Tuteurs intelligents : boucler la boucle

Yolaine Bourda, Claude Chaudet, Benoît Choffin, Jeanne Parmentier, Fabrice Popineau, Jill-Jênn Vie

▶ **To cite this version:**

Yolaine Bourda, Claude Chaudet, Benoît Choffin, Jeanne Parmentier, Fabrice Popineau, et al.. Tuteurs intelligents : boucler la boucle. 2018. hal-02197685

HAL Id: hal-02197685
https://hal.science/hal-02197685

Submitted on 30 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Tuteurs intelligents : boucler la boucle

Yolaine Bourda[1] Claude Chaudet[4] Benoît Choffin[1]
Jeanne Parmentier[3] Fabrice Popineau[1] Jill-Jênn Vie[2]

[1] CentraleSupélec/LRI
[2] RIKEN AIP
[3] Institut Villebon – Georges Charpak
[4] Webster University

{yolaine.bourda,benoit.choffin,fabrice.popineau}@lri.fr
jill-jenn.vie@riken.jp
chaudet@webster.ch
jeanne.parmentier@villebon-charpak.fr

Résumé

La réalisation d'un tuteur interactif personnalisé est un challenge extrêmement ardu. Les approches utilisant des techniques d'apprentissage automatique essayent fréquemment de prédire la prochaine action du tuteur face à un apprenant sur la base des comportements observés d'une cohorte passée. Nous argumentons ici en faveur d'un changement de perspective et nous proposons de nous intéresser à l'interaction entre le tuteur et l'apprenant, et plus particulièrement à l'échange d'information qui a lieu entre eux. Nous arguons qu'il faut « boucler la boucle » entre le tuteur et l'apprenant. Pour illustrer cette approche, nous présentons un exemple d'application sous la forme d'un outil d'entraînement en ligne et identifions quelques questions de recherche que cette vision soulève.

1 Problématique

Durant leur cours, les professeurs adaptent le contenu de leur enseignement et la manière dont il est dispensé, et ce dès le début, alors que les élèves leur sont encore inconnus. Cette adaptation se produit en temps réel à un niveau micro (celui de l'heure de cours), mais aussi à un niveau plus macro (celui de la semaine, du trimestre, ou même du semestre). Elle se fonde en outre sur des signaux de différentes natures émis par les élèves et reçus par le professeur : questions en cours, attention en classe, résultats à des tests, discussions avec les concernés,...

Sur une plateforme d'apprentissage en ligne (telle que Moodle, OpenEdX, Coursera,...) , il n'y a généralement pas de professeur pour adapter en cours de route l'enseignement à la cohorte d'apprenants. En outre, l'intelligence artificielle est une véritable opportunité pour réaliser l'adaptation au niveau de l'élève. Le rôle d'un tuteur intelligent est donc d'adapter son comportement de manière autonome, comme le ferait un enseignant. Il doit être réactif et choisir la meilleure manière de faire (contenu différent ou présentation différente) quand il se trouve face à un élève – que ce dernier réagisse selon les attentes du tuteur ou non.

Malgré l'essor récent de l'apprentissage automatique (*machine learning*), et la mise en évidence précoce des gains de l'apprentissage *humain* liés à la personnalisation de ce dernier, les LMS tardent encore à proposer une adaptation des activités pédagogiques à leurs apprenants [1].

Cette question de la personnalisation fait partie des sujets qui intéressent les disciplines de l'analytique de l'apprentissage et de la fouille de données éducatives. Une approche commune consiste à analyser des données historiques sur une cohorte passée d'étudiants : analyse de traces [3], captures diverses, tests adaptatifs [2]... dans l'objectif de construire un modèle d'apprenant. Ce modèle est ensuite utilisé pour prédire la prochaine action sur un nouvel apprenant : on suppose que la généralisation aura fonctionné. Nous sommes ici dans un contexte de *passive machine learning*. Cette approche ne nous paraît pas suffisante, pour deux raisons principales :

— Les données peuvent être biaisées par divers facteurs (préférences des apprenants, suivi de la trame d'un cours) et manquer de variabilité.
— Comment réagir face à un apprenant (ou une nouvelle activité) qu'on ne connaît pas ou bien dont le comportement est éloigné de ceux observés ? C'est le problème bien connu du *démarrage à froid*.

