



# Rétroactions dans un environnement numérique d'apprentissage : modèle de description et décision

► **Sébastien JOLIVET** (IUFÉ, UNIGE et LDAR, Université de Paris), **Amel YESSAD**, **Mathieu MURATET** (LIP6, Sorbonne Université), **Elann LESNES** (LDAR, Université Rouen-Normandie), **Brigitte GRUGEON-ALLYS** (LDAR, Université Paris-Est Créteil), **Vanda LUENGO** (LIP6, Sorbonne Université)

---

---

■ **RÉSUMÉ** • L'article introduit un modèle de description des rétroactions épistémiques et un modèle informatique de décision de ces rétroactions. Le modèle de description est fondé sur des connaissances didactiques et a pour objectif d'être suffisamment explicite pour aider à la décision automatique des rétroactions. Le modèle informatique de décision combine des connaissances expertes et un algorithme d'apprentissage par renforcement. La faisabilité de l'approche est évaluée avec la réification du modèle et son intégration à une implémentation du modèle décisionnel.

■ **MOTS-CLÉS** • modèle de rétroaction épistémique, modèle informatique de décision de rétroaction, apprentissage par renforcement

■ **ABSTRACT** • *This paper introduces a model for describing epistemic feedbacks and a computer model for deciding these feedbacks. The description model is based on didactic knowledge and aims to be sufficiently explicit to allow the automatic decision-making of feedbacks. The computational decision model combines expert knowledge and a RL algorithm. Initial evaluations of the description model were carried out to verify its descriptive capacity. In addition, the reification of the model and its integration into an implementation of the decision-making model made it possible to show the feasibility of the approach.*

■ **KEYWORDS** • *epistemic feedback model, computer model of feedback decision, reinforcement learning*

## **1. Introduction**

Les rétroactions jouent un rôle important tout au long de l'apprentissage (Hattie et Timperley, 2007), (Shute, 2008). Du point de vue des environnements informatiques pour l'apprentissage humain (EIAH) ces rétroactions peuvent prendre différentes formes (réponse correcte, rétroaction motivationnelle, vérification, explication de l'erreur, etc.) et intervenir à différents moments de l'interaction. Par ailleurs, plusieurs travaux cherchent à classer les rétroactions (produites par le système ou l'enseignant) à des fins explicatives ou d'aide pour l'enseignant. Cependant, ces différents travaux ne permettent pas de décrire le contenu de la rétroaction du point de vue des connaissances en jeu d'apprentissage. Enfin, il n'y a pas de véritable consensus sur les choix des rétroactions pour avoir un effet sur l'apprentissage.

Par ailleurs, les techniques informatiques pour produire des rétroactions adaptées sont variées : production des règles, réseaux bayésiens, traitement automatique de la langue, etc. (Bimba *et al.*, 2017). Ces méthodes symboliques basées sur l'expertise humaine sont complexes à mettre en place et difficiles à faire évoluer dynamiquement, au fil des interactions avec les apprenants. Elles peuvent être combinées avec des méthodes numériques pour permettre une meilleure évolution dans le temps et une meilleure adaptation aux profils des apprenants.

Ainsi, l'objet de cet article est de présenter un modèle de description des rétroactions épistémiques et un système de décision de ces rétroactions pour des environnements d'apprentissage numériques. Nous débutons en section 2 par un état de l'art qui nous permet de mettre en évidence différents verrous scientifiques et solutions liés aux rétroactions, en particulier : 1) comment décrire les rétroactions épistémiques en se fondant sur des connaissances didactiques et 2) comment modéliser un système décisionnel calculable s'appuyant à la fois sur cette description, sur des hypothèses didactiques relatives à la rétroaction et sur des connaissances issues des données d'interaction entre l'apprenant et le système.

Nous présentons en section 3 notre modèle de description fondé sur des connaissances didactiques qui sont également introduites. Puis nous abordons en section 4 la question de la décision des rétroactions, en tant que problème de décision dans l'incertain.

Dans la section 5, nous présentons les premières évaluations du modèle de description ainsi qu'une première preuve de concept, ou démonstration de faisabilité, dans l'étude du cas *Mindmath*, qui permet de mettre en relation le modèle de description avec le modèle de décision.

La conclusion, en section 6, nous permet alors d'aborder les limites et les perspectives ouvertes par ce travail, en particulier l'adéquation du modèle de description proposé par rapport aux questions initiales et les difficultés soulevées par la production effective des rétroactions.

## **2. État de l'art et verrous scientifiques**

De façon générale, la rétroaction peut être définie comme une intervention d'un agent extérieur afin de fournir de l'information concernant des aspects liés à l'exécution d'une tâche (Kluger et DeNisi, 1996, p. 255). Dans cet article nous nous intéressons aux rétroactions dans des situations d'apprentissage où l'apprenant interagit avec un environnement informatique. Nous définissons la rétroaction comme une intervention faite par un système informatique suite à l'exécution d'une tâche d'apprentissage.

### **2.1. Éléments généraux sur les rétroactions**

Une des revues de la littérature sur les rétroactions en situation d'apprentissage (Hattie et Gan, 2011) montre que différentes perspectives psychologiques fournissent des cadres distincts pour décrire différentes visions de l'apprentissage, ainsi que la nature, les caractéristiques et la fonction de la rétroaction.

Tel qu'indiqué par Mory (Mory, 2004, p. 745), les rétroactions peuvent faire référence non seulement à l'exactitude ou non d'une réponse, mais aussi à d'autres informations, telles que la précision de la réponse, l'accompagnement dans la tâche, des conseils plus généraux ou des messages liés à la motivation.

L'effet de ces rétroactions sur l'apprenant est un sujet sur lequel il n'y a pas de consensus.

Cependant, Wisniewski et ses collègues (Wisniewski *et al.*, 2020) signalent dans une revisite de l'article « *The power of feedback* » (Hattie et Timperley, 2007), que les rétroactions ont un impact plus important sur l'apprentissage que sur la motivation et le comportement. Stobart (Stobart, 2018, p. 46) pointe le fait que l'impact de la rétroaction n'est pas le même selon le niveau, en termes de compétences et d'expérience, de

l'apprenant. Par ailleurs, des travaux tels que ceux de Brooks *et al.* (Brooks *et al.*, 2019) et Small et Lin (Small et Lin, 2018) montrent l'importance de la prise en compte de l'erreur et les diverses formes de cette prise en compte.

## **2.2. Les rétroactions épistémiques**

Nous avons choisi de nous intéresser uniquement aux rétroactions épistémiques, définies comme relatives à la connaissance en jeu dans l'activité (Luengo, 2009, p. 14), ces rétroactions étant choisies et produites par le système informatique. C'est donc pour nous une problématique à la croisée de l'informatique et de la didactique comme l'a formulé Luengo (Luengo 2009, p. 1) : « pour un contenu spécifique, ayant analysé l'activité de l'apprenant, choisir la rétroaction à produire pour faire évoluer cette connaissance de façon optimale (choix du moment, de la nature, de la modalité) ».

## **2.3. Le choix du moment de la rétroaction**

En ce qui concerne le choix du moment, Van Lehn (Van Lehn, 2006) propose de distinguer les rétroactions selon deux types de boucles : boucle interne (*inner loop feedback*) et boucle externe (*outer loop feedback*). La rétroaction en boucle interne est produite pendant la résolution d'une tâche particulière, alors que la rétroaction en boucle externe est produite en ciblant une tâche finale ou un ensemble de tâches. La rétroaction de type boucle interne fournit généralement des informations sur l'exactitude d'une solution (partielle ou non), combinées à des conseils sur la façon de corriger les erreurs et/ou la manière de procéder pour résoudre la tâche en cours. Les tuteurs intelligents sont un exemple classique car ils proposent des rétroactions à chaque étape de résolution. La rétroaction en boucle externe utilise l'état actuel des connaissances de l'apprenant, dans un domaine donné, pour sélectionner les tâches ou activités à proposer pour la suite (VanLehn, 2011). Des effets positifs des rétroactions tant dans des boucles internes (VanLehn, 2011) que dans des boucles externes (Bull et Kay, 2016) ont été mesurés et ces deux types de rétroactions sont implémentés dans des environnements d'apprentissage. Dans notre cas, nous nous intéressons aux rétroactions en boucle interne.

## **2.4. La nature de la rétroaction**

La nature de la rétroaction est relative à l'information qu'elle contient. Ici, nous nous intéressons en particulier aux rétroactions formatives, car elles ciblent le processus d'apprentissage (Shute, 2008, p. 154). Pour ce type

de rétroactions plusieurs catégories ont été proposées. Leibold et Schwarz (Leibold et Schwarz, 2015) signalent quatre types de rétroactions selon la nature des informations qu'elles contiennent (feedback correctif - feedback épistémique - feedback suggestif - feedback épistémique et suggestif). Shute (Shute, 2008, p. 160), synthétisant de nombreux travaux, propose de multiples critères qui l'amène à une typologie des rétroactions ordonnées par complexité croissante (*No feedback, Verification, Correct response, Try again, Error-flagging, Elaborated, Attribut isolation, Topic Contingent, Response continent, Hints/cues/prompts, Bugs/misconceptions, Informative Tutoring*). J. Murray et al. (Murray et al., 2018, p. 86) proposent une taxonomie des rétroactions dans le cadre des rétroactions écrites sur des productions papier d'élèves, afin de pouvoir caractériser leur nature (*topic-specific feedback ; corrective feedback ; generic feedback, simple feedback about something correct, simple feedback about something incorrect, complex feedback about something correct, complex feedback about something incorrect, feed-forward feedback, connective feedback, dialogical feedback, personal feedback, positive feedback, negative feedback, nondescript feedback, discours feedback*).

Ces distinctions sont intéressantes, mais posent le problème de leur caractère opérationnel lors de la conception. Par exemple, une rétroaction de type « *topic specific feedback* » pourrait bien être également une rétroaction de type « *corrective feedback* ». De plus, ces catégories sont souvent produites à posteriori et ne sont pas suffisamment formelles pour les rendre calculables, à partir des connaissances expertes, lors du choix de la rétroaction.

## **2.5. La modalité de la rétroaction**

Enfin, la modalité de la rétroaction fait référence au registre auquel elle fait appel. Celui-ci peut être décrit du point de vue informatique au sens du type de sortie (son, image, écrit), mais peut aussi faire appel à la notion de registre de représentation sémiotique (Duval, 1993) telle qu'utilisée en didactique. Une rétroaction peut, par exemple, mobiliser le registre langue naturelle, le registre algébrique ou le registre figural. Les travaux proposant des taxonomies de rétroactions ne font pas référence à cette dimension.

## **2.6. Didactique des mathématiques et rétroactions**

Les recherches dans le champ de la didactique des mathématiques s'intéressant à la question des rétroactions épistémiques fournies par un environnement informatique sont peu nombreuses.

D'une part, certains environnements informatiques fréquemment utilisés pour l'apprentissage des mathématiques n'incluent pas de rétroactions explicites. En particulier, dans les logiciels de géométrie dynamique, ce sont les règles de la géométrie euclidienne régissant la construction et le déplacement qui fournissent, *de facto*, des rétroactions, par exemple lors de la déformation d'une figure.

