

Apprentissage par renforcement contraint guidé par un graphe de connaissances pour personnaliser les parcours d'apprentissage

Rania Ait Chabane^{1,2} Armelle Brun¹ Azim Roussanaly¹

(1) Université de Lorraine, LORIA (2) Stellia.ai

rania.ait-chabane@loria.fr, armelle.brun@loria.fr

RÉSUMÉ

Ce travail présente une architecture d'apprentissage adaptatif combinant graphes de connaissances enrichis et contraintes pédagogiques dans un cadre d'apprentissage par renforcement. Le graphe est construit à partir de ressources expertes (ex. manuel scolaire) et enrichi automatiquement par un modèle de langage pour compléter les relations et inférer des contraintes. Un module de knowledge tracing estime la progression de l'apprenant vers un objectif pédagogique donné. Un agent de renforcement, entraîné en environnement simulé, recommande des activités en maximisant la progression attendue tout en respectant les contraintes. Cette approche vise à renforcer la pertinence, la diversité et l'explicabilité des parcours proposés. Une évaluation sur des jeux de données réels est prévue en travaux futurs.

ABSTRACT

Constrained Reinforcement Learning Guided by a Knowledge Graph for Personalized Learning Pathways

This work presents an adaptive learning architecture that integrates enriched knowledge graphs and pedagogical constraints into a reinforcement learning framework. The graph is built from expert resources (e.g., textbooks) and enriched using a language model to infer semantic relations and constraints. A knowledge tracing module estimates the learner's progress toward a given objective. A reinforcement learning agent, trained in a simulated environment, recommends optimal activities by maximizing expected learning gains while respecting constraints. This approach aims to improve the relevance, diversity, and explainability of personalized learning paths. Future work will focus on evaluation using real-world educational datasets.

MOTS-CLÉS : Apprentissage adaptatif, Graphes de connaissances, Apprentissage par renforcement.

KEYWORDS: Adaptive learning, Knowledge graphs, Reinforcement learning.

ARTICLE : **Accepté à IA-ÉDU@CORIA-TALN 2025.**

1 Introduction

L'émergence des plateformes d'apprentissage en ligne a profondément modifié les pratiques éducatives, en ouvrant la voie à une personnalisation des parcours d'apprentissage. Le Knowledge Tracing (KT) est un levier central de cette personnalisation (Abdelrahman *et al.*, 2023). Il modélise l'évolution du niveau de maîtrise des concepts à partir des traces d'apprentissage. Les premières approches bayésiennes (Corbett & Anderson, 1994; Käser *et al.*, 2017) adoptent une approche probabiliste. Le

Deep KT (Piech *et al.*, 2015) a ensuite introduit des réseaux neuronaux récurrents pour capturer des dépendances complexes au sein des séquences d’interactions. Plus récemment, le Self-Attentive KT (Pandey & Karypis, 2019) utilise des mécanismes d’attention pour pondérer l’importance des interactions passées. Toutefois, ces modèles restent limités par leur faible interprétabilité et par une prise en compte partielle du contexte. Le Graph-based KT (GKT) proposé par (Nakagawa *et al.*, 2019) adresse ce point en apprenant une structure sémantique implicite, c’est-à-dire à partir des interactions, tout en prédisant le niveau de maîtrise grâce aux *Graph Neural Network* (GNN). En parallèle, la modélisation explicite du domaine d’apprentissage via des graphes de connaissance (*Knowledge Graphs*-KG) s’est imposée comme une solution pour représenter les concepts, les compétences, les activités, leurs relations et leurs dépendances (Ji *et al.*, 2022). Ces graphes, en apportant une connaissance experte structurée, guident les modèles prédictifs et renforcent à la fois leurs performances et leur explicabilité (Xian *et al.*, 2019). Des travaux récents mobilisent aussi les *Large Language Models* (LLMs) pour construire ou enrichir des KG éducatifs. Ainsi, (Jhaji *et al.*, 2024) utilisent GPT-4 pour générer des EduKG alignés sur des objectifs pédagogiques, et (Abu-Rasheed *et al.*, 2025) utilisent un LLM pour la complétion de graphes de parcours universitaires. Enfin, l’élément clé de la personnalisation est l’algorithme dit d’*adaptive learning* chargé d’identifier l’activité à proposer à l’apprenant à chaque étape de son parcours. Plusieurs approches sont explorées : bandits multi-bras (Mui *et al.*, 2021), classification non supervisée (Chen & Huang, 2023), ou systèmes à base de règles (Kolekar *et al.*, 2019). Toutefois, le *Reinforcement Learning* (RL) est privilégié grâce à sa capacité à optimiser séquentiellement les décisions sur le long terme selon l’évolution de l’état de l’apprenant. Des architectures hybrides combinant KG, KT et RL ont récemment vu le jour. Par exemple, TGKT-RL (Chen *et al.*, 2023) s’appuie sur un graphe question-compétence et un module KT fondé sur GAT+TCN pour guider l’apprenant vers une maîtrise ciblée.