Une pédagogie rationnelle[8] se fonde sur un dialogue entre tuteur et apprenant. Le tuteur possède un modèle initial de l'apprenant qu'il met à jour au fur et à mesure de ses interactions avec cet apprenant, en fonction des informations que ce dernier lui renvoie. Un ITS doit donc adopter une approche **itérative** et **interactive** pour lui permettre

1. Des exceptions existent, bien entendu, mais elles restent marginales ou propriétaires (e.g. Knewton).

de corriger à chaque étape son modèle de l'apprenant (par exemple, si cet apprenant est atypique ou si son comportement évolue), voire de changer de modèle. Nous privilégions donc des algorithmes d'*online machine learning* dans les ITS [6, 1]. Cependant, il est à noter que l'adaptativité *online* d'un algorithme peut elle aussi générer un biais dans l'estimation de paramètres [5].

Néanmoins, nous soutenons que la spécificité de l'apprenant n'est aujourd'hui pas suffisamment prise en compte, même avec des méthodes *online*. [7] rappelle ainsi qu'une séquence d'activités pédagogiques optimale pour l'étudiant moyen risque fort d'être sous-optimale pour *chaque* étudiant en particulier. Mais d'autre part, les cohortes d'apprenants disponibles sont rarement suffisantes pour étalonner un système reposant sur de nombreux paramètres nécessaires à une personnalisation fine. Dans cette mesure, nous pouvons dire que la boucle n'est pas suffisamment bouclée entre le tuteur et l'apprenant.

Pour ces raisons, nous argumentons pour un changement de perspective et nous proposons d'étudier l'*échange d'information* qui a lieu entre le tuteur et l'apprenant pour améliorer le choix de la prochaine action du tuteur. De cette façon, nous cherchons à « boucler la boucle » entre l'enseignant et l'étudiant. Nous nous plaçons ainsi dans un contexte d'*iterative machine teaching* [4] où nous cherchons à optimiser l'apprentissage de l'apprenant par une sélection soigneuse de chaque activité, et ce *en ligne*. Cette sélection se fonde sur l'historique des observations individuelles de l'apprenant et également sur les données de cohorte passée, ainsi que sur les connaissances du domaine. Ce paradigme est illustré à la section 2.

2 Description d'un exemple d'application : l'*exerciseur*

Dans le but de tester différentes politiques et de glaner des données d'usage, nous avons commencé à développer un outil que nous avons nommé *exerciseur*. Cet outil est destiné à être déployé en ligne. Il est alimenté par :

- un graphe de prérequis de composantes de connaissances ;
- un ensemble de *ressources* explicatives indexées par rapport aux composantes de connaissance ;
- un ensemble d'*exercices* indexés également par rapport aux composantes de connaissances.

La procédure est la suivante : dans un premier temps, le tuteur pose un exercice à l'étudiant. Cet exercice teste des composantes de connaissance dont les prérequis sont supposés au moins partiellement maîtrisés. Selon la validité de sa réponse à l'exercice, le tuteur choisit de lui proposer une autre question ou bien une ressource explicative. La procédure s'arrête quand la couverture du graphe des composantes de connaissances est estimée suffisante.

Schématiquement, dans ce qui suit f et g représentent des degrés de maîtrise de composantes de connaissance pour un apprenant :

- l'exerciseur est initialisé pour assurer la maîtrise d'un ensemble $C = \{c_1, \ldots, c_K\}$ de composantes de connaissances ;
- le nouvel apprenant a_i se connecte ;
- l'exerciseur propose une activité A_1 permettant d'acquérir la maîtrise d'un certain élément de connaissance c_k ; le degré de maîtrise de c_k espéré à l'issue de cette activité est $f(a_i, c_k)$;
- l'apprenant reçoit A_1, qui lui procure un degré de maîtrise $g(a_i, c_k)$, qui peut être différent de $f(a_i, c_k)$;
- l'exerciseur fait passer un test à l'apprenant sur la composante de connaissance c_k ;
- l'exerciseur s'attend à observer un degré de maîtrise $f(a_i, c_k) \pm \varepsilon$ où ε est un bruit gaussien, ou une quantité due à des erreurs de slip ou guess ;
- l'exerciseur reçoit en fait une indication de degré de maîtrise $g(a_i, c_k) \pm \varepsilon$;
- l'exerciseur va proposer une activité A_2 permettant d'acquérir la maîtrise d'un certain élément de connaissance $c_{k'}$, et ainsi de suite.

