THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE

PREPAREE A L'UNIVERSITE DE FRANCHE-COMTE

Ecole doctorale n° <numéro de l'Ecole doctorale>

<Nom de l'Ecole doctorale>

Doctorat de Informatique

Par

M. SOTO FORERO Daniel

Adaptation en temps réel d'une séance d'entraînement par intelligence artificielle

Thèse présentée et soutenue à Besançon, le <date>

Composition du Jury :

<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Président
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Rapporteur
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Rapporteur
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Examinateur
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Examinatrice
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Directeur de thèse
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Codirecteur de thèse
<Civilité><Nom><Prénom>	<Fonction et établissement d'exercice>	Invité

REMERCIEMENTS

SOMMAIRE

I

CONTEXTE ET PROBLÉMATIQUES

1

INTRODUCTION

Ce chapitre introductif vise à expliquer les termes du sujet et à introduire la thématique principale abordée. Il présente la problématique de cette thèse, introduit les différentes contributions proposées et annonce le plan du manuscrit.

Comme l'indique [Nkambou et al., 2010], les environnements informatiques pour l'apprentissage humain (EIAH) sont des outils proposant des services devant permettre aux apprenants d'acquérir des connaissances et de développer des compétences dans un domaine spécifique. Pour fournir des services efficaces, le système doit intégrer une représentation des connaissances du domaine et des mécanismes pour utiliser ces connaissances. Il doit également être en mesure de raisonner et de résoudre des problèmes.

Le système *Artificial Intelligence - Virtual Trainer* (AI-VT) est un EIAH générique développé au département d'informatique des systèmes complexes (DISC) de l'institut de recherche FEMTO-ST. Cet outil informatique propose un ensemble d'exercices aux apprenants dans le cadre de séances d'entrainement. AI-VT intègre le fait qu'une séance d'entraînement se situe dans un cycle de plusieurs séances. Les réponses apportées par l'apprenant à chaque exercice sont évaluées numériquement sur une échelle prédéfinie, ce qui permet d'estimer les progrès de l'apprenant et de déduire les sous-domaines dans lesquels il peut avoir des difficultés. Une séance est générée par un système multi-agents associé à un système de raisonnement à partir de cas (RàPC) [Henriet et al., 2017]. Un apprenant choisit le domaine dans lequel il souhaite s'entraîner et AI-VT lui propose un test préliminaire. Les résultats obtenus permettent de placer l'apprenant dans le niveau de maîtrise adéquate. Le système génère ensuite une séance adaptée veillant à l'équilibre entre l'entraînement, l'apprentissage et la découverte de nouvelles connaissances. L'actualisation du niveau de l'apprenant est effectuée à la fin de chaque séance. De cette façon l'apprenant peut avancer dans l'acquisition des connaissances ou s'entraîner sur des connaissances déjà apprises.

Un certain nombre d'EIAH utilisent des algorithmes d'intelligence artificielle (IA) pour détecter les faiblesses et aussi pour s'adapter à chaque apprenant. Les algorithmes et modèles de certains de ces systèmes seront analysés dans les chapitres 3 et 4. Ces chapitres présenteront leurs propriétés, leurs avantages et leurs limites.

Le système AI-VT initial était capable d'ajuster les paramètres de personnalisation d'une séance à l'autre, mais il ne pouvait pas modifier une séance en cours même si certains exercices de celle-ci était trop simples ou trop complexes. Chaque séance était figée et devait être déroulée jusqu'à son terme avant de pouvoir identifier des acquis et des lacunes. Les travaux de cette thèse ont eu pour objectif de pallier ce manque.

La problématique principale de cette thèse est la personnalisation en temps réel du parcours d'apprentissage des apprenants dans le système AI-VT (*Artificial Intelligence - Virtual Trainer*)

Ici *le temps réel* est considéré comme étant le moment où se déroule la séance que l'apprenant est en train de suivre. Par conséquent, l'objectif est de rendre AI-VT plus dynamique pour l'identification des difficultés et l'adaptation du contenu personnalisé en fonction des connaissances démontrées.

La partie suivante présente une liste des principales contributions apportées par cette thèse à la problématique générale énoncée plus haut.

Ce travail de thèse a été éffectué au sein l'Université de Franche-Comté (UFC) devenue depuis le 1er janvier 2025 l'Université Marie et Louis Pasteur (UMLP). Ces recherches ont été menées au sein de l'équipe DEODIS du département d'informatique des systèmes complexes de l'institut de recherche FEMTO-ST, unité mixte de recherche (UMR) 6174 du centre national de la recherche scientifique (CNRS).

1.1/ CONTRIBUTIONS PRINCIPALES

La problématique principale de ces travaux de recherche a été déclinée en plusieurs sous-parties. Ces dernières sont présentées ci-dessous sous la forme de questions. Pour chacune d'elles, une ou plusieurs propositions ont été faites. Voici les questions de recherche abordées et les contributions apportées :

1. **Comment permettre au système AI-VT d'évoluer et d'intégrer de multiples outils ?** Pour répondre à cette question, une architecture modulaire autour du moteur initial d'AI-VT a été conçue et implémentée (ce moteur initial étant le système de raisonnement à partir de cas (RàPC) développé avant le démarrage de ces travaux de thèse).