D'autre part, lorsque les environnements informatiques incluent des rétroactions épistémiques explicites, leur conception n'est pas toujours interrogée d'un point de vue didactique, comme le souligne Rezat qui reprend (Dawson *et al.*, 2018) : « *in most cases, there is no clear indication of how the feedback inputs (e.g., comments on the assessment performance) are designed to impact on subsequent assessment or how the impact is to be measured* » (Rezat, 2021, p. 1434). Cette difficulté est aussi identifiée par McKendree : « *the psychological theories that discuss the issue have offered no concrete recommendations about the form or content of feedback in order to be maximally effective. [...] Further, the systems that have been implemented or suggested have offered little evidence that the techniques used actually help the learning process.* » (McKendree, 1990, p. 382).

En géométrie comme en algèbre, différents travaux montrent que la conception des rétroactions en lien avec le diagnostic du raisonnement mathématique mené par l'élève est complexe, y compris dans le cadre d'un travail mathématique relativement codifié, voire algorithmique (raisonnement déductif en géométrie, calcul sur les expressions algébriques ou résolution d'équations du premier ou du second degrés). Ainsi, en géométrie, Tessier-Baillargeon, Leduc, Richard et Gagnon comparent onze « systèmes d'aide à la démonstration » (Tessier-Baillargeon *et al.*, 2017). Leur analyse porte notamment sur la composante tutorielle des systèmes et les rétroactions proposées aux élèves et pointe l'importance de l'analyse de l'activité de l'apprenant. Ils notent en particulier que l'aide « à la prochaine étape [...] qui implique que le système tutoriel puisse identifier la solution travaillée par l'élève pour le guider vers la prochaine action à poser » (Tessier-Baillargeon *et al.*, 2017, p. 113) n'est pas majoritaire puisqu'elle concerne quatre des systèmes étudiés

sur les onze. De plus, trois de ces quatre systèmes imposent une structure assez rigide au raisonnement de l'élève.

En algèbre, notamment dans le logiciel APLUSIX qui permet de travailler le calcul sur les expressions algébriques et la résolution d'équations, Bouhineau et Nicaud définissent trois types de rétroactions épistémiques implémentées dans l'environnement (Bouhineau et Nicaud, 2006):

- des *rétroactions syntaxiques*, qui répondent à « est-ce que l'expression algébrique est syntaxiquement correcte et bien définie ? » ;
- des *rétroactions sémantiques*, qui répondent à « est-ce que le raisonnement poursuivi est sémantiquement correct ? » ;
- des *rétroactions stratégiques*, qui sont relatives à l'avancement de la résolution.

Cette dernière catégorie correspond à des jauges affichées qui se remplissent au fur et à mesure de l'avancée dans la résolution. Ce type de rétroactions, plus global, nécessite une structure assez rigide du raisonnement mobilisé pour la résolution des problèmes.

Au-delà de la conception des rétroactions se posent la question de l'appropriation de ces rétroactions par les élèves et celle de leurs éventuels effets sur les apprentissages mathématiques. À ce propos, Rezat souligne : « *while there is a large body of mainly quantitative research on the effectiveness of feedback in general, very little is known about how feedback actually affects students' individual content specific learning processes and conceptual development* » (Rezat, 2021, p. 1433).

Dans une étude récente, il étudie l'impact des rétroactions sur les conceptions d'élèves résolvant une tâche de probabilité proposée dans un manuel numérique de mathématiques. Le manuel propose plusieurs types de rétroactions décrites à partir de la classification de Shute (Shute, 2008) présentée dans la section 2.4: lorsque l'élève appuie sur le bouton permettant de valider sa réponse, il reçoit une rétroaction de vérification (*verification feedback*) et, si la réponse est erronée, la possibilité de recommencer deux fois (*try again feedback*). Si l'élève échoue après le deuxième essai, une ampoule apparaît sur le côté, celle-ci donne plusieurs indices (*elaborated feedback hints/cues/prompts*). Si l'élève échoue une troisième fois, la bonne réponse lui est donnée (*correct feedback*). Rezat montre alors que dans les deux cas étudiés, la rétroaction de vérification incite les élèves à repenser leurs réponses. Cependant, puisque cette rétroaction n'est accompagnée d'aucune information, les élèves se

**Sébastien JOLIVET, Amel YESSAD, Mathieu MURATET, Elann LESNES, Brigitte GRUGEON-ALLYS, Vanda LUENGO**

contentent d'ajuster leur réponse sans faire évoluer leur procédure ou stratégie de résolution. Après avoir reçu la rétroaction élaborée, les élèves sont en mesure de proposer la réponse correcte. Cependant, cette réponse correcte n'apparaît pas à la suite d'une procédure de résolution correcte et attendue à ce niveau scolaire. En effet, « *the analysis revealed that in both cases students instrumentalize information that is not relevant for the task* » (Rezat, 2021, p. 1442). On observe par exemple une élève s'appuyer sur les formulations des deux rétroactions de vérification qu'elle reçoit (« *No, this is not absolutely correct* » et « *No, not yet correct* ») pour déduire qu'il faut proposer une réponse proche de celle qui l'a menée à recevoir la rétroaction « *No, this is not absolutely correct* ».

Nous constatons des observations comparables dans les analyses de Cazes et Vandebrouck (Cazes et Vandebrouck, 2008) qui étudient cinq bases d'exercices en ligne en s'intéressant aux effets sur des apprenants de lycée et du supérieur. Les bases d'exercices mettent en œuvre différents types de rétroactions : elles donnent ou non la bonne réponse, font référence ou non à la réponse erronée de l'élève, explicitent ou non la technique attendue, etc. On peut noter qu'aucune de ces bases n'exploite la réponse de l'élève pour le guider vers la réponse attendue. Cazes et Vandebrouck remarquent que les rétroactions ne sont souvent pas adaptées à l'activité de l'élève. Elles ne permettent à l'élève ni de comprendre son erreur, ni de proposer une meilleure solution. Soit la rétroaction ne reprend pas la procédure utilisée par l'élève pour lui expliquer pourquoi elle ne fonctionne pas, soit elle lui propose une procédure plus complexe (au moins en apparence) et l'élève ne parvient pas à s'en emparer. Cazes et Vandebrouck concluent en disant que lorsque la rétroaction est adaptée, si l'élève l'interprète correctement et s'il est dans une logique d'apprentissage (c'est-à-dire qu'il cherche « un peu plus que l'obtention du résultat »), alors la rétroaction peut jouer le rôle d'une aide constructive (Cazes et Vandebrouck, 2008, p. 186) et permettre à l'élève de trouver le résultat de l'exercice, mais aussi de transférer une partie des connaissances en jeu à d'autres exercices.

Rezat résume ainsi la situation : « *no matter how carefully the tasks in e-textbooks and the feedback messages are designed, it is very likely that they still contain irrelevant information, which might become salient in students' solution processes. Therefore, tasks, feedback messages, and diagrams need to be designed very carefully and as unambiguously as possible* » (Rezat, 2021, p. 1443).

Si ces recherches permettent à nouveau de catégoriser les rétroactions, elles ne permettent pas de dire comment construire des rétroactions adaptées à l'activité de l'élève.

## **2.7. Modélisation formelle et calculable des rétroactions épistémiques**

Nous avons ainsi l'objectif de proposer une formalisation permettant, d'une part, de pouvoir choisir automatiquement une rétroaction en tenant compte de ces différentes dimensions et, d'autre part, de pouvoir analyser de façon plus fine les rétroactions proposées par différents environnements informatiques d'apprentissage ainsi que leurs effets.

Tel qu'introduit précédemment, les connaissances didactiques semblent nécessaires pour concevoir des rétroactions épistémiques. Cependant, le choix de la rétroaction la plus pertinente pour un apprenant donné à un moment donné reste un verrou. De plus, l'évolution des systèmes prenant en compte tous ces éléments est difficile ou fastidieuse à mettre en place (Luengo, 2009).

Il semble donc nécessaire de concevoir un modèle capable de produire des rétroactions à partir des connaissances expertes, mais également capable d'évoluer, de façon informée, au fur et à mesure des interactions avec les apprenants, c'est-à-dire en utilisant les traces de ces interactions pour informer le modèle de décision et adapter l'importance de chaque rétroaction selon les contextes (c.-à-d. les situations d'apprentissage), les profils des élèves ou les modalités disponibles.

Une approche permettant de représenter le processus de décision de la rétroaction (Murray *et al.*, 2004) ainsi que le type de rétroaction, de façon explicite et découplée, peut permettre ces deux types d'évolution. Nous pouvons ainsi distinguer deux composantes : l'une relative à la modélisation de la rétroaction elle-même et l'autre au processus de décision permettant de choisir la ou les rétroactions les plus pertinentes. Pour cela il est nécessaire de caractériser de façon formelle les rétroactions et d'identifier par ailleurs les facteurs qui interviennent dans le choix de la rétroaction, en tenant compte de ces caractéristiques.

Plusieurs travaux se sont intéressés à la dimension décisionnelle. Mayo et Metrovic (Mayo et Metrovic, 2001) ont proposé une classification des rétroactions qui est relative à l'optimisation de la rétroaction - qu'ils qualifient de « *pedagogical action* » (Mayo et Metrovic, 2001, p. 131). Selon ces auteurs, étant donné un modèle de l'apprenant représenté sous forme

**Sébastien JOLIVET, Amel YESSAD, Mathieu MURATET, Elann LESNES, Brigitte GRUGEON-ALLYS, Vanda LUENGO**

de réseaux bayésiens, le principe est de calculer l'action pédagogique optimale. Ils distinguent ainsi trois approches d'optimisation : alternative, diagnostic et théorie de la décision (*ibid.*).

Dans tous ces cas, les rétroactions elles-mêmes sont produites de façon *ad hoc*. Certains travaux produisent les rétroactions automatiquement après le processus de décision, mais cela particulièrement dans le cas des boucles externes. Nous pouvons ainsi citer l'approche développée par Luengo *et al.* (Luengo *et al.*, 2011), qui proposent une procédure en quatre étapes : choix de la cible de la rétroaction (quelle connaissance ou erreur doit être ciblée suite au diagnostic), choix de l'intention de la rétroaction (déstabiliser, renforcer, diagnostiquer), choix de la modalité et enfin choix du contenu. Pour le contenu de la rétroaction, l'analyse et le calcul sont faits en fonction de la forme du support choisie (cours en ligne, simulateur ou base des cas cliniques).

Une autre approche, moins informée par les connaissances expertes, est celle de Clement *et al.* (Clement *et al.*, 2015) qui proposent des rétroactions de boucle longue de type « prochaine activité » en adaptant un algorithme d'apprentissage par renforcement bien connu appelé *Multi Armed Bandit* (MAB). Les algorithmes de type MAB mettent en œuvre un processus de décision markovien à un seul état ; l'objectif est de chercher (*i.e.* explorer) l'action qui produit la meilleure récompense immédiate, puis de continuer à sélectionner (*i.e.* exploiter) cette dernière tant que c'est le cas. Les premières implémentations d'approches basées sur des MAB pour le choix de nouveaux exercices, proposées par Clement *et al.* (Clement *et al.*, 2015), ont démontré l'efficacité de ces approches pour la sélection de nouveaux scénarios, étant donné des degrés plus ou moins importants de connaissances expertes à disposition. Plus récemment, Frenoy *et al.* (Frenoy *et al.*, 2016) ont utilisé un MAB basé sur une approche probabiliste *softmax* pour la sélection d'actions pédagogiques lors de l'apprentissage de la calligraphie dans un environnement virtuel informé, c'est-à-dire un environnement disposant des informations sur la tâche de l'utilisateur et son contexte. Le point fort d'une approche par renforcement tient au fait de pouvoir combiner les connaissances expertes à des données collectées pendant l'interaction entre l'apprenant et le système.