Ces approches présentent deux limites majeures : une sémantique des graphes de connaissances pauvre, et l’absence de contraintes pédagogiques encadrant l’apprentissage par renforcement. Pour y répondre, nous proposons (i) d’exploiter un KG, issu d’une expertise humaine et enrichi sémantiquement par LLM et (ii) d’intégrer des contraintes pédagogiques pour guider l’exploration dans le processus RL.

2 Contribution

L’architecture proposée dans ce projet de recherche repose sur trois étapes, présentées Figure 1 :

L’étape 1 vise à une représentation sémantique riche du domaine pour adapter l’apprentissage de manière pertinente. Elle repose sur la construction d’un KG à partir d’une source experte, comme un manuel scolaire. Ce choix limite l’intervention d’experts humains, souvent coûteuse et chronophage (Cukurova *et al.*, 2023). Les éléments pédagogiques clés ainsi que leurs relations de précédence sont extraits via des techniques de TAL en suivant la structure hiérarchique du manuel. Les traces d’apprentissage issues d’une plateforme e-learning structurée selon le manuel, sont intégrées au KG. De façon optionnelle, pour pallier d’éventuelles insuffisances dans les connaissances extraites du manuel, un LLM est utilisé pour ajouter des relations sémantiques entre concepts. En complément, le LLM est mobilisé pour inférer automatiquement des contraintes pédagogiques à partir du KG et de l’objectif visé par l’apprenant. Ces contraintes peuvent inclure, par exemple, une modulation des prérequis en fonction du niveau actuel de l’apprenant. Elles sont ensuite intégrées au graphe pour guider plus finement la personnalisation du parcours d’apprentissage.

L’étape 2 est un module de KT, adapté pour aller au-delà de l’évaluation du niveau de connaissance d’une activité, il permet également d’évaluer le niveau de connaissance d’un objectif pédagogique (par

exemple les dérivées). Le contexte ici est plus riche que dans (Nakagawa *et al.*, 2019) : l’identifiant de l’apprenant, le KG enrichi, et l’objectif pédagogique. Le défi ici sera d’adapter les méthodes de KT sur des objectifs qui ne sont pas des activités.

L’étape 3 repose sur un agent d’apprentissage par renforcement (RL) entraîné de façon classique en deux temps. Dans un premier temps, l’agent interagit avec un environnement simulé, représenté par un module de Graph-based Knowledge Tracing (GKT) (Nakagawa *et al.*, 2019). Ce dernier utilise la structure du graphe de connaissances enrichi pour estimer, à chaque étape, l’évolution du niveau de maîtrise de l’apprenant en réponse à une activité donnée. Le GKT joue donc ici le rôle d’environnement simulé, dans lequel les conséquences pédagogiques des actions peuvent être prédites avant d’être proposées à l’apprenant réel.

La politique de l’agent d’apprentissage par renforcement (RL) vise à sélectionner l’activité pédagogique la plus pertinente en s’appuyant sur un réseau de neurones à graphes (GNN), exploité à partir de la structure sémantique du graphe de connaissances (KG) et des contraintes pédagogiques inférées. Il est prévu de formaliser les contraintes (étape 1) de manière explicite. D’une part, les contraintes fortes, telles que le respect des prérequis, doivent être directement intégrées à la politique de l’agent en restreignant l’espace des actions admissibles. D’autre part, les contraintes faibles, comme la promotion de la diversité des types d’activités, interviennent dans la fonction de récompense, via une modulation de sa valeur. Cette récompense est définie comme la variation estimée du niveau de maîtrise de l’objectif d’apprentissage, évaluée suite à l’exécution simulée d’une activité par l’agent.

Dans un second temps, l’agent communique l’activité choisie à l’apprenant réel. Une fois l’activité réalisée, la nouvelle trace d’apprentissage est intégrée au KG. Ce dernier évolue dynamiquement en intégrant les nouvelles connaissances observées chez l’apprenant, ce qui lui permet de jouer un double rôle : à la fois représentation du domaine et le suivi temporel de la progression individuelle de l’apprenant. Un nouveau cycle d’interaction est alors relancé jusqu’à atteinte de l’objectif pédagogique.