2. **Comment déceler qu'un exercice est plus ou moins adapté aux besoins de l'apprenant ?** Choisir les exercices les plus adaptés nécessite d'associer une valeur liée à son utilité pour cet apprenant. Les modèles de régression permettent d'interpoler une telle valeur. Pour répondre à cette question, nous avons proposé de nouveaux modèles de régression fondés sur le paradigme du raisonnement à partir de cas et celui des systèmes multi-agents.
3. **Quel modèle et quel type d'algorithmes peuvent être utilisés pour recommander un parcours personnalisé aux apprenants ?** Pour apporter une réponse à cette question, un système de recommandation fondé sur l'apprentissage par renforcement a été conçu. L'objectif de ces travaux est de proposer un module permettant de recommander des exercices aux apprenants en fonction des connaissances demontrées et en se fondant sur les réponses apportées aux exercices précédents de la séance en cours. Ce module de révision de la séance en cours du modèle de RàPC est fondé sur un modèle Bayésien.
4. **Comment consolider les acquis de manière automatique ?** Une séance doit non seulement intégrer des exercices de niveaux variés mais également permettre à l'apprenant de renforcer ses connaissances. Dans cette optique, notre modèle Bayésien a été enrichi d'un processus de Hawkes incluant une fonction d'oubli.

1.2/ PLAN DE LA THÈSE

Ce manuscrit est scindé en deux grandes parties. La première partie contient trois chapitres et la seconde en contient quatre. Le premier chapitre de la première partie (chapitre 2) vise à introduire le sujet, les concepts, les algorithmes associés, présenter le contexte général et l'application AI-VT initiale. Le deuxième chapitre de cette partie présente différents travaux emblématiques menés dans le domaine des environnements informatiques pour l'apprentissage humain. Le chapitre suivant conclut cette première partie en présentant le raisonnement à partir de cas.

Dans la seconde partie du manuscrit, le chapitre 5 explicite l'architecture proposée pour integrer les modules développés autour du système AI-VT initial et élargir ses fonctionalités. Le chapitre 6 propose deux outils originaux fondés sur le raisonnement à partir de cas et les systèmes multi-agents pour résoudre des problèmes de régression de façon générique. Le chapitre 7 présente l'application de ces nouveaux outils de régression dans un système de recommandation intégré à AI-VT utilisant un modèle Bayésien. Ce chapitre montre de quelle manière il permet de réviser une séance d'entraînement en cours. Le chapitre 8 montre l'utilisation du processus de Hawkes pour renforcer l'apprentissage.

Enfin, une conclusion générale incluant des perspectives de recherche termine ce manuscrit.

2

CONTEXTE

Dans ce chapitre sont décrits plus en détails le contexte applicatif et le contexte technique de ces travaux de recherche. Il présente les concepts et algorithmes utilisés dans le développement des modules. Ces modules font partie des contributions de cette thèse à l'environnement informatique pour l'apprentissage humain (EIAH) appelé AI-VT (*Artificial Intelligence - Artificial Trainer*).

Ce chapitre commence par une présentation des EIAH suivie d'une présentation sommaire du fonctionnement d'AI-VT. AI-VT étant un système de raisonnement à partir de cas (RàPC) modélisé sous la forme d'un système multi-agents (SMA), cette présentation des EIAH est suivie d'une présentation du RàPC, puis des SMA. Ce chapitre se termine par une présentation de différents algoritmes et notions implémentés dans AI-VT pour la personnalisation et l'adaptation des séances d'entraînement proposées par l'EIAH.

2.1/ LES STRATÉGIES D'APPRENTISSAGE HUMAIN ET LES ENVIRONNEMENTS INFORMATIQUES POUR L'APPRENTISSAGE HUMAIN (EIAH)

2.1.1/ LES STRATÉGIES D'APPRENTISSAGE

Dans [Lalitha and Sreeja, 2020], l'apprentissage est défini comme une fonction du cerveau humain acquise grâce au changement permanent dans l'obtention de connaissances et de compétences par le biais d'un processus de transformation du comportement lié à l'expérience ou à la pratique. L'apprentissage implique réflexion, compréhension, raisonnement et mise en oeuvre. Un individu peut utiliser différentes stratégies pour acquérir la connaissance. Comme le montre la figure 2.1, les stratégies peuvent être classées entre mode traditionnel et mode en-ligne.

L'apprentissage traditionnel se réfère à l'apprentissage-type tel qu'il est fait dans une classe. Il est centré sur l'enseignant où la présence physique est l'élément fondamental

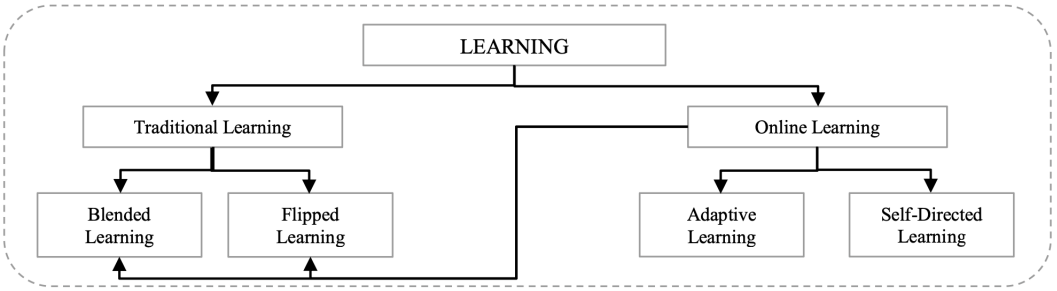

FIGURE 2.1 – Stratégies d'apprentissage ([Lalitha and Sreeja, 2020])

dans la mesure où elle implique une unité de temps et de lieu. Ici, l'enseignant interagit directement avec l'élève, et les ressources sont généralement des documents imprimés. L'apprentissage et la participation doivent y être actifs, la rétroaction y est instantanée et il existe des interactions sociales. En revanche, les créneaux horaires, les lieux et les contenus sont rigides (décidés par l'enseignant et la structure dans laquelle les enseignements sont dispensés).