## **2.8. Problématique**

Cet état de l'art permet de pointer trois éléments essentiels :

- la diversité des classifications des rétroactions et le côté peu opérant de ces classifications pour produire et décider des rétroactions adaptées à l'activité de l'apprenant ;

- les limites des connaissances issues des travaux de didactique permettant de répondre à la question « quelle est la rétroaction la plus pertinente ? » ;

- la diversité des approches informatiques relatives à la décision des rétroactions.

Suite à ces constats, nous faisons les hypothèses suivantes :

- il est possible de proposer un modèle formalisé de description des rétroactions ;

- ce modèle peut permettre d'intégrer les connaissances de la didactique sur les savoirs en jeu ;

- ce modèle est une base permettant d'aborder la décision des rétroactions épistémiques en boucle interne à l'aide d'algorithmes d'apprentissage par renforcement.

Une première étape, qui est l'objet principal du travail présenté, est de formaliser suffisamment les rétroactions à partir des connaissances expertes, tout en s'assurant que l'on puisse exploiter cette formalisation, notamment pour rendre calculable le choix de la rétroaction.

### **3. Modèle de rétroactions**

Nous présentons maintenant le modèle de rétroactions. Celui-ci permet de caractériser didactiquement une rétroaction dans la mesure où il permet d'explicitier à son propos :

- les éléments du savoir présents,
- les précisions apportées sur ces éléments de savoir,
- la présence de contenus relatifs à une erreur liée au savoir,
- sa ou ses fonctions possibles,
- sa relation avec la tâche donnée dans l'environnement.

La description des éléments de savoir mobilisés dans les rétroactions, des précisions apportées dessus et des classes d'erreurs usuelles nécessitent de disposer d'une modélisation du savoir. Nous exploitons l'approche praxéologique telle que définie dans la théorie anthropologique du didactique, ou TAD (Chevallard, 1992), et son extension T4-TEL (Chaachoua *et al.*, 2019). Nous présentons dans la sous-section suivante les éléments de ces modèles nécessaires à la compréhension de la suite de l'article. Pour une approche plus globale, le lecteur trouvera des

utilisations détaillées de ces cadres théoriques, pour des travaux en EIAH, dans (Mandin et Guin, 2014), (Vu et Tchounikine, 2020) et (Jolivet *et al.*, 2021a).

### **3.1. Modélisation du savoir**

L'approche praxéologique consiste à représenter toute activité humaine à l'aide d'un quadruplet [Type de tâches, Technique, Technologie, Théorie]. Pour préciser et illustrer les éléments de ce quadruplet nous nous plaçons dans le domaine des mathématiques.

- *Type de tâches*: les tâches à réaliser sont regroupées dans des ensembles de tâches, un ensemble contenant des tâches pouvant toutes être résolues d'au moins une manière commune. Un tel ensemble est appelé *type de tâches*. Les types de tâches sont définis à partir d'un verbe d'action (par exemple « résoudre ») et d'un objet mathématique sur lequel s'exerce l'action (par exemple « une équation du premier degré »).

- *Technique* (renommée *procédure* dans la suite du texte) : il s'agit d'un moyen existant pour réaliser les tâches d'un type de tâches ; par exemple, pour résoudre certaines équations du 1<sup>e</sup> degré : « transformer l'équation en une équation équivalente avec l'inconnue dans un seul membre, transformer l'équation en une équation équivalente avec les termes constants dans l'autre membre, calculer la valeur de l'inconnue ».

- *Technologie et théorie* (renommées *justification* dans la suite du texte) : il s'agit des propriétés, définitions, etc., qui permettent de justifier la validité mathématique des techniques ; par exemple, « une égalité est préservée quand on réalise la même opération dans les deux membres », « le produit d'un nombre et de son inverse est égal à un ».

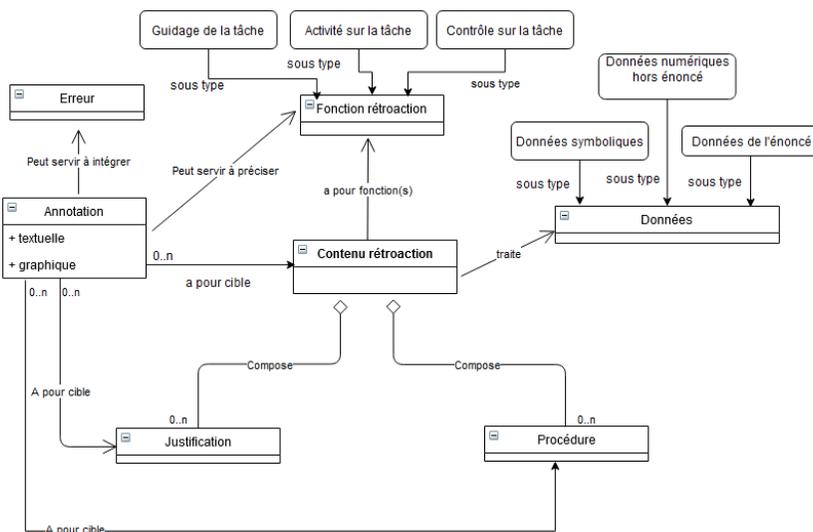
Pour prendre en compte les erreurs, nous nous appuyons sur des classes d'erreurs susceptibles de se manifester lors de la résolution des tâches. Elles sont identifiées à partir de travaux issus de la didactique et sont reconnues par la mobilisation par l'apprenant d'une ou de plusieurs technologies erronées, qui va s'exprimer par la mise en œuvre d'une procédure erronée.

Afin de limiter les confusions, ou difficultés de lecture, liées à la dimension fortement polysémique dans ce contexte interdisciplinaire du vocabulaire utilisé dans la TAD, nous utilisons dans la suite de ce texte le terme *procédure* en remplacement du terme *technique* (c'est-à-dire tout élément permettant de réaliser un ensemble de tâches) et le terme *justification* en remplacement des termes *technologie et théorie* (c'est-à-dire les éléments du savoir qui permettent de justifier la validité d'une

procédure). Ainsi, pour une tâche donnée, il existe une ou plusieurs procédures qui permettent de réaliser cette tâche. Les justifications permettent de garantir la validité des procédures et de guider leur mise en œuvre.

### 3.2. Le modèle de rétroaction

Notre modèle de rétroaction est synthétisé dans le diagramme présenté dans la figure 1. Nous détaillons dans la suite de cette section chacune de ses composantes et explicitons les relations qui les organisent. Nous l'illustrons par quelques exemples qui sont enrichis au fil de l'introduction des différentes composantes.



**Figure 1 • Diagramme du modèle de rétroaction**

#### 3.2.1. Composante « Contenu rétroaction »

La composante « Contenu rétroaction » est au cœur de la représentation de la rétroaction, elle est nécessaire à son existence. Elle peut être composée d'une ou plusieurs procédures de résolution et/ou d'une ou plusieurs justifications.

*Exemples :* pour la tâche « résoudre l'équation  $2y+1=5y-5$  » on peut envisager, par exemple, les contenus proposés dans le tableau 1.

**Tableau 1 • Exemples de contenus d'une rétroaction**

Exemples de procédures	Exemples de justifications
<p><i>Procédure 1</i>  <math>2y+1 = 5y - 5</math>  <math>2y+1-2y = 5y - 5 -2y</math>  <math>1+5 = 3y - 5 + 5</math>  <math>6 = 3y</math>  <math>y = 2</math></p>	<p><i>Justification 1</i>                      On obtient une équation équivalente à l'équation initiale en réalisant la même opération dans les deux membres de l'équation</p>
<p><i>Procédure 2</i>                      2 est une racine évidente de l'équation. Or une équation du 1<sup>er</sup> degré admet zéro, une ou une infinité de solutions. Donc 2 est la solution de l'équation.</p>	<p><i>Justification 2</i>                      On appelle solution d'une équation toute valeur qui, substituée à l'inconnue, rend l'égalité vraie.</p>

*Remarque:* nous proposons ici des contenus possibles, nous ne discutons ni de leur pertinence mathématique, ni de leur formulation, ni de leur mise en forme, ni de leur adaptation à tel ou tel élève.

### **3.2.2. Composante « Annotation »**

La composante « Annotation » correspond à des éléments facultatifs qui viennent enrichir la rétroaction. Une annotation peut concerner (relation *a pour cible* dans la figure 1) le niveau global de la rétroaction ou porter spécifiquement sur un des éléments de contenu. Une annotation peut être textuelle ou graphique et avoir plusieurs finalités : détailler un aspect du contenu de la rétroaction, intégrer des éléments relatifs à une erreur, expliciter une fonction de la rétroaction. Nous illustrons la première finalité dans les exemples qui suivent et revenons sur les deux autres après avoir défini les composantes *erreur* et *fonction de la rétroaction*.

*Exemples:* pour les contenus proposés dans le tableau 1 on peut envisager, par exemple, les annotations proposées dans le tableau 2.

**Tableau 2 · Exemples d'annotations pour une rétroaction**

Exemple d'annotation avec pour cible le niveau global	Exemple d'annotation avec pour cible la procédure 1 du tableau 1.
« Nous te proposons un exemple de résolution d'une équation dont tu peux t'inspirer » <i>qui annote la rétroaction de contenu la procédure 1 ou 2 du tableau 1.</i>	$2y+1 = 5y - 5$ $2y+1-2y = 5y - 5 -2y$ (l'objectif de cette étape est d'obtenir une équation équivalente à l'équation initiale contenant l'inconnue dans un seul membre) $1+5 = 3y - 5 + 5$ (l'objectif de cette étape est d'obtenir une équation équivalente tous les termes constants étant dans l'autre membre) $6 = 3y$ $y = 2$

Dans la section 3.2.6 nous proposons une typologie d'annotations selon leur cible.

### **3.2.3. Composante « Erreur »**

La composante « Erreur » décrit un contenu facultatif qui est porté par le moyen d'une annotation. L'étude didactique réalisée sur le savoir (procédures et justifications) peut permettre d'identifier des erreurs « classiques ». Il est alors possible d'intégrer à une rétroaction, par le moyen d'une annotation, des éléments relatifs à une erreur. Ceci peut se faire de deux manières : d'une part, en présentant une erreur effectivement réalisée (mise en œuvre d'une procédure erronée, formulation d'une justification erronée) et en la mettant en regard de l'élément valide ; d'autre part, en pointant explicitement une étape d'une procédure ou un élément d'une justification comme étant l'origine ou un lieu de manifestation fréquent d'erreur.

#### *Exemples*

La rétroaction définie par le contenu « dans un triangle ABC rectangle en A,  $BC^2=AB^2+AC^2$  (théorème de Pythagore) » peut être augmentée :

- de l'annotation « Attention, avant d'utiliser le théorème de Pythagore, il faut bien vérifier que ton triangle est rectangle » qui fait référence à l'erreur classique qui est d'utiliser ce théorème hors de son domaine de validité ;

- de l'annotation « Il est fréquent, lorsque la longueur cherchée n'est pas celle de l'hypoténuse, de se tromper en ne faisant pas la soustraction » qui fait référence à la non identification du fait que, dans l'énoncé du

théorème, c'est la longueur de l'hypoténuse qui est seule dans un des membres de l'égalité.

