

Modélisation d'un Apprentissage humain via ses Réponses Imparfaites et Elicitation

Mots-clés : Apprentissage actif, fonctions de croyance, données imparfaites

Unité de recherche : IRISA - UMR 6074

Equipe : DRUID

Site web de l'équipe : <http://www.irisa.fr/druid>

Directeur de thèse : Arnaud MARTIN

UMR 6074, Département D7 - Gestion des données et de la connaissance, Equipe DRUID

Co-encadrante : Constance THIERRY

UMR 6074, Département D7 - Gestion des données et de la connaissance, Equipe DRUID

Contact : constance.thierry@irisa.fr

Date de début de thèse : septembre 2023

Lieu : Lannion

Description du sujet de la thèse

De plus en plus d'apprenant, se tournent vers un nouveau système de cours à distance appelé MOOCs (Massive Open Online Course). D'après la plateforme française FunMOOC [1] en mars 2021, plus de 2,5 millions de personnes ont suscité près de 11 millions d'inscriptions à plus de 1800 sessions de cours en ligne diffusés. Il est possible d'étudier presque tous les sujets grâce à cette méthode d'apprentissage. Ces cours sont dispensés en ligne par de nombreuses universités du monde entier et sont généralement accessibles à tout public. Par ailleurs, des dizaines d'universités célèbres dans le monde proposent désormais des MOOC. L'Université de Rennes 1 propose elle-même plusieurs MOOC aussi bien pour le personnel que les étudiants. Nous pouvons par exemple citer le MOOC « Cybersécurité : comment sécuriser un site web » créé par l'ENSSAT en partenariat avec l'IMT [2].

Les cours en ligne consistent en des vidéos ou en des sessions de visioconférences auxquelles participe un très large public. Cette distance avec le participant, présente l'avantage pour ce dernier d'apprendre en tout lieu et à tout moment sans contraintes géographique et temporelle forte. Il peut ainsi par exemple visionner plusieurs fois l'enregistrement d'un cours. L'inconvénient pour le maître d'apprentissage qui conçoit ces cours est que l'apprentissage distanciel complexifie l'estimation de l'assimilation de l'information. Parmi les méthodes existantes pour évaluer l'acquisition de connaissances chez l'apprenant, l'utilisation de questions fermées comme des questionnaires à choix multiple (QCM) est la plus fréquente. Les QCM sont un type spécifique de questionnaire qui propose un ensemble restreint de réponses à une question. Ils peuvent être classifiés en deux catégories : les QCM à réponses uniques (QCM-RU) et les QCM à réponses multiples (QCM-RM). Pour les QCM-RU, l'étudiant ne peut choisir qu'une unique réponse parmi l'ensemble proposé.

Un problème inhérent aux QCM-RU est leur rigidité qui limite la capacité des répondants à s'exprimer. Ces questionnaires ne donnent pas la possibilité aux étudiants de rendre compte de leur ignorance, leur imprécision et leur incertitude. Par conséquent, si les répondants hésitent entre deux (ou plusieurs) réponses, ils en choisiront une au hasard. Le problème dans le domaine de l'éducation est qu'une réponse aléatoire ne reflète pas exactement le niveau réel des connaissances de l'apprenant.

Afin de surmonter ce problème, plusieurs études ont été menées avec des QCM-RU où les répondants pouvaient fournir une certitude estimée concernant leur réponse. Par exemple, Khan et al. [3] ont développé un questionnaire en ligne dans lequel les étudiants répondent à une question par vrai ou faux et rapportent un niveau de certitude dans leur réponse d'après une échelle proposée par les auteurs. Après avoir soumis leurs réponses, les étudiants peuvent demander un retour qui leur indique si les affirmations étaient vraies ou fausses. L'objectif de ce questionnaire est d'aider les étudiants à auto-évaluer leurs connaissances et les enseignants à adapter leurs cours en fonction des résultats des étudiants. Dans un but similaire, Farrell [4] utilise un questionnaire où le répondant pouvait fournir une valeur numérique de certitude. Bien que ces méthodes soient intéressantes les réponses des étudiants restent très restreintes puisqu'elles se cantonnent à un choix booléen. Les QCM-RU ne laissent pas assez de possibilité au répondant pour s'exprimer contrairement aux QCM-RM. C'est pourquoi Diaz et al. [5] proposent dans leur approche un QCM-RM permettant aux étudiants d'exprimer une réponse plus flexible. Pour ce faire, pour chaque question, l'étudiant peut sélectionner une ou plusieurs réponses, et, pour chaque réponse sélectionnée, associer une valeur numérique correspondant à sa confiance dans la réponse choisie. Les réponses sont modélisées par les auteurs grâce à la théorie des fonctions de croyance [6]. Les réponses imparfaites permettent d'identifier les questions qui posent problème aux étudiants. Dans cet article, les QCM-RM utilisés pour les tests n'ont qu'une seule réponse correcte mais il est également possible de travailler avec des QCM-RM qui incluent un ensemble de choix comme bonne réponse attendue. C'est ce que font Diaz et al. [7] qui considèrent des QCM-RM similaires à ceux mentionnés précédemment mais présentent cette fois la possibilité que la bonne réponse ne soit pas un singleton mais potentiellement un ensemble de propositions. Les auteurs proposent toujours à l'étudiant de renseigner sa confiance en sa réponse. Ils lui accordent également la possibilité d'exprimer son ignorance ce qui est ici un concept intéressant. L'étudiant reconnaît ainsi son ignorance partielle ou non sur la réponse à la question. Les auteurs utilisent des fonctions de croyance conditionnelles afin de déterminer si un étudiant a bien acquis les connaissances ou non. Diaz et al. mentionnent également que la difficulté de la question est à prendre en compte dans l'agrégation globale des réponses de l'étudiant car la difficulté impacte le répondant dans son choix.

