



Suite d'outils pour QCM++ automatiques

Contexte

Les questionnaires à choix multiples (QCM) font partie des outils d'enquête utilisés dans l'enseignement pour évaluer les apprentissages. Une caractéristique intéressante de ces outils est qu'ils permettent d'automatiser l'évaluation des questionnaires rendus, et cette automatisation peut offrir de nombreux avantages.

- Évaluer extrêmement rapidement des copies. Ce facteur est non seulement très important lorsque le nombre d'apprenants est grand (*e.g.*, en licence ou sur un MOOC), mais il permet également de réduire considérablement le temps entre le passage d'un test et le résultat de ce test, ce qui permet aux apprenants d'avoir un retour très rapidement.
- Améliorer l'impartialité lors de l'attribution de notes aux copies. En particulier, si l'algorithme qui associe une note à chaque copie est déterministe alors il ne sera pas biaisé par l'ordre de parcours des copies (alors qu'un correcteur humain est soumis à la fatigue et aura tendance à ajuster son évaluation à partir des copies précédemment évaluées). Si de plus cet algorithme ne se base que sur les réponses fournies aux questions sans les relier à des informations personnelles sur l'apprenant ou à d'autres évaluations (et que les questionnaires sont *bien* conçus), l'algorithme permet de donner des notes sans a priori personnalisé à chaque apprenant ce qui peut éviter de très nombreux biais humains (*e.g.*, évaluer différemment la copie d'un *bon* ou d'un *mauvais* apprenant, être soumis à des biais sexistes/racistes/etc. implicites ou explicites).
- Permettre une analyse fine de la compréhension de savoirs et de la maîtrise de savoir-faire. Selon le type de questionnaire utilisé, il est possible d'associer à chaque choix des savoirs/savoir-faire/compétences, et il est possible d'agréger et d'analyser (pseudo-)automatiquement ces données pour réaliser un bilan de la compréhension de certaines notions, que ce soit sur l'ensemble des apprenants, par groupe ou de manière individualisée. Ces bilans peuvent être utilisés par une équipe pédagogique pour orienter au mieux la suite d'un cours (*e.g.*, accélérer le rythme car les notions sont maîtrisées, revoir une notion qui n'a pas été transmise, faire un sujet de TD/TP plutôt qu'un autre car une notion est jugée plus prioritaire qu'une autre vu la compréhension actuelle de la population d'apprenants), mais aussi au niveau individuel pour détecter et signaler aux apprenants les notions au niveau de maîtrise insuffisant.

Des outils libres existent déjà pour sonder et évaluer automatiquement des apprenants, on peut en particulier citer **AMC** pour la création de QCM *papier*, et les plateformes pédagogiques comme **Moodle** ont une sous-partie dédiée aux formulaires d'évaluation. Ces outils ne permettent cependant pas d'exploiter pleinement les avantages offerts par cette automatisation actuellement, en particulier dans leur analyse des données. Voici les propriétés additionnelles que l'on aimerait voir dans ces outils.

- Permettre la définition et la composition de questions dans différents questionnaires à but d'évaluation sur papier et/ou plateformes numériques (*e.g.*, Moodle), mais aussi comme outil d'apprentissage et d'autoévaluation pour les apprenants. En particulier, permettre la génération de questions et de réponses automatiquement serait très intéressant pour correspondre aux particularités de la licence flexible à l'UT3 (plusieurs sujets d'examen similaires à concevoir chaque semestre).

- Permettre l'utilisation facile des différents formats de questionnaires automatiques recommandés par la littérature en pédagogie. En particulier, l'utilisation de questionnaires de types Multiple-True-False [F92] [BC19] [HDR10] et de degrés de certitude [P15] [L17] n'est pas satisfaisante.

Objectifs du projet

L'objectif général de ce projet est de fournir une solution logicielle pour répondre aux besoins listés en section précédente. Ce projet est naissant, le travail de cette année se concentre donc sur les phases initiales du développement : clarification des besoins, veille technologique, réalisation de prototypes et conception de l'architecture générale.

Les objectifs de ce projet sont nombreux et les objectifs principaux sont listés ci-dessous. **Il n'est pas attendu que tous ces objectifs soient réalisés par des étudiant·e·s, et encore moins qu'un seul groupe d'étudiant·e·s réalise tous ces objectifs !** Selon le profil des candidat·e·s et leur nombre, des objectifs seront définis pour chaque groupe travaillant sur le projet.

- Identifier les types de QCM, voire de questionnaires automatisables au sens large, que l'on souhaite rendre disponible dans la solution logicielle créée. Cela se fera en réalisant un (petit) état de l'art scientifique de la littérature en pédagogie et en sciences humaines (psychologie cognitive...) autour du sujet, mais également en listant les types de questionnaires disponibles dans les différents outils (libres ou non) existants autour du problème.
- Réaliser une veille technologique des outils libres existants de création et d'évaluation de QCM, en commençant par **AMC** et **Moodle**. Quels types de QCM permettent de faire chaque outil ? Quel format de données utilisent-ils ? Ces formats de données sont-ils compatibles ? Comment rendre compatibles les formats de données si ce n'est pas le cas ? Quelle sous-partie du travail réalisé dans les outils existant est-il possible de réutiliser pour répondre à nos besoins ? En analysant le design souhaité et celui des outils existants, ainsi que la dette technique des outils existants, vaut-il mieux faire nos développements au sein de l'existant ou vaut-il mieux s'investir dans un nouveau projet ?
- Si nécessaire, définition d'un format de données descriptif pour représenter une question et ses réponses associées. Il sera important de permettre d'associer des métadonnées aux réponses, afin de permettre plus tard une analyse des résultats.
- Si nécessaire, définition d'un format de données descriptif pour représenter une feuille à compléter, dans le but de modulariser le scan automatique de copies *papier*. Ce format contiendrait notamment l'emplacement des cases à cocher et des champs à remplir sur chaque page, et la question/réponse correspondant à ces cases/champs.
- Conception de l'architecture logicielle de la solution logicielle complète (composants, pipelines de données, interface graphique)...
- Implémentation de prototypes pour les différents composants identifiés. Possibilité de développement de nouveaux composants *ad hoc*, mais aussi de modularisation et d'amélioration d'outils existants.

Bibliographie

- [F92] Frisbie, D. A. (1992). The multiple true-false item format: A status review. *Educational Measurement: Issues and Practice*.
- [BC19] Brassil, C. E., & Couch, B. A. (2019). Multiple-true-false questions reveal more thoroughly the complexity of student thinking than multiple-choice questions: a Bayesian item response model comparison. *International Journal of STEM Education*, 6(1), 1-17.
- [HDR10] Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education*, 15(3), 309-333.
- [P15] Prosperi, O. (2015). Le réalisme avec degrés de certitude. *Mesure et évaluation en éducation*, 38(1), 121-140.
- [L17] Leclercq, D. (2017). Une meta-analyse des degrés de certitude exprimés en mots. *Evaluer: Journal International de Recherche en Education et Formation*, 2(3).