### **3.2.4. Composante « Fonction de la rétroaction »**

La composante « Fonction de la rétroaction » est un attribut que l'on peut associer à une rétroaction. Galpérine identifie trois types d'opérations pour une action : orientation, exécution, contrôle (Galpérine, 1966). En nous inspirant de ce résultat, nous proposons trois fonctions possibles pour une rétroaction.

- *Fonction de guidage* : il s'agit d'une rétroaction destinée à un apprenant qui ne s'est pas encore engagé dans l'activité de résolution de la tâche (peu importe la raison). La rétroaction vise à lui fournir un moyen de débiter son activité. Ce moyen peut être le rappel d'une justification (définition ou propriété) ou la présentation d'une (partie d'une) procédure permettant l'entrée dans l'activité de résolution de la tâche. Dans ce cas, il serait plus juste de parler de rétro-inaction que de rétroaction, on retrouve l'idée de *feed-forward* proposée dans la taxonomie de J. Murray *et al.* (Murray *et al.*, 2018, p. 86).

- *Fonction d'aide à la reprise de la tâche* : il s'agit d'une rétroaction destinée à un apprenant qui a réalisé, au moins partiellement, la tâche. Il peut avoir terminé la tâche, avec une ou plusieurs erreurs, ou ne pas être arrivé à finaliser son travail. Dans ce cas, la rétroaction porte directement sur la tâche, et contient donc une ou des justifications et/ou une ou des procédures mobilisées dans la réalisation de la tâche. Elle peut être augmentée par l'intégration d'éléments relatifs à une erreur.

- *Fonction de contrôle* : il s'agit d'une rétroaction destinée à un apprenant qui a finalisé la tâche avec au moins une erreur, sans toutefois lui fournir une rétroaction d'aide à la reprise de la tâche. On va l'amener à contrôler, de manière autonome, la validité de son résultat, par exemple en lui donnant une tâche supplémentaire.

#### *Exemples*

- Dans la rétroaction « Pour commencer n'hésite pas à réaliser un schéma à partir des données de l'énoncé », le contenu « réaliser un schéma à partir des données de l'énoncé » est une procédure relative à la résolution de certains problèmes. L'annotation « Pour commencer n'hésite pas à » attribue la *fonction de guidage* à la rétroaction.

- Dans la rétroaction « Voici un exemple détaillé qui devrait t'aider à corriger ta production [procédure 1 du tableau 1] », l'annotation « Voici un exemple détaillé qui devrait t'aider à corriger ta production » attribue la

*fonction d'aide à la reprise de la tâche* à la rétroaction. La procédure 1 est un modèle de solution proposé à l'élève comme moyen de retravailler sa tâche qui est proche de la tâche utilisée pour écrire la rétroaction.

- Dans la rétroaction « Nous te rappelons que [justification 2 du tableau 1]. Ta réponse est-elle bien une solution de l'équation ? », les annotations « Nous te rappelons que » et « Ta réponse est-elle bien une solution de l'équation ? » attribuent la *fonction contrôle* à la rétroaction.

*Remarque* : un même contenu de rétroaction peut être pertinent pour plusieurs fonctions. Par exemple une rétroaction composée simplement de la justification « on appelle solution d'une équation (...) » peut avoir une fonction de guidage pour un élève qui ne commence pas la résolution d'une équation, mais peut aussi avoir la fonction de contrôle pour permettre à un élève qui a finalisé sa résolution de vérifier lui-même sa réponse. Ainsi une rétroaction peut avoir une ou plusieurs *fonctions potentielles*. Une annotation permet éventuellement d'en spécifier une.

### **3.2.5. Composante « Données »**

La dernière composante de notre modèle permettant de décrire une rétroaction caractérise le type de données utilisées pour la formulation de la rétroaction et leur rapport avec les données de la tâche à résoudre.

Nous distinguons tout d'abord deux cas :

- soit la rétroaction utilise les données de l'énoncé (valeurs des coefficients d'une équation, nom des points de la figure, longueurs des segments...), on parle d'une *rétroaction instanciée à la tâche* ;

- soit la rétroaction n'utilise pas les données de l'énoncé, on distingue alors à nouveau deux cas :

- la rétroaction est rédigée en utilisant des données numériques qui ne sont pas celles de l'énoncé, on parle alors d'une *rétroaction instanciée à une tâche* ;
- elle est rédigée en utilisant des données symboliques (lettres pour les coefficients d'une expression algébrique, lettres pour désigner des longueurs en géométrie...), on parle alors de *rétroaction générique*.

Dans le Tableau 3, nous proposons des exemples de rétroactions des trois types pour la tâche « Développer  $(y+3)(4y-5)$  ».

**Tableau 3 • Exemples de variations de rétroactions selon la composante « données »**

Rétroaction instanciée à la tâche	Rétroaction instanciée à une tâche	Rétroaction générique
Voici la première étape de la résolution, à toi de continuer : $(y+3)(4y-5) =$ $yx4y-5xy+3x4y-3x5$	Voici un exemple dont tu peux t'inspirer : $(z-4)(2z+5) =$ $zx2z+zx5-4x2z-4x5=$ $2z^2-3z-20$	Nous te rappelons la propriété de distributivité suivante : $(a+b)(c+d)=ac+ad+bc+bd$

### 3.2.6. Typologie pour la composante « Annotation »

Comme nous l'avons illustré dans les exemples proposés ci-dessus, les annotations peuvent avoir différentes cibles. Sans viser à l'exhaustivité, nous indiquons trois cibles et, pour chacune de ces cibles, nous proposons dans le Tableau 4 une typologie des annotations selon leur objectif. Les différentes colonnes indiquent la cible de l'annotation (niveau global, procédure ou justification) puis chaque ligne correspond aux différentes finalités possibles des annotations selon la cible.

**Tableau 4 • Typologie d'annotations selon la cible de la rétroaction**

Exemples d'objectifs pour une annotation qui a pour cible...		
... le niveau global de la rétroaction	... une procédure de la rétroaction	... une justification de la rétroaction
Expliciter le contenu de la rétroaction (p. ex. : voici un rappel de cours, voici un exemple résolu...)	Mettre en relation la procédure et la justification associée	Expliciter le domaine de validité de la justification (« je ne peux utiliser cette justification que si... »)
Expliciter la fonction de la rétroaction : - Guidage - Aide à la reprise de la tâche - Contrôle	Identifier les changements de types de tâches (la procédure est considérée comme un ensemble de types de tâches), ce qui vise à mettre en évidence les grandes étapes de la procédure	Expliciter la finalité de la justification (« je peux utiliser cette justification pour... »)

	Expliciter la mise en œuvre d'une étape particulière de la procédure, ce qui vise à préciser un type de tâches ingrédient de la procédure	Proposer une autre formulation de la justification (p. ex., exprimer la justification dans un autre registre)
	Présenter une mise en œuvre erronée de la procédure avec identification de l'erreur (en relation avec la mise en œuvre correcte)	Présenter une justification erronée en lien avec la justification valide
	Signaler explicitement une étape de la procédure comme lieu fréquent d'erreur	

### **3.3. Premières conclusions sur le modèle de description des rétroactions épistémiques**

Dans la section 3.2, nous avons proposé un modèle permettant de décrire des rétroactions épistémiques (voir sections 2.2 et 2.6) dans la mesure où il permet de décrire :

- les éléments du savoir présents dans la rétroaction ;
- leur nature en termes de procédure et de justification ;
- les informations complémentaires présentes sur les procédures et les justifications ;
- la prise en compte des erreurs (dont l'importance a été rappelée dans les sections 2.1 et 2.6) ;
- l'utilisation ou non des données de l'énoncé dans la rétroaction.

Il permet aussi de caractériser la rétroaction par une ou des fonctions, explicitées ou potentielles.

Nous avons donc présenté un modèle formalisé de description de rétroactions épistémiques qui permet d'intégrer les connaissances de la didactique sur les savoirs en jeu. Conformément aux objectifs fixés à la fin de la section 2.8, nous présentons dans la section suivante comment ce modèle permet d'aborder la question de la décision des rétroactions épistémiques. Une première validation du modèle de description sera l'objet de la section 5.

## **4. Décision des rétroactions**

### **4.1. Problématique de la décision**

Le modèle de rétroaction que nous avons présenté permet de décrire les rétroactions, ce qui est une étape importante pour permettre la décision de ces rétroactions. Comme nous le verrons dans la suite de cette section, il permet d'orienter l'expert dans l'élicitation des connaissances nécessaires pour le système de décision et permet également d'orienter le choix de la décision.

La décision des rétroactions dépend de multiples variables latentes sur l'apprenant et son activité qui ne sont pas observables directement (Lan *et al.*, 2017), (Murray *et al.*, 2004), ce qui rend difficile, même pour des experts du domaine, la décision de la rétroaction la plus adaptée à un élève résolvant une tâche et la prévision des conséquences possibles de cette rétroaction en termes de gain d'apprentissage pour l'élève. *À fortiori*, cette difficulté est accentuée dans le contexte des plateformes d'entraînement en ligne où le système dispose d'informations partielles voire même bruitées sur l'état de l'apprenant A, sur sa résolution R et sur les caractéristiques didactiques de la tâche T en cours. Dans la suite, nous désignons ces informations par le triplet  $\langle A|R|T \rangle$ .

Dans le cadre de cette recherche, nous tentons de répondre aux questions suivantes :

- Peut-on faire appel à des algorithmes d'apprentissage automatique (AA) pour aider à réduire l'incertitude de la décision des rétroactions ?
- Est-il possible d'entraîner un modèle d'AA sur des données éducatives pour décider des rétroactions adaptées aux élèves ?
- Quel type de modèles d'AA est le mieux adapté pour la décision des rétroactions ?

L'approche que nous proposons combine l'AA et l'expertise de didacticiens pour décider des rétroactions adaptées. Cette notion d'adaptation dépend des objectifs attachés aux rétroactions. Doivent-elles permettre la progression de l'apprenant dans le domaine enseigné ou « juste » sa réussite dans la tâche ? L'approche proposée doit être suffisamment flexible pour permettre la réalisation d'objectifs différents.

## **4.2. Modélisation et caractéristiques du problème de décision des rétroactions**

Les algorithmes d'AA ont été abondamment utilisés dans le domaine des EIAH (en particulier dans les communautés AIED et EDM) pour la modélisation de l'apprenant, la prédiction de son comportement, l'adaptation des parcours d'apprentissage, les tests adaptatifs, etc. Il existe trois types d'algorithmes d'apprentissage automatique : l'apprentissage non supervisé, l'apprentissage supervisé et l'apprentissage par renforcement (*Reinforcement Learning* ou RL) (Sutton et Barto, 2018). C'est ce dernier type d'algorithmes que nous avons choisi d'utiliser. Ce choix est motivé par les caractéristiques du problème de décision des rétroactions. En effet, nous le modélisons comme un problème de décision dans l'incertain puisqu'il existe une incertitude d'une part sur le type et le contenu de la rétroaction la plus adaptée et d'autre part sur les conséquences de la rétroaction décidée par le système sur l'apprenant. Ainsi, les caractéristiques intrinsèques des modèles par renforcement correspondent bien aux caractéristiques d'un problème de décision des rétroactions. Nous résumons ces caractéristiques comme suit.

