

Système intelligent de décision de feedbacks au sein de plateformes d'apprentissage humain : une approche apprentissage par renforcement

Contexte du projet de recherche doctoral

Ce projet de recherche doctoral a pour objectif de travailler sur un système de décision adaptative de feedbacks à destination d'élèves au sein de plateformes d'apprentissage humain (par ex. pour l'enseignement des mathématiques ou de la pensée informatique). Dans une plateforme d'apprentissage, l'élève réalise des exercices, organisés au sein de parcours, durant lesquels il bénéficie de feedbacks. La production des exercices, des parcours et la décision des feedbacks sont fondés épistémologiquement et didactiquement. C'est sur la dimension feedbacks adaptatifs que nous centrons ce projet doctoral.

Nous considérons un feedback comme toute décision prise par le système (aides, rappel de cours, tests, exercices à faire, etc.) et qui permet à l'élève de progresser et d'améliorer sa maîtrise des concepts du domaine à enseigner (Hattie et al., 2007). Nous travaillerons en particulier sur l'aide « à la prochaine étape [...] qui implique que le système de décision puisse identifier la solution travaillée par l'élève pour le guider vers la prochaine action à poser » (Tessier-Baillargeon et al., 2017) tels que définis dans plusieurs travaux (Aleven, 2013) (Efremov et al., 2020). Il existe de multiples approches pour décrire et catégoriser les feedbacks. Par exemple, Leibold & Schwarz (2015) signalent quatre types de feedbacks selon la nature des informations qu'ils contiennent (feedback correctif - feedback épistémique - feedbacks suggestif - feedback épistémique et suggestif).

Problématique du projet de recherche doctoral

Ce projet de recherche doctoral traite de la problématique de la décision du feedback le plus adapté aux caractéristiques de l'élève, de l'état de sa résolution, de l'exercice et d'autres attributs de l'environnement qui sont à définir. Il s'agit, de déterminer quel feedback est adapté pour un élève afin de lui permettre une progression optimale dans son parcours d'apprentissage. Ce problème de décision des feedbacks adaptatifs peut être modélisé comme un problème de décision dans l'incertain puisque les données sur l'état de l'environnement (l'élève, la résolution de l'élève, les caractéristiques de l'exercice etc.) sont imprécises, bruitées et incertaines. La décision est également incertaine due à l'absence de consensus entre les experts en didactique sur le feedback le plus adapté dans une situation de résolution de problèmes donnée.

Le travail de thèse s'appuiera sur les travaux menés dans le cadre du projet MindMath (FUI, 2017-2021) et qui ont permis d'explorer différentes pistes de recherche. Nous travaillerons sur la décision des feedbacks adaptatifs en adaptant des algorithmes d'apprentissage par renforcement. Plusieurs travaux (He-Yueya and al., 2021, Efremov et al., 2020, Bassen and al. 2020, Chi and al. 2011, Doroudi et al. 2019) ont montré l'intérêt de l'apprentissage par renforcement pour l'élaboration de politiques pédagogiques, car les modèles d'apprentissage par renforcement peuvent apprendre des relations complexes et souvent latentes entre les tâches pédagogiques, les actions des élèves et leurs acquis en termes de connaissances.

Il s'agit ici d'apprendre une politique de décision de feedbacks qui maximise les gains d'apprentissage des élèves. La modélisation du problème et le choix de la fonction de récompense sont déterminants. Un travail important d'explicitation des connaissances des experts en didactique est nécessaire en amont de l'apprentissage automatique du modèle de décision sur les données des élèves. Nous faisons l'hypothèse que l'association entre des connaissances symboliques issues de l'expertise humaine et des connaissances apprises implicitement par l'algorithme de renforcement permettrait une décision des feedbacks, d'une part, mieux adaptée aux élèves et facilitant leur progression dans le domaine à enseigner et, d'autre part, explicable pour les enseignants.