On voit ici que l'exerciseur possède un modèle *a priori* de l'apprenant. Ce modèle se révèle incorrect, cependant l'exerciseur doit corriger son évaluation pour le choix de la prochaine action. Les questions qui se posent sont :

- comment l'exerciseur doit-il modifier son estimation de f ?
- comment combiner l'exploitation des données sur les étudiants déjà vus avec la mesure de l'apprentissage de l'étudiant a_i ?
- quelle partie de f ne dépend que de c_k ?

L'établissement d'une politique consiste à dire quel objectif l'exerciseur doit viser pour la prochaine activité. Plusieurs politiques de sélection d'une nouvelle activité sont disponibles. Un nouvel étudiant qui se connecte à l'exerciseur se voit assigner une politique au hasard pour la durée de son entraînement.

Cet exerciseur a pour vocation essentielle la collecte de données et le test de politiques de sélection. Il pourra toutefois être utilisé en complément de cours.

3 Questions pour la recherche

Ce changement de perspective ouvre un nouveau regard sur quelques questions de recherche que nous listons ci-dessous.

- Comment mesurer le gain d'apprentissage ? Pour citer Martin Riopel (UQÀM) : est-ce qu'un étudiant qui passe d'une maîtrise d'une composante de connaissance de 20% à 30% gagne autant, plus ou moins que celui qui passe de 80% à 90% ?
- Quel est le compromis entre raccourcir le temps passé sur un cours et assurer la durabilité de l'apprentissage ?
- Quel niveau de difficulté viser pour la prochaine activité ? Il y a un champ d'investigation en lien avec les sciences cognitives.
- Peut-on s'adapter efficacement à quelqu'un qui

comprend vite ou possède déjà un bagage sur le domaine tout comme à quelqu'un de plus lent ou plus novice ?
— Comment gérer la problématique liée à la dimension de l'espace de recherche ?
— Le protocole décrit plus haut suppose que l'apprenant suit la recommandation proposée par le système. Comment intégrer la possibilité que l'apprenant soit acteur et choisisse son activité ?
— Faut-il utiliser un modèle explicite de l'apprenant ou bien privilégier une approche « *model-free* » ?

Références

[1] Reddy, S., Levine, S., & Dragan, A. Accelerating Human Learning with Deep Reinforcement Learning.

[2] Vie, J. J., Popineau, F., Bruillard, É., & Bourda, Y. (2017). A review of recent advances in adaptive assessment. In *Learning Analytics : Fundaments, Applications, and Trends* (pp. 113-142). Springer, Cham.

[3] Wilson, K. H., Xiong, X., Khajah, M., Lindsey, R. V., Zhao, S., Karklin, Y., ... & Heffernan, N. (2016). Estimating student proficiency : Deep learning is not the panacea. In *In Neural Information Processing Systems, Workshop on Machine Learning for Education.*

[4] Zhu, X. (2015, January). Machine Teaching : An Inverse Problem to Machine Learning and an Approach Toward Optimal Education. In *AAAI* (pp. 4083-4087).

[5] Nie, X., Tian, X., Taylor, J., & Zou, J. (2017). Why adaptively collected data have negative bias and how to correct for it. *arXiv preprint arXiv :1708.01977.*

[6] Lan, A. S., & Baraniuk, R. G. (2016, June). A Contextual Bandits Framework for Personalized Learning Action Selection. In *EDM* (pp. 424-429).

[7] Clement, B., Roy, D., Oudeyer, P. Y., & Lopes, M. (2014). Online optimization of teaching sequences with multi-armed bandits. In *7th International Conference on Educational Data Mining.*

[8] Milli, S., Abbeel, P., & Mordatch, I. (2017). Interpretable and Pedagogical Examples. *arXiv preprint arXiv :1711.00694.*