- L'absence de superviseur : contrairement aux modèles d'AA supervisé, ici le modèle n'apprend pas sur des exemples annotés par un superviseur humain puisque les experts semblent en difficulté dans le choix de la rétroaction la mieux adaptée pour le triplet  $\langle A|R|T \rangle$ . Le système dispose juste d'un signal de récompense qui, généralement, dépend de la conséquence de la rétroaction décidée par le système sur l'apprenant.

- La récompense du système peut être retardée : la rétroaction décidée par le système ne doit pas donner nécessairement lieu à une récompense immédiate puisque l'effet d'une rétroaction ne peut être évalué positivement simplement du fait que l'apprenant a réussi la tâche en cours, des effets sur les apprentissages à moyen et long terme peuvent être privilégiés.

- La rétroaction décidée par le système à l'instant  $t$  impacte les données qu'il reçoit du triplet  $\langle A|R|T \rangle$  à l'instant  $t+1$ . En effet, la rétroaction décidée par le système peut influencer l'apprenant  $A$  et sa résolution  $R$ . Ceci est une caractéristique intrinsèque des modèles RL en comparaison avec les deux autres types de modèles d'AA.

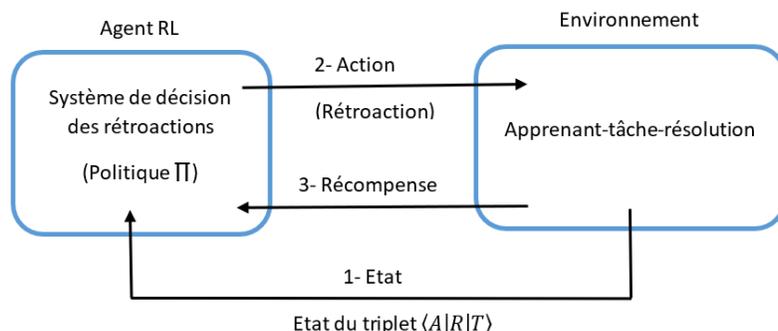
### **4.3. Apprentissage par renforcement : apprendre une politique de rétroaction adaptative**

Plusieurs travaux (He-Yueya *et al.*, 2021), (Efremov *et al.*, 2020), (Bassen *et al.*, 2020), (Chi *et al.*, 2011), (Doroudi *et al.*, 2019), (Rowe et Lester, 2015) ont montré l'intérêt de l'apprentissage par renforcement (RL) pour l'élaboration de politiques pédagogiques. Ils montrent que les modèles RL peuvent apprendre des relations complexes et souvent latentes entre les tâches pédagogiques, les actions des apprenants et leurs acquis en termes de connaissances.

De manière générale, un agent logiciel RL apprend à optimiser des décisions dans l'incertain. Il apprend une politique de sélection d'actions dans un environnement incertain, guidé par des récompenses différées, afin d'atteindre un objectif (Sutton et Barto, 2018). L'agent RL utilise un signal de récompense basé sur l'environnement afin d'apprendre une politique, notée  $\Pi$ , qui associe les états observés aux actions et maximise l'espérance de la récompense totale accumulée. Les agents, dans les problèmes d'apprentissage par renforcement, sont généralement modélisés avec des processus de décision markoviens.

Les éléments clés d'un algorithme d'apprentissage par renforcement (voir figure 2) sont les constituants du triplet (*état, récompense, action*). Dans le cas de la décision de rétroactions :

- L'*état* que reçoit l'agent RL est un ensemble d'informations du triplet  $\langle A|R|T \rangle$  (c'est-à-dire les connaissances de l'apprenant A sur le(s) sujet(s) de la tâche T et l'état de sa résolution R qui comprend, entre autres, ses erreurs) ;
- la *récompense* se rapporte aux objectifs de la stratégie (par exemple la réussite dans la tâche en cours ou l'amélioration de l'état des connaissances des apprenants) ;
- l'*action* fait référence à la décision du système en termes de rétroactions épistémiques (par exemple, présenter une rétroaction à l'élève pour l'aider à réussir la tâche en cours, comprendre son erreur, lui rappeler des éléments du cours, etc.), décision qui, sur la base des informations de l'état  $\langle A|R|T \rangle$ , vise à maximiser l'espérance mathématique de la récompense attendue.



(Niveau de l'apprenant, caractéristiques de la tâche, nombre de tentatives, erreur, etc.)

**Figure 2 · Fonctionnement du modèle RL dans le cas de la décision des rétroactions**

#### **4.4. Interventions des experts humains dans le processus de décision des rétroactions**

Les experts humains interviennent dans deux étapes de la décision des rétroactions : d'une part, pour atténuer le problème du démarrage à froid de l'algorithme de décision et, d'autre part, pour la définition de la fonction de récompense.

En éducation, l'absence de données massives antérieures disponibles et pertinentes pose le problème du démarrage à froid de l'algorithme de décision des rétroactions. Ce problème peut être critique puisque, avant que l'agent RL ait suffisamment appris une bonne politique de décision, il peut recommander des rétroactions qui ne sont pas pertinentes pour le contexte des apprenants. Ces derniers peuvent ainsi passer plus de temps dans l'environnement avant d'atteindre les objectifs ou être désorientés par les décisions de l'agent.

Pour répondre à cette problématique, dans le cadre de notre recherche, des experts ont construit un arbre de décision où les nœuds de chaque branche de l'arbre sont des variables observables du triplet  $\langle A|R|T \rangle$  intervenant dans la décision des rétroactions et où la feuille d'une branche est un ensemble de rétroactions. Cet ensemble contient les rétroactions considérées comme pertinentes didactiquement par les experts sans qu'ils aient pour autant une certitude quelconque sur leur efficacité pour un triplet observé  $\langle A|R|T \rangle$ . Il évite au moins la recommandation de rétroactions aberrantes. Ainsi, cet arbre de décision articule différentes composantes du modèle (procédure, justification, erreur, etc.) avec la connaissance de l'apprenant.

Le modèle formel de rétroactions permet aux experts de définir la fonction de récompense de l'algorithme RL. En effet, grâce aux différentes composantes du modèle de rétroaction (contenu, annotation, etc.), il est possible d'affecter une récompense immédiate plus ou moins importante à la rétroaction. Ainsi, les composantes du modèle vont permettre de guider plus finement la décision par rapport à une fonction de récompense qui considérerait la rétroaction dans sa globalité.

Ces deux interventions sont présentées plus en détail dans la section 5.2 consacrée à la présentation d'un cas d'étude.

## **5. Validations et applications**

Dans cette section nous présentons deux évaluations des modèles : tout d'abord une première évaluation qualitative de la capacité du modèle de rétroactions, présenté dans la section 3, à les décrire ; puis nous présentons une preuve de concept en explicitant l'utilisation du modèle et la mise en œuvre du processus de décision dans le cadre de l'environnement numérique d'apprentissage *Mindmath*.

### **5.1. Validation qualitative du modèle de description**

Pour exploiter l'aspect calculable du modèle de description des rétroactions, il faut tout d'abord s'assurer qu'il permet effectivement de les décrire et que cette description est identique lorsqu'elle est réalisée par plusieurs annotateurs. Nous nous sommes donc tout d'abord placés dans le cas où les rétroactions sont existantes. Pour pouvoir évaluer la portée du modèle nous avons travaillé avec des rétroactions provenant de deux environnements numériques d'apprentissage : la plateforme Carnegie Learning et la plateforme *Mindmath*. Quelques exemples de ces rétroactions sont proposés en figures 3 et 4.

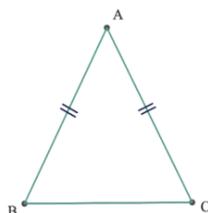
## Sticef – Vol. 29, n°2 – 2022 Sélection de la conférence EIAH 2021

### EXEMPLES DE FACTORISATIONS ✕

Énoncés	Résolutions
Factoriser $4y - 4z$	$4y - 4z = 4(y - z)$ ← Ici, 4 est un <b>facteur commun</b> aux deux termes de la différence. <span style="color: red; font-size: small;">↑    ↑ facteur commun</span>
Factoriser $4 + 4y$	$4 + 4y = 4 \times 1 + 4 \times y$ ← On écrit $4 = 4 \times 1$ pour bien faire apparaître les <b>facteurs</b> du premier terme. $4 + 4y = 4(1 + y)$
Factoriser $12 - 4y$	$12 - 4y = 4 \times 3 - 4 \times y$ ← On fait apparaître le <b>facteur 4</b> qui est <b>commun</b> aux deux termes de la différence. $12 - 4y = 4(3 - y)$
Factoriser $y^2 + 12y$	$y^2 + 12y = y \times y + 12 \times y$ ← On écrit $y^2 = y \times y$ pour bien faire apparaître les <b>facteurs</b> du premier terme. $y^2 + 12y = (y + 12)y = y(y + 12)$

### PROPRIÉTÉS GÉNÉRALES D'UN TRIANGLE ISOCÈLE ✕

Un triangle est **isocèle** si et seulement si il possède **deux côtés de même longueur**.



$$AB = AC$$

Le triangle  $ABC$  est isocèle en  $A$ .

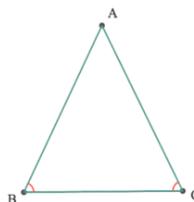
Le point  $A$  est le **sommet** du triangle isocèle.

Le segment  $[BC]$  est sa **base**.

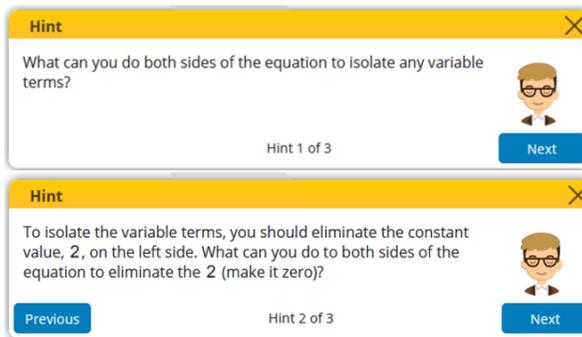
Si un triangle est isocèle, alors **les angles à la base ont même mesure**.

$$\widehat{ABC} = \widehat{BCA}$$

Réciproquement : si un triangle a deux angles de même mesure, alors il est isocèle.



**Figure 3 • Exemples de rétroactions issues de Mindmath**



**Figure 4 • Exemples de rétroactions issues de la plateforme Carnegie Learning**

### **5.1.1. Évaluation qualitative du modèle**

L'évaluation qualitative a eu lieu en deux temps.

Premier temps : lors de la 21<sup>e</sup> école d'été de didactique des mathématiques, à l'occasion d'une séance de travail réunissant cinq chercheurs confirmés et les auteurs de l'article, nous avons évalué la compréhension du modèle, la lisibilité des formulations et la complétude du modèle. Ces cinq chercheurs ne sont impliqués ni dans la conception du modèle ni dans le projet *Mindmath*. Ils disposent en revanche d'une expertise sur les cadres didactiques mobilisés dans le modèle. Dans cette évaluation nous avons brièvement présenté à l'oral le modèle, puis nous leur avons demandé de l'exploiter pour décrire un ensemble de rétroactions. Ils disposaient pour cela d'un tableau (constitué de trois zones permettant de décrire le contenu, les annotations et les informations associées, et les données) et d'une version antérieure de la figure 1.