L'autre stratégie globale est l'apprentissage en-ligne, qui s'appuie sur les ressources et l'information disponible sur le web. Cette stratégie incite les apprenants à être actifs pour acquérir de nouvelles connaissances de manière autonome grâce aux nombreuses ressources disponibles. Les points positifs de cette stratégie sont par exemple la rentabilité, la mise à niveau continue des compétences et des connaissances ou une plus grande opportunité d'accéder au contenu du monde entier. Pour l'apprenant, l'auto-motivation et l'interaction peuvent être des défis à surmonter. Pour l'enseignant, il est parfois difficile d'évaluer l'évolution du processus d'apprentissage de l'individu. L'apprentissage en-ligne fait aussi référence à l'utilisation de dispositifs électroniques pour apprendre (*e-learning*) où les apprenants peuvent être guidés par l'enseignant à travers des liens, du matériel spécifique d'apprentissage, des activités ou exercices.

Il existe également des stratégies hybrides combinant des caractéristiques des deux stratégies fondamentales, comme un cours en-ligne mais avec quelques séances en présentiel pour des activités spécifiques ou des cours traditionnels avec du matériel d'apprentissage complémentaire en-ligne.

2.1.2/ LES EIAH

Les EIAH sont des outils pour aider les apprenants dans le processus d'apprentissage et l'acquisition de la connaissance. Ces outils sont conçus pour fonctionner avec des stratégies d'apprentissage en-ligne et mixtes. Fondamentalement, l'architecture de

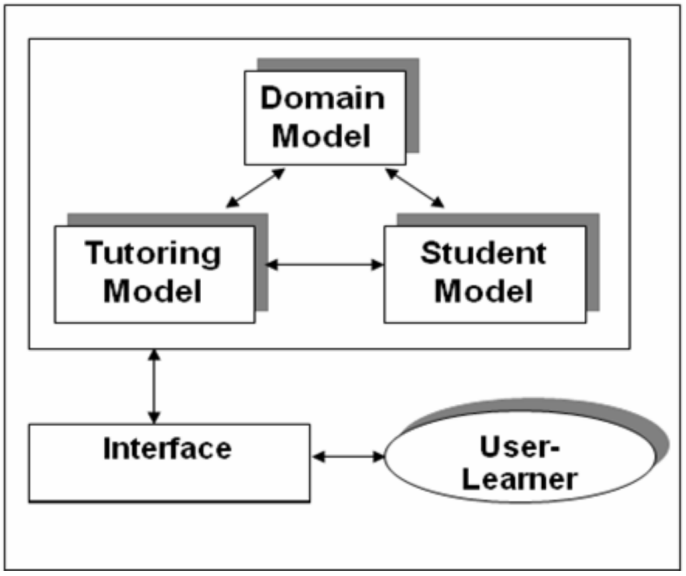

FIGURE 2.2 – L'architecture générale des EIAH, les composantes et leurs interactions ([Nkambou et al., 2010])

ces systèmes est divisée en quatre composants comme le montre la figure 2.2. Ces composants sont :

— Le *domaine* (*Domain Model* sur la figure 2.2), qui contient les concepts, les règles et les stratégies pour résoudre des problèmes dans un domaine spécifique. Ce composant peut détecter et corriger les erreurs des apprenants. En général, l'information est divisée en séquences pédagogiques.
— Le composant *apprenant* (*Student Model* sur la figure 2.2), qui est le noyau du système car il contient toute l'information utile à l'EIAH concernant l'apprenant (*User-Learner* sur la figure 2.2). Ce composant contient également les informations sur l'évolution de l'acquisition des compétences de l'apprenant. Ce module doit avoir les informations implicites et explicites pour pouvoir créer une représentation des connaissances acquises, non et partiellement acquises et l'évolution de l'apprentissage de l'apprentant. Ces informations doivent permettre de générer un diagnostic de l'état de la connaissance de l'apprenant, à partir duquel le système peut prédire les résultats dans certains domaines et choisir la stratégie optimale pour présenter les nouveaux contenus.
— L'*enseignant* (*Tutoring Model* sur la figure 2.2) est également modélisé sous la forme d'un composant. Il reçoit l'information des modules précédents, information grâce à laquelle il peut prendre des décisions sur le changement de parcours ou sur la stratégie d'apprentissage. Il peut également interagir avec l'apprenant.
— L'*interface* (*Interface* sur la figure 2.2) est le module chargé de la gestion des

configurations de l'EIAH et des interactions entre ses composants.

Le développement et la configuration de ce type de systèmes relèvent de multiples disciplines et impliquent plusieurs domaines de recherche parmi lesquels figurent la pédagogie, l'éducation, la psychologie, les sciences cognitives et l'intelligence artificielle.

2.1.3/ L'EXERCISEUR INITIAL AI-VT

Le DISC travaille depuis plusieurs années sur l'utilisation potentielle de l'intelligence artificielle dans le cadre d'un exerciseur couvrant plusieurs domaines d'apprentissage. Cet exerciseur, appelé AI-VT(Artificial Intelligence Virtual Trainer) est fondé sur différents concepts d'intelligence artificielle et ses domaines d'application sont l'apprentissage de l'aïkido, des bases de l'algorithmique et de l'anglais. Un premier réseau de convolution détermine les lacunes de l'apprenant en analysant les résultats de quelques exercices préliminaires. Puis un système de raisonnement à partir de cas distribué prend le relais et détermine une liste d'exercices à proposer en regard de ses lacunes et en tenant compte des séances qui ont déjà été proposées par le système à cet apprenant.