Des algorithmes d'apprentissage par renforcement seront utilisés pour apprendre certains paramètres de la décision des feedbacks qui sont difficiles à déterminer par les experts et qui probablement dépendent du profil des élèves et leur rapport à la matière enseignée.

L'agenda du travail de thèse sera le suivant :

- Réaliser un état de l'art exhaustif des algorithmes et des techniques de l'intelligence artificielle utilisés pour décider des feedbacks,
- Adapter et tester différents algorithmes d'apprentissage par renforcement (Qlearning, Bandit manchot, etc.),
- Modéliser une fonction de récompense adaptée et apprendre une politique décisionnelle adaptée au contexte des plateformes d'apprentissage humain,
- Evaluer la solution proposée, vis-à-vis, d'autres solutions de l'état de l'art et sur des corpus de données différents.

Références bibliographiques

- V. Aleven (2013). Help seeking and intelligent tutoring systems: Theoretical perspectives and a step towards theoretical integration. In: Azevedo, R., & Aleven, V. (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 311–335). New York: Springer. doi: [10.1007/978-1-4419-5546-3_21](https://doi.org/10.1007/978-1-4419-5546-3_21).
- J. Bassen, B. Balaji, M. Schaarschmidt, C. Thille, J. Painter, D. Zimmaro, A. Games, E. Fast, and J. C. Mitchell (2020). Reinforcement learning for the adaptive scheduling of educational activities. In CHI, pages 1–12.
- M. Chi, K. VanLehn, D. Litman, and P. Jordan (2011). Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction*, 21(1):137–180
- A. Efremov, A. Ghosh, and A. Singla. Zero-shot learning of hint policy via reinforcement learning and program synthesis. In EDM, 2020.
- J. Hattie & H. Timperley (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112.
- J. He-Yueya, & A. Singla (2021). Quizzing Policy Using Reinforcement Learning for Inferring the Student Knowledge State. EDM. International Educational Data Mining Society
- S. Doroudi, V. Aleven, and E. Brunskill. Where's the reward? *International Journal of Artificial Intelligence in Education*, 29(4):568–620, 2019
- N. Leibold & L.M. Schwarz (2015). The art of giving online feedback. *Journal of Effective Teaching*, 15(1), 34-46.

- J. Nouri, M. Saqr & Fors, U. (2019). Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback. In *10th International Conference on Education, Training and Informatics (ICETI 2019)* (Vol. 17, No. 4, pp. 17-21).
- A. Pardo, O. Poquet, R. Martínez-Maldonado & S. Dawson (2017). Provision of data-driven student feedback in la & EDM. *Handbook of learning analytics*, 163-174.
- V. Shute, (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.

Résumé en anglais

The objective of this PhD project is to investigate the decision process of epistemic feedbacks in training platforms. We aim to model the decision process of feedback as a decision under uncertainty and focus the research on the study and the adaptation of reinforcement algorithms to adapt feedbacks to the learner's profile and its problem solving. AI technics will be combined to human expertise in order to learn efficiently the parameters of the decision model.

Résumé en français

Ce projet de recherche doctoral traite de la problématique de la décision des feedbacks adaptatifs au sein de plateformes d'apprentissage humain. Il s'agit, de déterminer quel contenu de feedback est adapté pour un élève afin de lui permettre une progression optimale dans son parcours d'apprentissage. Ce problème de décision des feedbacks adaptatifs peut être modélisé comme un problème de décision dans l'incertain puisque les données sur l'état de l'environnement (l'élève, la résolution de l'élève, les caractéristiques de l'exercice etc.) sont imprécises, bruitées et incertaines. La décision est également incertaine due à l'absence de consensus entre les experts en didactique sur le feedback le plus adapté dans une situation de résolution de problèmes donnée. Des techniques d'apprentissage par renforcement seront combinées à de l'expertise humaine pour apprendre efficacement les paramètres de la décision des feedbacks adaptatifs.

Commenté [AY1]: Si besoin