Deuxième temps : les auteurs de l'article ont procédé, en parallèle et sans interactions entre eux, à l'annotation de cinq rétroactions (trois issues de Carnegie et deux issues de *Mindmath*) à partir d'un guide de description donné préalablement. Ce guide était constitué d'une version antérieure de la figure 1 et d'une procédure à suivre pour chaque rétroaction :

- 1. décrire le contenu principal,
- 2. identifier la présence d'éléments relevant de la composante annotations,
- 3. préciser la fonction,

- 4. caractériser le type de données.

Même si tous les auteurs n'ont pas le même niveau de familiarité avec le modèle et/ou avec la didactique, choisir les auteurs comme annotateurs permet de limiter les effets liés à des difficultés d'appropriation du modèle.

Les enseignements suivants peuvent être tirés de ces deux évaluations.

- La première évaluation nous a amenés à repréciser la description de certaines composantes du modèle et à affiner des choix de vocabulaire, en particulier en ce qui concerne les fonctions des rétroactions.

- Le modèle semble permettre d'identifier tous les éléments présents dans une rétroaction (aucun expert n'a souhaité ajouter une composante supplémentaire), mais ne permet pas de décrire certaines caractéristiques de ces éléments. Ceci provient du fait que pour passer de la modélisation de la rétroaction à sa réification des choix ont été réalisés. Ainsi, les registres sémiotiques dans lesquels sont représentés certains éléments de la rétroaction, des choix typographiques pour mettre en évidence certains morceaux de texte, la tournure des phrases (forme interrogative ou affirmative) sont apparus comme des éléments de la rétroaction ne pouvant être décrits avec le modèle.

- Comme on peut le constater sur les figures 3 et 4, les rétroactions proposées dans *Mindmath* sont beaucoup plus composites que celles proposées dans Carnegie Learning. Conformément à l'intuition, il y a plus de divergences inter-annotateurs pour les rétroactions avec beaucoup de contenus. Si les composantes *contenu* et *données* ne provoquent globalement pas de divergences, les annotations présentent plus de difficultés. Celles-ci proviennent essentiellement du niveau de granularité de ce qui est considéré comme une seule annotation. Ainsi, si on considère la rétroaction relative aux propriétés générales d'un triangle isocèle (issue de *Mindmath*) proposée en figure 3, l'ensemble des annotateurs identifient deux contenus : chacune des phrases en rouge est considérée comme un contenu de type *justification*. En revanche, pour la composante annotation, il est proposé soit de considérer les deux ensembles « figure + formule » comme deux annotations, soit de considérer quatre annotations distinctes (chaque figure et chaque formule). Ces écarts sur le niveau de granularité de ce qui est considéré comme annotation impliquent des réponses différentes sur la fonction de l'annotation. Cependant, quand les annotations considérées sont les mêmes, il y a accord sur les fonctions.

- Un dernier élément identifié lors de ce premier travail d'évaluation est la prise en compte inégale, par les annotateurs, de deux éléments : le titre de la rétroaction (quand il existe) est soit considéré comme une annotation au niveau global de la rétroaction, soit ignoré ; la formulation de la rétroaction (forme affirmative ou interrogative) est parfois considérée comme une annotation qui confère une fonction particulière à la rétroaction (fonction de guidage dans le cas de la forme interrogative par exemple). Cette observation fait écho à l'observation réalisée avec les cinq chercheur.e.s didacticien.ne.s.

### **5.1.2. Perspectives ouvertes par cette validation qualitative, autres critères de validation**

Cette première validation qualitative nous amène à identifier différentes pistes pour poursuivre ce travail d'évaluation de la capacité du modèle à décrire des rétroactions de manière non dépendante de l'annotateur, en particulier, les trois pistes qui suivent.

- Evaluer certaines composantes de manière indépendante, par exemple, en proposant des rétroactions dont les composantes « contenu » et « annotation » sont déjà explicitées, et en demandant à l'annotateur de proposer une ou plusieurs fonctions, en indiquant les raisons de son choix. Ceci pourrait permettre d'améliorer la typologie des annotations proposée dans le tableau 4 et d'étudier comment déduire de manière automatique, au moins partiellement, la fonction d'une rétroaction à partir des autres composantes.

- D'envisager un ou des moyens pour permettre la description des éléments qui ont manqué lors des deux évaluations.

- De reprendre le travail d'annotation de rétroactions par différents annotateurs en augmentant le nombre de rétroactions annotées afin d'évaluer, composante par composante et d'une manière globale, la fiabilité des descriptions obtenues. Ce travail pourrait être mené en ajoutant des rétroactions issues d'autres plateformes d'apprentissage des mathématiques, afin d'évaluer aussi la portée du modèle.

En complément de cette première validation qualitative, un autre critère de validation du modèle est développé dans les sections suivantes : une validation de la qualité des informations didactiques obtenues sur la rétroaction par l'utilisation qui en est faite, dans le cadre du projet *Mindmath*, en exploitant ces informations pour décider les rétroactions et calculer diverses informations nécessaires à la mise en œuvre d'un algorithme de *machine learning* (sections 5.2.2 et 5.2.3).

Ces premières expérimentations nous ont permis, d'une part, de réaliser un premier travail de reprise de certaines définitions ou du vocabulaire. D'autre part, elles nous incitent à construire un guide à destination de potentiels utilisateurs voulant décrire des rétroactions avec le modèle, en proposant des stratégies d'utilisation du modèle et des points d'attention (présence d'une forme interrogative dans la formulation de la rétroaction ; présence d'un titre ; etc.). La réalisation effective de ces éléments et leur expérimentation à une plus large échelle sont en cours et feront l'objet d'une publication ultérieure.

## **5.2. Preuve de concept : le cas *Mindmath***

En complément de cette première validation qualitative, le projet *MindMath* nous a permis de réifier la proposition dans un environnement informatique. Il nous a permis également une validation de la qualité des informations didactiques obtenues sur la rétroaction par l'utilisation qui en est faite en exploitant ces informations pour décider les rétroactions et calculer diverses informations nécessaires à la mise en œuvre d'un algorithme de *machine learning*.

Le projet *MindMath*<sup>1</sup> vise à la construction d'une plateforme d'entraînement aux mathématiques du collège (algèbre et géométrie). Il regroupe deux laboratoires de recherche (LDAR et LIP6) et trois entreprises (Tralalère, CabriLog et Domoscio). Une spécificité du projet est d'articuler les expertises, du LDAR en didactique, du LIP6 et de Domoscio en IA et de Tralalère et CabriLog en matière éditoriale. Dans la plateforme, l'élève réalise des exercices, organisés au sein de parcours d'entraînement adaptatifs, durant lesquels il bénéficie de rétroactions. La production des exercices et des parcours, le contenu et la décision des rétroactions sont fondés épistémologiquement et didactiquement (Jolivet *et al.*, 2021.a).

Dans cet article, nous nous concentrons sur la réification du modèle de description de rétroaction et sur la décision des rétroactions d'une manière générale. Le lecteur intéressé par d'autres éléments, en particulier sur les hypothèses didactiques qui ont fondé le travail et de premiers retours d'expérimentations peut se référer à (Jolivet *et al.*, 2021b).

---

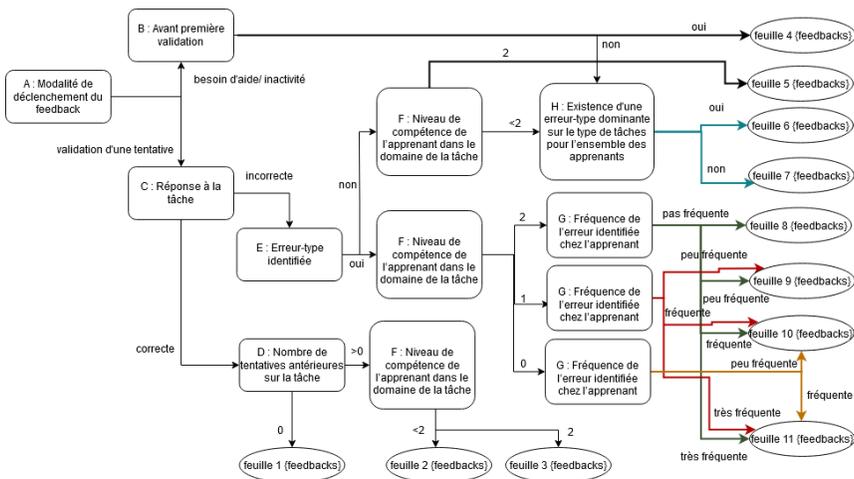
<sup>1</sup> Projet lauréat du Fond Unique Interministériel 23. <http://mindmath.education/>

Nous illustrons dans les sous-sections suivantes, pour le cas de *Mindmath*, les deux enjeux identifiés dans la section 4 : l'arbre de décision et la fonction de récompense de l'algorithme d'apprentissage.

### 5.2.1. Arbre de décision dans *Mindmath*

Le processus de décision des rétroactions s'appuie sur une implémentation de l'arbre de décision introduit dans la section 4.4. Plusieurs variables sur les apprenants (leur niveau de raisonnement, leurs erreurs classiques, etc.), les caractéristiques de la tâche et l'état de la résolution (nombre de tentatives, nature de la demande de rétroaction, nature de l'erreur, etc.) sont observées par l'agent RL.

La figure 5 présente un extrait de l'arbre de décision utilisé pour décider les rétroactions. Les valeurs 0, 1 et 2 présentes sur certaines branches spécifient le rapport de l'apprenant au savoir travaillé.



**Figure 5 • Extrait de l'arbre de décision des feedbacks construit avec des didacticiens**

Le peuplement des feuilles est réalisé avec différentes rétroactions décrites selon le modèle présenté dans la section 3.2. Les tableaux 5 et 6 présentent les peuplements réalisés respectivement pour les feuilles 5 et 11 de la figure 5, avec les pondérations précisées dans la dernière colonne. La feuille 5 correspond au cas d'une réponse erronée, sans erreur type identifiée, pour un apprenant qui possède de bonnes compétences. La feuille 11 correspond au cas d'une réponse erronée, associée à une erreur

qui est de plus très fréquente chez l'apprenant dont le niveau de compétence est faible.

**Tableau 5 • Description des rétroactions et pondérations associées pour la feuille 5 de l'arbre de décision**

Contenu	Annotation	Erreur	Fonction	Donnée	Pondération
Justification	Non	Non	Retour	Gén.	0,3
Justification	Oui	Non	Retour	Gén.	0,3
Procédure	Non	Er <sub>1</sub>	Retour	Gén.	0,2
Procédure	Non	Non	Retour	Gén.	0,2

Er<sub>1</sub> est l'erreur la plus fréquente sur des tâches similaires à celle à traiter.

**Tableau 6 • Description des rétroactions et pondérations associées pour la feuille 11 de l'arbre de décision**

Contenu	Annotation	Erreur	Fonction	Donnée	Pondération
Procédure	Oui	Er <sub>1</sub>	Retour	Inst. à une tâche	0,3
Procédure	Oui	Non	Retour	Idem	0,3
Procédure	Oui	Er <sub>1</sub>	Retour	Idem	0,2
Procédure	Oui	Non	Retour	Idem	0,2

L'approche que nous avons implémentée dans le cadre du projet *Mindmath* consiste à faire apprendre à l'agent RL une politique de décision des rétroactions au fur et à mesure de la réception des données et en maximisant l'espérance de la récompense telle que définie par les experts (voir section 4.4).