Dans le système AI-VT, une séance d'entraînement est proposée dans le cadre d'un cycle d'entraînement. Ainsi, un exercice proposé en première séance du cycle peut être reproposé ensuite dans une autre séance, afin de consolider les acquis et de remettre en mémoire certaines connaissances à l'apprenant. Une séance est entièrement consacrée à une seule et même compétence, déclinée en plusieurs sous-compétences. Des exercices dans différentes sous-compétences sont proposées à chaque séance. Les exercices d'une même sous-compétence sont regroupés. Un même exercice pouvant permettre de travailler deux sous-compétences différentes, un mécanisme de règlement des conflits a été implémenté dans AI-VT afin qu'un même exercice ne puisse être proposé plus d'une seule fois dans une même séance d'entraînement. Dans AI-VT, l'énoncé d'un exercice est constitué de deux parties : le contexte et la question. Cette structure modulaire permet d'associer plusieurs questions à un même contexte et inversement. Les réponses apportées par les apprenants à chaque exercice sont notées sur 10 points et le temps mis pour répondre à chaque question est comptabilisée. L'étudiant dispose d'une interface présentant un tableau de bord des exercices, sous-compétences et compétences qu'il ou elle a travaillé.

Pour l'apprentissage de l'Anglais, un robot est chargé d'énoncer les exercices de la séance. Au démarrage de cette thèse, les apprenants ont une liste personnalisée d'exercices proposée par le système AI-VT, mais c'est à l'enseignant de vérifier la validité des algorithmes/solutions proposés par les étudiants et d'identifier leurs difficultés. L'amélioration du système et de la personnalisation du parcours de l'apprenant implique une

correction automatique des exercices et cela sans utiliser un entrainement spécifique à chaque exercice. La définition de profils d'apprenants par le recueil de traces reste également à travailler pour optimiser les aides et séances d'exercices.

Le système AI-VT a été implémenté en se fondant sur deux paradigmes de l'IA : la séance d'entraînement est construite par différents modules suivant la philosophie du raisonnement à partir de cas, et l'architecture logicielle a été modélisée selon un système multi-agents. Une présentation sommaire de ces deux paradigmes de l'IA sont présentés dans cette section, et celles-ci sont suivies de la présentation de différents algorithmes et fonctions implémentés dans l'EIAH AI-VT. Des états de l'art plus complets sur les EIAH et le RàPC sont présentés dans le chapitre suivant. Le système AI-VT, son architecture et les évolutions qui ont été ralisées sur celle-ci sont détaillés dans le chapitre 4.

2.2/ LE CONTEXTE TECHNIQUE

2.2.1/ LE RAISONNEMENT À PARTIR DE CAS (RÀPC)

Le raisonnement à partir de cas est un modèle de raisonnement fondé sur l'hypothèse que les problèmes similaires ont des solutions similaires. Ainsi, les systèmes de RàPC infèrent une solution à un problème posé à partir des solutions mises en oeuvre auparavant pour résoudre d'autres problèmes similaires [Roldan Reyes et al., 2015].

Un cas est défini comme étant la représentation d'un problème et la description de sa solution. Les cas résolus sont stockés et permettent au système de RàPC de construire de nouveaux cas à partir de ceux-ci. Formellement, si P est l'espace des problèmes et S l'espace des solutions, alors un problème x et sa solution y appartiennent à ces espaces : $x \in P$ et $y \in S$. Si y est solution de x alors, un cas est représenté par le couple $(x, y) \in P \times S$. Le RàPC a besoin d'une base de N problèmes et de leurs solutions associées. Cette base est appelée base de cas. Ainsi tout cas d'indice $n \in [1, N]$ de cette base de cas est formalisé de la manière suivante : (x^n, y^n). L'objectif du RàPC est, étant donné un nouveau problème x^z (, de trouver sa solution y^z en utilisant les cas stockés dans la base de cas [Lepage et al., 2020].

Le processus du RàPC est divisé en quatre étapes formant un cycle. La figure 2.5 montre le cycle, le flux d'informations ainsi que les relations entre chacune des étapes [Leikola et al., 2018] : rechercher, adapter, réviser et stocker. Chacune des étapes est décrite dans [Richter and Weber, 2013] de la manière suivante.

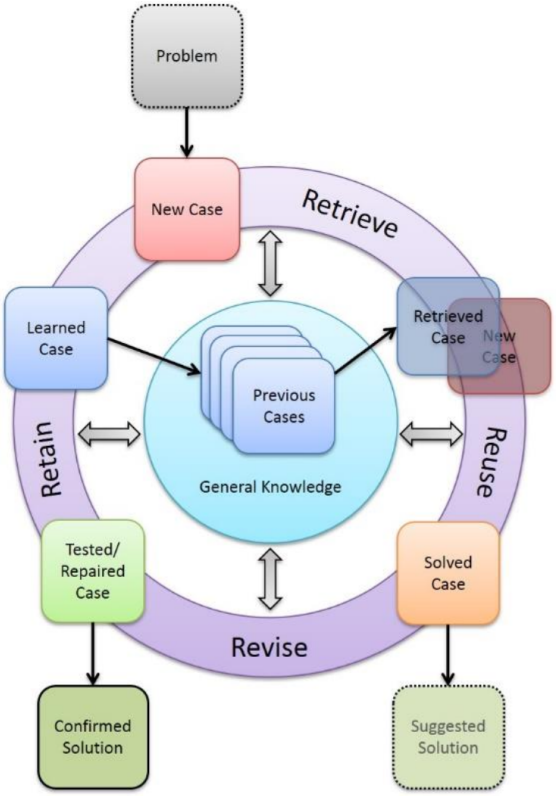

FIGURE 2.3 – Cycle fondamental du raisonnement à partir de cas ([Leikola et al., 2018])

2.2.1.1/ RECHERCHER

L'objectif de cette étape est de rechercher dans la base de cas, les cas similaires à un nouveau cas donné. Cette similarité n'est pas un concept général, mais dépend du contexte, de l'objectif et du type de données. La question fondamentale qui se pose dans cette étape est : quel est le cas le plus approprié dans la base de cas qui permet de réutiliser sa solution pour résoudre le nouveau problème donné ?