Nous constatons que pour offrir la possibilité à l'étudiant de rendre compte au mieux de ses connaissances les QCM-RM sont plus pertinents que les QCM-RU. L'intérêt des approches présentées ici est qu'elles permettent non seulement à l'enseignant une estimation des acquis de l'étudiant, mais également une auto-évaluation des connaissances du côté de l'apprenant. En revanche, les méthodes proposées ne s'appliquent pas dans un contexte d'apprentissage dynamique qui induit une évaluation en temps réel de l'apprenant. En effet, le retour à l'apprenant est généralement fait après évaluation de ce dernier. Des QCM d'apprentissage avec des retours après les questions seraient intéressants dans un contexte de cours en ligne pour faciliter l'apprentissage à l'étudiant. De plus, les modèles se contentent de rendre compte si une notion est acquise ou non-acquise par un étudiant, ceci manque de granularité, un concept peut être partiellement appris.

Il est essentiel de parvenir à définir une approche de contrôle des connaissances qui permette un retour qui rende mieux compte de l'apprentissage réel de l'individu inscrit aux cours en ligne. Pour ce faire, nous proposons dans cette thèse un système où l'apprenant répond à un questionnaire d'évaluation avec une possibilité d'être imprécis en cas d'hésitation tout en donnant sa certitude. L'objectif est de parvenir à estimer la connaissance du contributeur grâce à ses réponses imparfaites, mais aussi d'interagir avec lui afin de dynamiser cet apprentissage en lui donnant par exemple un retour sur sa réponse. Il faudrait ainsi estimer l'apprentissage de façon dynamique afin de déterminer le moment où l'apprenant a acquis les connaissances promises par le cours. Il s'agit également d'identifier son degré d'acquisition des connaissances pour l'optimisation du retour qui lui sera fait.

La théorie des fonctions de croyance pourra être utilisée dans cette thèse pour la modélisation des réponses des apprenant et l'estimation de leur connaissance. Il sera notamment intéressant de comparer l'utilisation de cette approche à d'autres comme par exemple les ensembles flous [8]. En effet, Biswas [9] et Chen et al. [10] utilisent des approches qui se fondent sur la théorie des ensembles flous, néanmoins bien qu'elles aident le professeur dans son évaluation de l'étudiant elles ne permettent pas à ce dernier de s'auto-évaluer.

Bibliographie

[1] (28/10/2022) <https://www.fun-mooc.fr/fr/a-propos/>

[2] (28/10/2022) <http://blog.enssat.fr/2022/10/mooc-cybersecurite-comment-securiser-un.html>

[3] K. Khan, D. Davies, J. Gupta. Formative self-assessment using multiple true-false questions on the internet: feedback according to confidence about correct knowledge. *Med Teach*. 23 (2), 158–163 (2001).

[4] G. Farrell. A comparison of an innovative web-based assessment tool utilizing confidence measurement to the traditional multiple choice, short answer and problem solving questions." (2006).

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES
 Campus de Beaulieu, Bât IRISA/Inria 12 F • 263 Avenue du Général Leclerc CS74205
 35042 Rennes Cedex, France • Tél. : +33 (0)2 99 84 71 00 • www.irisa.fr

- [5] J. Diaz, M. Rifqi, B. Bouchon-Meunier, S. Jhean-Larose, and G. Denhière. Imperfect answers in multiple choice questionnaires. In European Conference on Technology Enhanced Learning, pages 144–154. Springer, 2008.
- [6] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. The Annals of Mathematical Statistics, 38:325–339, 1967.
- [7] J. Diaz, M. Rifqi, et B. Bouchon-Meunier. Qcm évidentiels pour le diagnostic des apprenants. Actes des rencontres francophones sur la Logique Floue et ses Applications (LFA2007), 2007.
- [8] L.A. Zadeh. Fuzzy sets, Inform. and Control 8 (1965) 338-353
- [9] R. Biswas. An application of fuzzy sets in students' evaluation. Fuzzy Sets & Systems, 74 : 187-194, 1995
- [10] S. M. Chen, C. H. Lee. New methods for students' evaluation using fuzzy sets. Fuzzy Sets & Systems, 104 : 209-218, 1999.