Nous avons choisi d'implémenter le *Q-learning* qui est un algorithme RL particulier (Watkins, 1989). Un des avantages du *Q-learning* est qu'il permet de comparer les récompenses probables de sélectionner telle ou telle rétroaction sans avoir un modèle de l'environnement (*Model free*). Le système de décision apprend une fonction de valeur état-action notée *Q* qui permet de maximiser la récompense sur le long terme :

$$Q: S \times A \rightarrow R$$

où  $S$  est l'ensemble des états du triplet  $\langle A|R|T \rangle$ ,  $A$  est l'ensemble des rétroactions possibles pour  $S$  et  $R$  l'espérance des récompenses futures.

L'algorithme du *Q-learning* fonctionne comme suit : au début de l'apprentissage automatique, la fonction  $Q$  est initialisée avec des pondérations données par les experts (voir Tableaux 5 et 6) qui représentent leur évaluation de la pertinence des rétroactions pour les différents états. Ensuite, à chaque choix d'action du système (dans notre cas une rétroaction), l'agent observe la récompense et le nouvel état (qui dépend de l'état précédent et de l'action actuelle). Le cœur de l'algorithme est une mise à jour de la fonction de valeur  $Q$  à chaque étape comme suit :

$$Q[s, a] := (1 - \alpha)Q[s, a] + \alpha \left( r + \gamma \max_{a'} Q[s', a'] \right)$$

où  $s'$  est le nouvel état,  $s$  est l'état précédent,  $a$  est l'action choisie,  $r$  est la récompense reçue par l'agent,  $\alpha$  est un nombre entre 0 et 1, appelé facteur d'apprentissage, et  $\gamma$  est le facteur d'actualisation.

Des expérimentations sont en cours pour comparer cette approche, qui combine l'expertise humaine avec l'algorithme du *Q-learning*, à une approche purement experte, c'est-à-dire en utilisant l'arbre et les valeurs décrites initialement.

### **5.2.2. Fonction de récompense dans *Mindmath***

Nous avons aussi fait le choix de faire appel à l'expertise humaine dans la politique des rétroactions en définissant des fonctions de récompense appropriées (Bunel *et al.*, 2018), (Gupta *et al.*, 2019), (Efremov *et al.*, 2020).

Dans le cadre de la plateforme *MindMath*, le modèle que nous avons proposé apprend à décider des rétroactions à destination d'apprenants en privilégiant leur réussite dans la tâche en cours et en minimisant la quantité d'information contenue dans ces rétroactions. La quantité d'information d'une rétroaction dépend de sa modélisation selon le modèle présenté précédemment (voir section 3.2.1). Pour chaque élément de contenu (procédure ou justification) et chaque annotation, nous avons associé une quantité d'information qui dépend du caractère plus ou moins informatif de l'élément. La quantité d'information globale d'une rétroaction est la somme des quantités d'information associées à ses composantes. Par conséquent, plus le contenu et les annotations d'une rétroaction sont nombreux, plus la récompense à l'issue de celle-ci diminue. Ce choix est motivé par le fait que plus la rétroaction est

informative, plus elle va permettre la réussite de l'élève dans la tâche sans pour autant garantir son apprentissage.

Par exemple, une rétroaction où des éléments de la procédure et de l'erreur sont affichés à l'apprenant a une quantité d'information supérieure à une rétroaction où on a uniquement un élément de justification. La première pourrait davantage favoriser la réussite de l'élève dans la tâche par rapport à la deuxième rétroaction, donc sa récompense est affaiblie pour éviter d'apprendre une politique  $\Pi$  où seules les rétroactions les plus informatives sont sélectionnées.

Bien entendu, d'autres objectifs, basés sur le modèle de rétroaction proposé, peuvent orienter la définition de cette fonction de récompense (par ex. maximiser le gain d'apprentissage des apprenants sur le long terme) et influencer ainsi la politique de décision des rétroactions. Ces objectifs n'ont pas été mis en œuvre dans le cadre du projet *Mindmath*.

### **5.3. Premier bilan du cas d'étude *Mindmath***

Lors de la conception de l'arbre, nous avons pris en compte les éléments clés relatifs à l'efficacité des rétroactions pointés dans l'état de l'art (section 2) tels que la prise en compte de l'erreur de l'apprenant et son niveau. La mise en relation de ces éléments avec les rétroactions disponibles a été permise par la description des rétroactions fondée sur des éléments didactiques.

Il faut noter que les choix réalisés pour associer les rétroactions aux feuilles ont notamment été contraints par certaines spécificités du projet (rétroactions effectivement produites, tâches présentes dans le projet, format des rétroactions). Cependant la démarche à retenir est l'exploitation du modèle pour caractériser didactiquement les rétroactions et leur associer une pondération. Cette démarche est développée dans (Jolivet *et al.*, 2021b). Il nous semble raisonnable de penser que le processus de mise en relation d'un arbre de décision avec des rétroactions (produites ou à produire) décrites à l'aide du modèle peut être exploité dans d'autres cas que celui de *Mindmath*.

Le processus de production des rétroactions, présenté dans la section 5.2.2, a aussi permis de faire émerger différentes difficultés qui ne sont pas prises en charge par le modèle dans son état actuel. Il s'agit d'ailleurs d'éléments qui ont aussi été mis en évidence lors la première évaluation réalisée avec les experts didacticiens : le modèle, s'il permet de définir le contenu de la rétroaction, ne permet en revanche pas de rendre compte

de la réification de ces contenus. Or, par exemple, le choix du registre de représentation sémiotique pour formuler une justification a un impact didactique connu dans de nombreux cas et peut aussi avoir un impact sur l'utilisabilité de la rétroaction par l'apprenant, selon qu'il est familier ou non du registre utilisé (par exemple le théorème de Pythagore donné en langue naturelle ou avec une figure accompagnée d'une formule).

Enfin, la modélisation formelle des rétroactions a permis d'aider à définir les éléments du processus de décision (l'arbre de décision, la fonction de récompense). Cependant, il reste à évaluer, du point de vue informatique, la convergence de l'algorithme de décision sur des données réelles.

## **6. Limites et perspectives**

Nous concluons cet article par deux propositions d'ouverture, l'une relative à l'exploitation du modèle, l'autre en lien avec les questions soulevées sur un plan informatique.

### **6.1. Un modèle pour étudier des environnements numériques d'apprentissage, ou aider à leur conception**

Dans l'article nous avons présenté notre modèle de rétroaction comme moyen de décrire des rétroactions, avec une dimension didactique, et son exploitation pour leur décision.

Nous soulignons deux autres exploitations pertinentes du modèle :

- guider la production de rétroactions dans le cadre de la conception ou de l'enrichissement d'un environnement numérique d'apprentissage ;
- évaluer des EIAH en décrivant, analysant, comparant les rétroactions, selon les différentes composantes du modèle, dans un ou plusieurs environnements numériques d'apprentissage.

Si cette seconde perspective n'a été qu'ébauchée lors de l'analyse de rétroactions de *Carnegie Learning* et *MindMath* dans le cadre de la validation du modèle, la première a été largement explorée dans le cadre du projet *MindMath*.

Ainsi, dans le projet *MindMath*, le modèle de description a été exploité pour guider la définition et la production des rétroactions. Plus précisément, il a été utilisé comme langage de communication entre trois groupes d'acteurs du projet :

- les experts en didactiques qui souhaitaient caractériser les rétroactions pertinentes de leur point de vue ;
- le producteur de rétroactions, non didacticien, qui devait articuler attentes des didacticiens et contraintes de production de la plateforme (contraintes temporelles, éditoriales et techniques) ;
- les informaticiens en charge de la décision des rétroactions qui avaient besoin de rétroactions décrites d'une manière standardisée.

La capacité du modèle à remplir cette fonction auprès des trois groupes d'acteurs a débouché sur la production effective de 142 rétroactions, en lien avec 5 thèmes mathématiques (la réduction, le développement et la factorisation d'expressions algébriques ; la résolution des équations du 1<sup>er</sup> degré ; la construction de triangles à partir d'informations sur les côtés et les angles). Cette production a permis la prise en compte des attentes des experts en didactique et la possibilité pour les informaticiens de gérer de manière systématique, et avec un coût de développement moindre, leur recommandation. Pour des précisions sur les limites et difficultés liées à ce processus nous renvoyons le lecteur vers (Jolivet *et al.*, 2021b). Il s'agit pour nous d'un critère supplémentaire de validation du modèle.

Un enseignement important peut être tiré de cette expérience pour des travaux de conception et de développement d'EIAH : lors de la production effective des rétroactions, même si elles sont décrites assez finement, par exemple à l'aide de notre modèle, il y a un nombre important de réifications possibles de la description. En effet, pour un même contenu, l'attribution de valeurs différentes aux composantes *annotations, données, fonctions*, permet de définir un ensemble de rétroactions variées. Certains de ces choix peuvent ne pas s'avérer neutres sur un plan didactique, d'autres sont contraints par les spécificités techniques de l'environnement.

Au final cela met en évidence deux éléments : tout d'abord qu'une analyse des effets des rétroactions sur l'apprenant ne peut être menée uniquement en s'appuyant sur leur description à partir du modèle mais peut nécessiter, en particulier en cas de résultats contradictoires ou difficilement interprétables, d'examiner les rétroactions réellement formulées et pas seulement leur catégorie. D'autre part, si on souhaite produire des rétroactions, il faut un prolongement du modèle qui permettrait de préciser la forme, et pas uniquement les contenus, des rétroactions à produire. L'exploitation effective d'un tel modèle étendu

nécessiterait la conception d'outils auteurs, à la fois pour l'expert du domaine qui souhaite décrire les rétroactions à produire, et pour le producteur qui doit assurer la production effective de celles-ci. Il s'agit de considérations qui sont hors du champ de notre travail actuel.

Il reste enfin important de garder à l'esprit qu'en complément des problématiques liées à la production des rétroactions que nous avons esquissées, il n'est pas possible d'avoir des certitudes sur l'effet d'une rétroaction, ni sur les causes de cet effet, comme le pointe Rezat : « *no matter how carefully the tasks in e-textbooks and the feedback messages are designed, it is very likely that they still contain irrelevant information, which might become salient in students' solution processes. Therefore, tasks, feedback messages, and diagrams need to be designed very carefully and as unambiguously as possible* » (Rezat, 2021, p. 1443).

Cet élément et la diversité des rétroactions potentielles motivent l'utilisation d'un algorithme de décisions de type *machine learning* qui permettra, au fil des itérations, de préciser quelles sont les rétroactions les plus efficaces. Notre problématique est réellement à l'intersection de questions informatiques et didactiques : conception, initialisation, choix des paramètres, etc. des algorithmes ; puis capacité à interpréter les résultats. Nous développons brièvement ces perspectives informatiques dans la section suivante.

## **6.2. Perspectives informatiques**

Nous avons abordé le problème difficile de la décision des rétroactions en combinant l'expertise humaine et l'algorithme *Q-learning*. L'expertise humaine intervient à deux niveaux de la décision :

- d'abord, dans la construction de l'arbre de décision permettant de relier des variables sur l'apprenant, sur sa résolution et sur la tâche en cours à un ensemble de rétroactions possibles ;
- ensuite, dans la définition de la fonction de récompense du *Q-learning*.