Le nouveau cas est comparé à chaque cas de la base afin d'en évaluer la similitude. La comparaison entre les cas est différente selon la structure de la base et la manière dont les problèmes sont décrits dans celle-ci. Généralement, un problème est représenté par un ensemble d'attributs aux valeurs spécifiques. Cette représentation est connue comme la représentation attribut-valeur. La similitude entre deux cas de ce type est calculée suivant l'équation 2.1. La similitude entre deux cas attribut-valeur $c_1 = a_{1,1}, a_{1,1}, ..., a_{1,n}$ et $c_2 = a_{2,1}, a_{2,1}, ..., a_{2,n}$ est la somme podérée des similitudes entre les attributs considérés individuellement. La pondération détermine l'importance de chaque attribut.

$$\sum_{i=1}^{n} \omega_i \; sim(a_{1,i}, a_{2,i}) \tag{2.1}$$

Dans cette étape, il est possible de sélectionner k différents cas similaires au nouveau cas donné.

2.2.1.2/ ADAPTER (RÉUTILISER)

L'idée du RàPC est d'utiliser l'expérience pour essayer de résoudre de nouveaux cas (problèmes). La figure 2.4 montre le principe de réutilisation d'un cas (problème) résolu (appelé *cas source*) auparavant pour le proposer en solution d'un nouveau cas (dénommé *cas cible*) après adaptation.

Si certains cas cibles sont très similaires au cas source le plus proche, il est cependant souvent nécessaire d'adapter la solution du cas source le plus similaire afin d'arriver à une solution satisfaisante pour le cas cible. De la même manière que pour la phase de recherche du cas le plus similaire, il existe de nombreuses stratégies d'adaptation. Celles-ci dépendent de la représentation des cas, du contexte dans lequel le RàPC est mis en oeuvre et des algorithmes utilisés pour l'adaptation. L'une des stratégies les plus utilisées est la définition d'un ensemble de règles transformant la ou les solutions des cas sources les plus similaires au cas cible. L'objectif ici est d'inférer une nouvelle solution y^z du cas x^z à partir des k cas sources retrouvés dans la première étape.

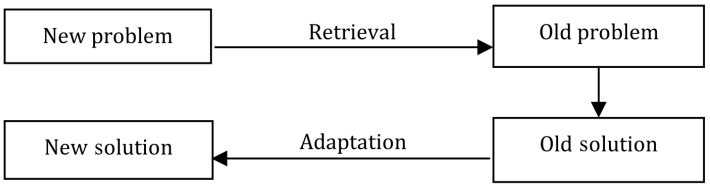

FIGURE 2.4 – Principe de réutilisation dans le RàPC ([Richter and Weber, 2013])

2.2.1.3/ RÉVISER ET RÉPARER

La solution adaptée proposée doit ensuite être évaluée et révisée afin de satisfaire certains critères de validation. L'étape de révision est chargée d'évaluer l'applicabilité de la solution cible obtenue suite à l'étape 'adaptation'. Cette évaluation peut être réalisée dans le monde réel ou dans une simulation.

L'objectif de cette phase de révision est la validation du couple (x^z, y^z). Autrement dit, le but est ici de vérifier si la solution trouvée y^z résout le problème x^z et satisfait les règles de validité et d'applicabilité. Si la solution n'est pas correcte, l'expert mettant en oeuvre la solution doit ajuster ou modifier la solution cible proposée.

Le processus d'évaluation, de révision et de réparation peut être mis en oeuvre via un processus d'apprentissage permettant d'améliorer la détection et la correction des failles des futures solutions générées. Dans les travaux de cette thèse, nous avons envisagé la possibilité d'utiliser différents outils d'apprentissage pour évaluer et réviser les solutions cibles dans certains cas particuliers.

L'évaluation peut être faite :

— par un expert humain capable d'évaluer la solution et sa pertinence dans le contexte applicatif,
— par un système fonctionnant en production et renvoyant le résultat de l'application de la solution,
— par un modèle statistique ou un système de test.

La correction des solutions peut être automatique, mais elle dépend du domaine et de l'information disponible. Un ensemble de règles peut aider à identifier les solutions non valides.

2.2.1.4/ STOCKER (RETENIR)

Si le cas cible résolu est jugé pertinent, alors celui-ci peut être stocké dans la base de cas pour aider ensuite à la résolution de futurs cas cible. La décision de stocker ou non un nouveau cas dans la base de cas doit tenir compte de la capacité de cette nouvelle base de cas à proposer de futures solutions pertinentes, évaluer et corriger les solutions cibles générées d'une part, et maintenir une base de cas d'une taille ne gaspillant pas inutilement des ressources de stockage et de calculs du système de RàPC d'autre part.

2.2.1.5/ CONTENEURS DE CONNAISSANCE

Pour pouvoir exécuter le cycle complet, le RàPC dépend de quatre sources différentes d'information. Ces sources sont appelées les conteneurs de connaissances par [Richter, 2009]. Cet article définit les conteneurs suivants dans les systèmes de RàPC : le conteneur de cas, le conteneur d'adaptation, le conteneur du vocabulaire et le conteneur de similarité :

— Le conteneur de cas contient les expériences passées que le système peut utiliser

pour résoudre les nouveaux problèmes. L'information est structurée sous la forme d'un couple (p, s), où p est la description d'un problème et s est la description de sa solution.

