rich information during learners' learning interactions and proposed a novel model named KTM-DLF that traces the evolution of learners' knowledge acquisition over time by explicitly modeling their learning and forgetting behaviors as well as the item difficulty. Extensive experiments confirmed the effectiveness of this model as it takes more and precise information into the modeling procedure. For the second issue of data sparseness and information loss, we explored to incorporate the knowledge structure (KS) into the learner knowledge assessment procedure to potentially resolve both the sparseness and information loss. We explored methods to infer the domain KS from the learner response data and integrated it with the original question–skill relation graph to enrich the data and alleviate the data sparseness. Then we proposed a novel KS-enhanced graph representation learning model (KSGKT) to learn the dense question and skill embeddings, and fused these embeddings with other distinctive features to obtain the comprehensive question representations, thus alleviating the information loss in the existing skill-level KT models. The distinctive representations are then utilized for KT with an attention mechanism. Extensive experiments demonstrated the superiority of the KSGKT model and the results proved it to be a good trial to alleviate the data sparseness and the information loss in conducting learner knowledge assessment. To cope with the third issue, we proposed a dynamic CDA model called KIEDCDA that incorporates not only the ability to trace the evolution of learners' knowledge proficiencies over time for large-scale assessments, but also the interpretability to explain learner performance in terms of their current knowledge proficiency and item characteristics. Experiments on several real-world datasets demonstrated the superiority and interpretability of the KIEDCDA model for learner performance modeling, suggesting that it is worthy of a good trial to track and explain learners' fine-grained and evolving knowledge states simultaneously.

This thesis presents our trials on dynamic learner knowledge assessment that enhanced the existing techniques. To further stimulate new ideas in the field of intelligent education, some remaining issues and future work are also described.

## 博士論文審査結果

Name in Full  
氏名 Wenbin Gan

Title  
論文題目 Dynamic Learner's Knowledge Assessment by Incorporating Learner and Domain Modeling in Intelligent Tutoring Systems

本学位論文は、知的学習支援システム (Intelligent Tutoring System: ITS) における学習者の動的知識習得状況の評価に関する課題に取り組んでいる。近年オンライン学習の普及に伴い、ITS の利用度も高まっている。知識や習熟度などにおいて学習者は多様であるが、ITS を利用することで、それぞれに適応した学習コンテンツによって個別化された学習活動を行い、学習内容を効率的に習得することができる。学習者に最適化した個別化学習を実現するためには、学習者の知識習得状況を適切に評価することが不可欠である。ITS を用いたオンライン学習活動の増加によって大量の学習ログが収集され、学習者の知識を正確に評価できる高度なモデル構築が、教育学、心理学、コンピューターサイエンス、および認知科学全体における学際的な研究ホットトピックとなっている。

しかし、ITS の長期学習ログから学習者の潜在的な知識状態を特定することは、かなり困難な作業である。人間の知識の構築は動的な現象であり、学習者は時間の経過とともに動的に習得し忘却するため、知識状態は絶えず変化している。さらに、学習中の知識の獲得は、学習者と学習ドメインの両者からの多様な要因の影響を受ける可能性がある。知識習得状況の評価について、これまで、認知診断アセスメント (Cognitive Diagnostic Assessment: CDA) と知識追跡 (Knowledge Tracing: KT) という 2 つの方向からさまざまな効果的なアプローチが提案されているが、従来の研究では解決されていない 3 つの重要な課題が残されている。すなわち、学習要因のモデリングが不十分であること、データの稀薄性と情報の損失、およびアセスメントのきめ細かさと解釈可能性という課題である。

本学位論文は、ITS における学習者とドメインモデリングの両方を統合し、学習者の動的な知識評価のための枠組みと、その枠組みに基づいた 3 つの重要課題に対処可能なアプローチを提案した。本論文は、7 章から構成されている。第 1 章では、近年盛んになってきたオンライン学習や個別化学習の意義と学習者知識状態の動的評価 (Dynamic Learner's Knowledge Assessment: DLKA) の概要、本論文で取り組んだ 3 つの研究課題および貢献を示している。第 2 章では、ITS の概要や、ITS における学習者とドメインモデリング、学習者の知識アセスメント方法、そして関連研究および問題点をまとめている。第 3 章では、論文で使用するデータセットと評価指標を説明している。続く 3 つの章で 3 つの課題を解決する 3 つの手法を提案し、本博士論文の主たる貢献を述べている。第 4 章では、不十分な学習要因モデリングという一つ目の研究課題を対処するため、学習パフォーマンスに影響する要因を定量化して学習者の知識の動的進化をモデル化する方法 (Knowledge Tracing Machine by modeling item Difficulty and Learning and Forgetting: KTM-DLF) を提案している。具体的には、動的知識の構築における学習と忘却効果、問題項目と学習

者の交互作用による適応型項目困難度をモデリングするアプローチの提案である。第 5 章では、学習ログから知識構造を自動生成した上、**KS-enhanced graph representation learning model for KT with an attention mechanism (KSGKT)** を提案し、二つ目の研究課題である学習データの稀薄性、情報損失を軽減できることを示している。第 6 章では、三つ目の研究課題であるきめ細かいアセスメントと解釈可能性に対して、学習の進化を評価できる動的認知診断モデル (**Knowledge Interaction-Enhanced Dynamic Cognitive Diagnostic Assessment: KIEDCDA**) を提案し、予測精度だけでなく結果の解釈可能性が向上することを示している。第 7 章では、これらの成果をまとめ、個々の提案モデルを統合して本論文の **DLKA** 全体に対する貢献について考察し、ITS における学習者知識アセスメントの将来課題について議論を行なっている。

公開発表会では博士論文の章立てに従って発表が行われ、その後に行われた論文審査会及び口述試験では、審査員からの質疑に対して適切に回答がなされた。

博士論文審査の結果、出願者は情報学分野の十分な知識と研究能力を持つと認められ、また研究内容は学位論文として十分なレベルの新規性や有効性があると認められた。また、本論文の内容に関し、査読付きジャーナル論文 2 編、査読付き国際会議論文 4 編を出版済みである。以上の理由により、審査委員会は、本学位論文が学位の授与に値すると判断した。