L'implémentation de l'algorithme de décision a été testée sur des données synthétiques pour vérifier la convergence et déterminer la valeur de certains hyperparamètres de l'algorithme d'apprentissage du *Q-learning* (les facteurs  $\alpha$  et  $\gamma$  pour la fonction de valeur par exemple). Ces données synthétiques ont été générées en considérant des comportements théoriques d'élèves. La démarche qui a guidé cette génération de données est la suivante : considérer trois modes de

raisonnement des élèves théoriques (idoine, en construction, non adapté) - voir (Jolivet et Grugeon, 2022) pour plus de précisions - et définir pour chaque mode une probabilité d'échec de l'élève dans la tâche, qui dépend également du type de rétroaction reçue (procédure ou justification). Nous avons par exemple défini une forte probabilité d'échec pour des élèves ayant un mode de raisonnement non adapté et ayant reçu des rétroactions de justification peu informatives. Ces données synthétiques ont été utilisées pour rechercher les hyperparamètres du *Q-learning* et non pas pour l'apprentissage de la politique de décision.

Des expérimentations en situation écologique avec des élèves dans le cadre du projet *MindMath* sont prévues prochainement et vont permettre de recueillir à grande échelle des données issues de la plateforme. En particulier, nous travaillons sur différentes fonctions de récompense et ces expérimentations permettront de les tester sur des données réelles.

Il est également nécessaire de comparer cette approche mixte avec d'autres travaux récents sur la décision de rétroactions (Price *et al.*, 2017), (Efremov *et al.*, 2020).

## RÉFÉRENCES

- Bassen, J., Balaji, B., Schaarschmidt, M., Thille, C., Painter, J., Zimmaro, D., Games, A., Fast, E. et Mitchell, J. C. (2020). Reinforcement learning for the adaptive scheduling of educational activities. *CHI*, 1-12.
- Bimba, A. T., Idris, N., Al-Hunaiyyan, A., Mahmud, R. B. et Shuib, N. L. B. M. (2017). Adaptive feedback in computer-based learning environments: A review. *Adaptive Behavior*, 25(5), 217-234. <https://doi.org/10.1177/1059712317727590>
- Bouhineau, D. et Nicaud, J.-F. (2006). Environnements informatiques pour l'apprentissage humain. Dans M. Grandbastien et J.-M. Labat (dir.), *Environnements informatiques pour l'apprentissage humain* (p. 333-350). Hermès-Lavoisier.
- Brooks, C., Carroll, A., Gillies, R. M. et Hattie, J. (2019). A Matrix of Feedback for Learning. *Australian Journal of Teacher Education*, 44(4), 14-32.
- Bull, S., et Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education*, 26, 293- 331.
- Bunel, R., Hausknecht, M., Devlin, J., Singh, R. et Kohli, P. (2018). *Leveraging grammar and reinforcement learning for neural program synthesis*. arXiv. <http://arxiv.org/abs/1805.04276>
- Cazes, C. et Vandebrouck, F. (2008). L'activité des élèves sur les bases d'exercices en ligne. Dans F. Vandebrouck (dir.), *La classe de mathématiques : activités des élèves et pratiques des enseignants* (p. 149-218). Octarès.
- Chaachoua, H., Bessot, A., Romo, A. et Castela, C. (2019). Developments and functionalities in the praxeological model. Dans M. Bosch, Y. Chevallard, F. Javier Garcia et J. Monaghan (dir.), *Working with the anthropological theory of the didactic: A comprehensive casebook* (p. 41-60). Routledge.

**Sébastien JOLIVET, Amel YESSAD, Mathieu MURATET, Elann LESNES, Brigitte GRUGEON-ALLYS, Vanda LUENGO**

Chevallard, Y. (1992). Concepts fondamentaux de la didactique : Perspectives apportées par une approche anthropologique. *Recherche en didactique des mathématiques*, 12(1), 83-121.

Chi, M., VanLehn, K., Litman, D. et Jordan, P. (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.

Clement, B., Roy, D., Oudeyer, P.-Y. et Lopes, M. (2015). Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining*, 7(2), 20-48. <https://doi.org/10.5281/zenodo.3554668>

Dawson, P., Henderson, M., Ryan, T., Mahoney, P., Boud, D., Phillips, M. et Molloy, E. (2018). Technology and feedback design. Dans M. J. Spector, B. B. Lockee et M. D. Childress (dir.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (p. 1-45). Springer. [https://doi.org/10.1007/978-3-319-17727-4\\_124-1](https://doi.org/10.1007/978-3-319-17727-4_124-1)

Doroudi, S., Aleven, V. et Brunskill, E. (2019). Where's the reward? *International Journal of Artificial Intelligence in Education*, 29(4), 568-620.

Duval, R. (1993). Registres de représentation sémiotique et fonctionnement cognitif de la pensée. *Annales de didactique et de sciences cognitives*, 5, 37-65.

Efremov, A., Ghosh, A. et Singla, A. (2020). Zero-shot learning of hint policy via reinforcement learning and program synthesis. Dans *Proceedings of the International Conference on Educational Data Mining (EDM 2020)* (p. 388-394).

Frenoy, R., Soullard, Y., Thouvenin, I. et Gapenne, O. (2016). Adaptive training environment without prior knowledge: Modeling feedback selection as a multi-armed bandit problem. Dans *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (p. 131-139). ACM.

Galpérine, P. (1966). Essai sur la formation par étapes des actions et des concepts. Dans A.-N Léontiev, A. Luria et A. Smirnov (dir.), *Recherches psychologiques en URSS* (p. 114-132). Moscou : Éditions du Progrès.

Gupta, R., Kanade, A. et Shevade, S. K. (2019). Deep reinforcement learning for syntactic error repair in student programs. Dans *Proceedings of the 33th AAAI Conference on Artificial Intelligence* (p. 930-937).

Hattie, J. et Gan, M. (2011). Instruction based on feedback. Dans *Handbook of research on learning and instruction* (p. 263-285). Routledge.

Hattie, J. et Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>

He-Yueya, J. et Singla, A. (2021). Quizzing Policy Using Reinforcement Learning for Inferring the Student Knowledge State. Dans *Proceedings of the International Conference on Educational Data Mining (EDM 2021)* (p. 533-539).

Jolivet, S., Lesnes-Cuisiniez, E. et Grugeon-Allys, B. (2021). Conception d'une plateforme d'apprentissage en ligne en algèbre et en géométrie : Prise en compte et apports de modèles didactiques. *Annales de didactique et de sciences cognitives*, 26, 117-156.

Jolivet, S., Yessad, A., Muratet, M., Luengo, V., Reiter, B., Grugeon-Allys, B. et Lesnes-Cuisiniez, E. (2021). Feedbacks épistémiques dans une plateforme d'entraînement aux mathématiques : modèle et décision, apports croisés de l'informatique et de la didactique. Dans M. Lefevre, C. Michel, T. Geoffre, M. Rodi, L.

Alvarez et A. Karoui (dir.), *Actes de la 10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain* (p. 82-93).

Jolivet, S. et Grugeon-Allys, B. (2022). Modélisation de parcours d'apprentissage adaptés à l'apprenant dans un EIAH. Dans *Pré-Actes de la 7e conférence internationale sur la théorie anthropologique du didactique (CITAD7)* (p. 92-106).

Kluger, A. N. et DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.

Lan, A. S., Brinton, C. G., Yang, T. Y. et Chiang, M. (2017). Behavior-Based Latent Variable Model for Learner Engagement. Dans *Proceedings of the International Conference on Educational Data Mining (EDM 2017)* (p. 64-71).

Leibold, N. et Schwarz, L. M. (2015). The art of giving online feedback. *Journal of Effective Teaching*, 15(1), 34-46.

Luengo, V. (2009). *Les rétroactions épistémiques dans les Environnements Informatiques pour l'Apprentissage Humain* [Habilitation à Diriger des Recherches, Université Joseph Fourier, Grenoble I, France]. <https://hal.science/hal-00699802>

Luengo, V., Vadcard, L., Tonetti, J. et Dubois, M. (2011). Diagnostic des connaissances et rétroaction épistémique adaptative en chirurgie. *Revue des Sciences et Technologies de l'Information - Série RIA : Revue d'Intelligence Artificielle, Lavoisier*, 25(4), 499-524.

Mandin, S. et Guin, N. (2014). Basing learner modelling on an ontology of knowledge and skills. Dans *Proceedings of the 2014 IEEE 14th International Conference on Advanced Learning Technologies* (p. 321-323).

Mayo, M. et Mitrovic, A. (2001). Optimising ITS behaviour with bayesian networks and decision theory. *International Journal of Artificial Intelligence in Education*, 12, 124-153.

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human-Computer interaction*, 5(4), 381-413.

Mory, E. H. (2004). Feedback research revisited. Dans D. Jonassen (dir.), *Handbook of Research on Educational Communications and Technology* (p. 745-783). Erlbaum.

Murray J., Gasson R. et Smith J. (2018). Toward a taxonomy of written feedback messages. Dans A. Lipnevich et J. Smith (dir.), *The Cambridge Handbook of Instructional Feedback* (p. 79-96). Cambridge University Press.

Murray, R., VanLehn, K. et Mostow, K. (2004). Looking ahead to select tutorial actions: A decision-theoretic approach. *International Journal of Artificial Intelligence in Education*, 14, 235-278.

Price, T., Zhi, R. et Barnes, T. (2017). Evaluation of a data-driven feedback algorithm for open-ended programming. Dans *Proceedings of the International Conference on Educational Data Mining (EDM 2017)* (p. 192-197).

Rezat, S. (2021). How automated feedback from a digital mathematics textbook affects primary students' conceptual development: Two case studies. *ZDM–Mathematics Education*, 53, 1433-1445. <https://doi.org/10.1007/s11858-021-01263-0>

Rowe, J. P. et Lester, J. C. (2015). Improving student problem solving in narrative-centered learning environments: A modular reinforcement learning framework. Dans *Proceedings of the 17th International conference on artificial intelligence in education (AIED 2015)* (p. 419-428). Springer.

**Sébastien JOLIVET, Amel YESSAD, Mathieu MURATET, Elann LESNES, Brigitte GRUGEON-ALLYS, Vanda LUENGO**

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153-189.

Small, M. et Lin, A. (2018). Instructional feedback in mathematics. Dans A. Lipnevich et J. Smith (dir.), *The Cambridge Handbook of Instructional Feedback* (p. 169-190). Cambridge University Press.

Stobart, G. (2018). Becoming proficient: An alternative perspective on the role of feedback. Dans A. Lipnevich et J. Smith (dir.), *The Cambridge Handbook of Instructional Feedback* (p. 29-51). Cambridge University Press.

Sutton, R. S. et Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Tessier-Baillargeon, M., Leduc, N., Richard, P. R. et Gagnon, M. (2017). Etude comparative de systèmes tutoriels pour l'exercice de la démonstration en géométrie. *Annales de didactique et de sciences cognitives*, 22, 91-117.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, 16, 227- 265.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197-221.

Vu, T. M. H. et Tchounikine, P. (2020). Supporting teacher scripting with an ontological model of task-technique content knowledge. *Computers & Education*, 163, 104098.

Watkins, C. J. C. H. (1989). *Learning from delayed rewards* [Thèse de Doctorat, Université de Cambridge, Royaume Unil.

Wisniewski, B., Zierer, K. et Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.03087>