— Le conteneur d'adaptation stocke la ou les stratégies d'adaptation ainsi que les règles et paramètres nécessaires pour les exécuter.

— Le conteneur du vocabulaire contient l'information, sa signification et sa terminologie. Les éléments comme les entités, les attributs, les fonctions ou les relations entre les entités peuvent y être décrits.

— Le conteneur de similarité contient l'ensemble des connaissances liées et nécessaires au calcul de la similarité entre deux cas : les fonctions de calcul de similarité et les paramètres de mesure de similarité $sim(p_1, p_2)$ entre les cas p_1 et p_2.

La figure 2.5 montre les flux d'informations entre les étapes du RàPC et les conteneurs. Les flèches continues de la figure 2.5 décrivent l'ordre dans lequel les phases du cycle du RàPC sont exécutées. Les flèches avec des lignes discontinues matérialisent les flux d'information, c'est-à-dire les liens entre les étapes du cycle et les conteneurs de connaissance.

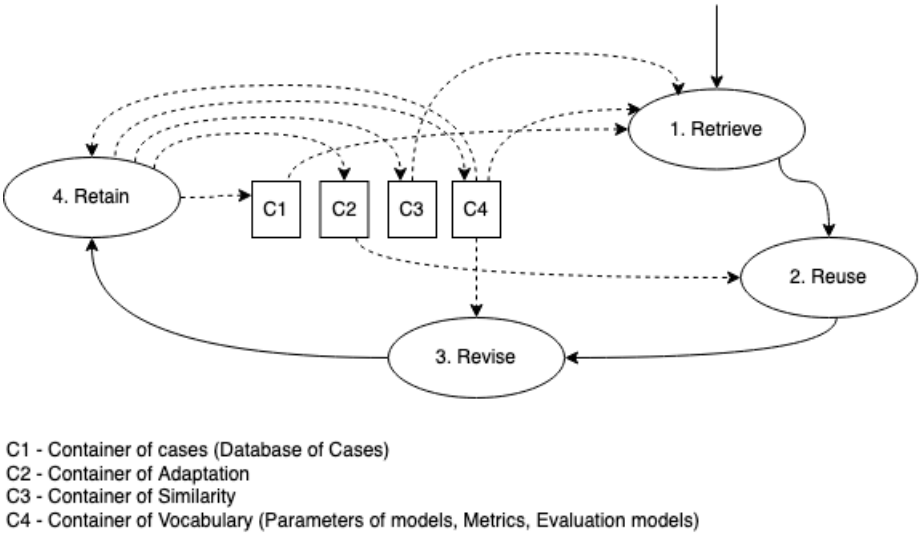

FIGURE 2.5 – Cycle du RàPC, les étapes, les conteneurs et leurs flux de données

2.2.2/ LES SYSTÈMES MULTI-AGENTS

Les systèmes multi-agents sont des systèmes conçus pour résoudre des problèmes en combinant l'intelligence artificielle et le calcul distribué. Ces systèmes sont composés de multiples entités autonomes appelées agents, qui ont la capacité de communiquer entre

elles et également de coordonner leurs comportements [Hajduk et al., 2019].

Les agents ont les propriétés suivantes :

- Autonomie : les agents fonctionnent sans intervention externe des êtres humains ou d'autres entités, ils ont un mécanisme interne qui leur permet de contrôler leurs états.
- Réactivité : les agents ont la capacité de percevoir l'environnement dans lequel ils sont et peuvent réagir aux changements qui se produisent.
- Pro-activité : les agents peuvent effectuer des changements, réagir à différents stimuli provenant de leur environnement et s'engager dans un processus cognitif interne.
- Coopération : les agents peuvent communiquer les uns avec les autres, échanger des informations afin de se coordonner et résoudre un même problème.
- Apprentissage : il est nécessaire qu'un agent soit capable de réagir dans un environnement dynamique et inconnu, il doit donc avoir la capacité d'apprendre de ses interactions et ainsi améliorer la qualité de ses réactions et comportements.

Il existe quatre types d'agent en fonction des capacités et des approches :

- Réactif : c'est l'agent qui perçoit constamment l'environnement et agit en fonction de ses objectifs.
- Basé sur les réflexes : c'est l'agent qui considère les options pour atteindre ses objectifs et développe un plan à suivre.
- Hybride : il combine les deux modèles antérieurs en utilisant chacun d'eux en fonction de la situation et de l'objectif.
- Basé sur le comportement : l'agent a à sa disposition un ensemble de modèles de comportement pour réaliser certaines tâches spécifiques. Chaque comportement se déclenche selon des règles prédéfinies ou des conditions d'activation. Le comportement de l'agent peut être modélisé avec différentes stratégies cognitives de pensée ou de raisonnement.

2.2.3/ Différents algorithmes et fonctions implémentés dans AI-VT pour la personnalisation et l'adaptation des séances d'entraînement proposées

2.2.3.1/ Pensée Bayesienne

D'après les travaux de certains mathématiciens et neuroscientifiques, toute forme de cognition peut être modélisée dans le cadre de la formule de Bayes (equation 2.2), car elle permet de tester différentes hypothèses et donner plus de crédibilité à celle qui est confir-

mée par les observations. Étant donné que ce type de modèle cognitif permet l'apprentissage optimal, la prédiction sur les événements passés, l'échantillonnage représentatif et l'inférence d'information manquante ; il est appliqué dans quelques algorithmes de machine learning et d'intelligence artificielle [Hoang, 2018].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.2}$$

La formule de Bayes calcule la probabilité a posteriori $P(A|B)$ de la plausibilité de la théorie A compte tenu des données B, et requiert trois termes : (1) une probabilité a priori $P(A)$ de la plausibilité de l'hypothèse A, (2) le terme $P(B|A)$ qui mesure la capacité de l'hypothèse A à expliquer les données observées B et (3) la fonction de partition $P(B)$ qui met en compétition toutes les hypothèses qui ont pu générer les données observées B. À chaque nouvelle évaluation de la formule, la valeur du terme a priori $P(A)$ est actualisée par la valeur du terme a posteriori $P(A|B)$. Ainsi, à chaque évaluation, le degré de plausibilité de chaque hypothèse est ajusté [Hoang, 2018].