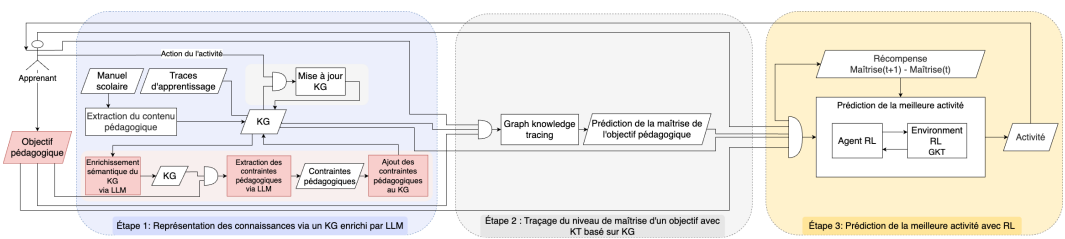


FIGURE 1 – Architecture générale du système proposé.

3 Conclusion

Les prochaines étapes concerneront l’implémentation et l’évaluation de l’architecture à partir de jeux de données comme ASSISTments. Les performances seront évaluées selon trois critères : (1) l’amélioration du niveau de maîtrise, mesurée par l’AUC et le RMSE entre les prédictions du module de Knowledge Tracing et les réponses réelles des apprenants ; (2) la diversité des parcours recommandés, évaluée par l’entropie moyenne des types d’activités proposées et la distance de Jaccard entre séquences d’apprenants ; (3) l’explicabilité des recommandations, analysée qualitativement par des experts et quantitativement via la longueur moyenne des chaînes de raisonnement extraites du graphe.

Références

- ABDELRAHMAN G., WANG Q. & NUNES B. (2023). Knowledge tracing : A survey. *ACM Computing Surveys*, **55**(11), 1–37.
- ABU-RASHEED H. *et al.* (2025). Llm-assisted knowledge graph completion for curriculum and domain modelling in personalized higher education recommendations. *arXiv preprint arXiv :2501.12300*.
- CHEN L. & HUANG H. (2023). Adaptive e-learning system based on learner portraits and knowledge graph. In *2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, volume 3.
- CHEN Z., WANG Q. & NUNES B. (2023). Tgkt-based personalized learning path recommendation with reinforcement learning. In *International Conference on Knowledge Science, Engineering and Management* : Springer Nature Switzerland.
- CORBETT A. T. & ANDERSON J. R. (1994). Knowledge tracing : Modeling the acquisition of procedural knowledge. In *User Modeling and User-Adapted Interaction*, volume 4, p. 253–278.
- CUKUROVA M. *et al.* (2023). Adoption of artificial intelligence in schools : Unveiling factors influencing teachers' engagement. In *International conference on artificial intelligence in education* : Springer Nature Switzerland.
- JHAJJ G. *et al.* (2024). Educational knowledge graph creation and augmentation via llms. In *International Conference on Intelligent Tutoring Systems* : Springer Nature Switzerland.
- Ji S., PAN S., CAMBRIA E., MARTTINEN P. & YU P. S. (2022). A survey on knowledge graphs : Representation, acquisition and applications. *IEEE Transactions on Neural Networks and Learning Systems*, **33**(2), 494–514.
- KÄSER T., KLINGLER S., SCHWING A. & GROSS M. (2017). Dynamic bayesian networks for student modeling. *IEEE Transactions on Learning Technologies*, **10**(4), 450–462.
- KOLEKAR S. V., PAI R. M. & PAI M. M. (2019). Rule based adaptive user interface for adaptive e-learning system. *Education and Information Technologies*, **24**, 613–641.
- MUI J., LIN F. & DEWAN M. A. A. (2021). Multi-armed bandit algorithms for adaptive learning : A survey. In *International Conference on Artificial Intelligence in Education* : Springer International Publishing.
- NAKAGAWA H., IWASAWA Y. & MATSUO Y. (2019). Graph-based knowledge tracing : Modeling student proficiency using graph neural network. In *IEEE/WIC/ACM International Conference on Web Intelligence*.
- PANDEY S. & KARYPIS G. (2019). A self-attentive model for knowledge tracing. *arXiv preprint arXiv :1907.06837*.
- PIECH C., BASSEN J., HUANG J., GANGULI S., SAHAMI M., GUIBAS L. & SOHL-DICKSTEIN J. (2015). Deep knowledge tracing. In *Advances in Neural Information Processing Systems*, volume 28.
- XIAN Y., FU Z., WANG S., ZHANG Z., YAO Y. & SUN S. (2019). Reinforcement knowledge graph reasoning for explainable recommendation. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*.