Par la suite, nous explicitons quelques algorithmes, intègrés généralement dans l'étape "Rechercher" du RàPC, pour la recherche des cas les plus proches d'un nouveau cas.

2.2.3.2/ Méthode des k plus proches voisins (K-Nearest Neighborhood - KNN)

Comme est défini dans [Cunningham and Delany, 2021], la méthode des 'k' plus proches voisins est une méthode pour classifier des données dans des classes spécifiques. Pour faire cela, il est nécessaire de disposer d'une base de données avec des exemples déjà identifiés. Pour classer de nouveaux exemples, la méthode attribue la même classe que celle de leurs voisins les plus proches. Généralement, l'algorithme compare les caractéristiques d'une entité avec plusieurs possibles voisins pour essayer d'obtenir des résultats plus précis. 'k' représente le nombre de voisins, sachant que l'algorithme peut être exécuté à chaque fois avec un nombre différent de voisins.

Étant donné un jeu de données D constitué de $(x_i)_{i\in[1,n]}$ données (où $n = |D|$). Chacune des données est décrite par F caractéristiques qui sont des valeurs numériques normalisées $[0, 1]$, et par une classe de labellisation $y_j \in Y$. Le but est de classifier une donnée inconnue q. Pour chaque $x_i \in D$, il est possible de calculer la distance entre q et x_i selon l'équation 2.3.

$$d(q, x_i) = \sum_{f \in F} w_f \delta(q_f, x_{if}) \tag{2.3}$$

La distance entre q et x_i est la somme pondérée de toutes les distances élémentaires $delta$ calculées pour chaque caractéristique. La fonction de distance δ peut être une métrique générique. Pour des valeurs numériques discrètes il est possible d'utiliser la définition 2.4,

$$\delta(q_f, x_{if}) = \begin{cases} 0 & q_f = x_{if} \\ 1 & q_f \neq x_{if} \end{cases} \tag{2.4}$$

et si les valeurs sont continues l'équation 2.5.

$$\delta(q_f, x_{if}) = |q_f - x_{if}| \tag{2.5}$$

Un poids plus important est généralement attribué aux voisins les plus proches. Le vote pondéré en fonction de la distance est une technique couramment utilisée (équation 2.6).

$$N(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \tag{2.6}$$

Le vote est couramment calculé selon l'équation 2.7.

$$vote(y_i) = \sum_{c=1}^{k} \frac{1}{d(q, x_c)^p} N(y_j, y_c) \tag{2.7}$$

une autre alternative est basée sur le travail de Shepard, équation 2.8.

$$vote(y_i) = \sum_{c=1}^{k} e^{d(q, x_c)} N(y_j, y_c) \tag{2.8}$$

La fonction de distance peut être n'importe quelle mesure d'affinité entre deux objets, mais cette fonction doit répondre à quatre critères (équations 2.9).

$$\begin{aligned} &d(x, y) \geq 0 \\ &d(x, y) = 0, \ \textit{seulement si } x = y \\ &d(x, y) = d(y, x) \\ &d(x, z) \geq d(x, y) + d(y, z) \end{aligned} \tag{2.9}$$

2.2.3.3/ K-MOYENNES

Selon [Sinaga and Yang, 2020], K-Moyennes est un algorithme d'apprentissage utilisé pour partitionner et grouper des données. Considérons $X = \{x_1, ..., x_n\}$ un ensemble de vecteurs dans un espace Euclidien $\mathbb{R}^d$, et $A = \{a_1, ..., a_c\}$ où c est le nombre de groupes. Considérons également $z = [z_{ik}]_{n \times c}$, où z_{ik} est une variable binaire (i.e. $z_{ik} \in \{0, 1\}$) qui indique si une donnée x_i appartient au k-ème groupe, $k = 1, , c$. La fonction bijective k-moyenne est définie selon l"équation 2.10.

$$J(z, A) = \sum_{i=1}^{n} \sum_{k=1}^{c} z_{ik} \parallel x_i - a_k \parallel^2 \tag{2.10}$$

Cet algorithme minimise la fonction objectif des k moyennes $J(z, A)$: à chaque itération, les termes a_k et z_{ik} sont calculés selon les équations 2.11 et 2.12.

$$a_k = \frac{\sum_{i=1}^{n} z_{ik} x_{ij}}{\sum_{i=1}^{n} z_{ik}} \tag{2.11}$$

$$z_{ik} = \begin{cases} 1 & if \parallel x_i - a_k \parallel^2 = min_{1 \leq k \leq c} \parallel x_i - a_k \parallel^2 \\ 0, & otherwise \end{cases} \tag{2.12}$$

2.2.3.4/ MODÈLE DE MÉLANGE GAUSSIEN GMM (*Gaussian Mixture Model*)

Le modèle de mélange gaussien (GMM) est un modèle probabiliste composé de plusieurs modèles gaussiens simples. Ce modèle est décrit dans [Wang et al., 2021]. En considérant une variable d'entrée multidimensionnelle $x = \{x_1, x_2, ..., x_d\}$, GMM multivarié est défini selon dans l'équation 2.13.

$$p(x|\theta) = \sum_{k=1}^{K} \alpha_k g(x|\theta_k) \tag{2.13}$$

Dans cette équation, K est le nombre de modèles gaussiens uniques dans GMM, également appelés composants ; α_k est la probabilité de mélange de la k-ème composante respectant la condition $\sum_{k=1}^{K} \alpha_k = 1$.

θ_k représente la valeur moyenne et la matrice de covariance de chaque modèle gaussien unique : $\theta_k = \{\mu_k, \Sigma_k\}$. Pour un seul modèle gaussien multivarié, nous avons la fonction de densité de probabilité $g(x)$ (équation 2.14).

$$g(x; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right) \tag{2.14}$$

La tâche principale est d'obtenir les paramètres α_k, θ_k pour tout k défini dans GMM en utilisant un ensemble de données avec N échantillons d'entrainement. Une solution classique et disponible pour l'estimation des paramètres requis utilise l'algorithme de maximisation des attentes (Expectation-Maximization EM), qui vise à maximiser la vraisemblance de l'ensemble de données. Il s'agit d'un algorithme itératif durant lequel les paramètres sont continuellement mis à jour jusqu'à ce que la valeur delta log-vraisemblance entre deux itérations soit inférieure à un seuil donné.

2.2.3.5/ FUZZY-C

Fuzzy C-Means Clustering (FCM) est un algorithme de clustering flou non supervisé largement utilisé [Xu et al., 2021]. Le FCM utilise comme mesure de distance la mesure euclidienne. Supposons d'abord que l'ensemble d'échantillons à regrouper est $X = \{x_1, x_2, ..., x_n\}$, où $x_j \in R^d (1 \leq j \leq n)$ dans un espace Euclidien à d dimensions, et c le nombre de clusters. L'équation 2.15 montre la fonction objectif de FCM.

$$J(U, V) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} (x_j - v_i)^T A (x_j - v_i) \tag{2.15}$$

où A est la matrice métrique, m est un nombre quelconque ($m > 1$) qui dénote le degré de flou, u_{ij} est le degré d'appartenance du j-ème échantillon x_j qui appartient au i-ème cluster, dont le centre est v_i. $U = (u_{ij})$, $V = [v_1, v_2, ..., v_c], 1 \leq i \leq c, 1 \leq j \leq n, 2 \leq c < n$ satisfaisant les conditions de l'équation 2.16.

$$\sum_{i=1}^{c} u_{ij} = 1, u_{ij} \geq 0 \tag{2.16}$$

Pour compléter le modèle, les équations de mise à jour pour le centre du cluster v_i (équation 2.17) et des degrés d'appartenance u_{ij} (équation 2.18) sont définies.

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}} \tag{2.17}$$

$$u_{ij} = \frac{((x_j - v_i)^T A (x_j - v_i))^{\frac{2}{m-1}}}{\left(\sum_{h=1}^{c} (x_j - v_h)^T A (x_j - v_h)\right)^{\frac{2}{m-1}}} \tag{2.18}$$

Les algorithmes qui se trouvent dans la suite de cette section sont des algorithmes qui font partie de l'intelligence artificielle et sont utilisés dans certains EIAH pour améliorer les recommandations, essayer de corriger et de détecter les faiblesses des apprenants de façon automatique.

2.2.3.6/ BANDIT MANCHOT MAB (*Multi-Armed Bandits*)

Dans [Gupta et al., 2021] MAB est décrit comme un problème qui appartient au domaine de l'apprentissage par renforcement. Celui-ci représente un processus de prise de décision séquentielle dans des instants t de temps, où un utilisateur doit sélectionner une action $k_t \in K$ à réaliser dans un ensemble K fini d'actions possibles et donnant une récompense inconnue de l'utilisateur. L'objectif est de maximiser la récompense cumulée globale dans le temps. La clé pour trouver une solution au problème est de trouver l'équilibre optimal entre l'exploration (exécuter des actions inconnues pour collecter de l'information complémentaire) et l'exploitation (continuer à exécuter les actions déjà connues pour cumuler plus de gains). L'un des algorithmes utilisés pour résoudre ce problème c'est l'échantillonnage de Thompson.

2.2.3.7/ ÉCHANTILLONNAGE DE THOMPSON TS (*Thompson Sampling*)

Comme indiqué dans [Lin, 2022], l'algorithme d'échantillonnage de Thompson est un algorithme de type probabiliste utilisé généralement pour résoudre le problème MAB. Il s'appuie sur un modèle Bayesien dans lequel une distribution de probabilités *Beta* est initialisée. Cette distribution de probabilités est ensuite affinée de manière à optimiser la valeur résultat à estimer. Les valeurs initiales des probabilités de Beta sont définies sur l'intervalle $[0, 1]$. Elle est définie à l'aide de deux paramètres α et β.

L'échantillonnage de Thompson peut être appliqué à la résolution du problème du Bandit Manchot(MAB). Dans ce cas, les actions définies dans le MAB ont chacune une incidence sur la distribution de probabilités Beta de l'échantillonnage de Thompson. Pour chaque action de MAB, les paramètres de Beta sont initialisés à 1. Ces valeurs changent et sont calculées à partir de récompenses obtenues : si au moment d'exécuter une action spécifique le résultat est un succès, alors la valeur du paramètre α de sa distribution Beta augmente mais si le résultat est un échec alors c'est la valeur du paramètre β de sa distribution Beta qui augmente. De cette façon, la distribution pour chacune des actions possibles est ajustée en privilégiant les actions qui génèrent le plus de récompenses.

L'échantillonnage de Thompson est un cas particulier de la loi de Dirichlet comme le montre la figure 2.6. L'équation 2.19 décrit formellement la famille des courbes générées par la distribution Beta.

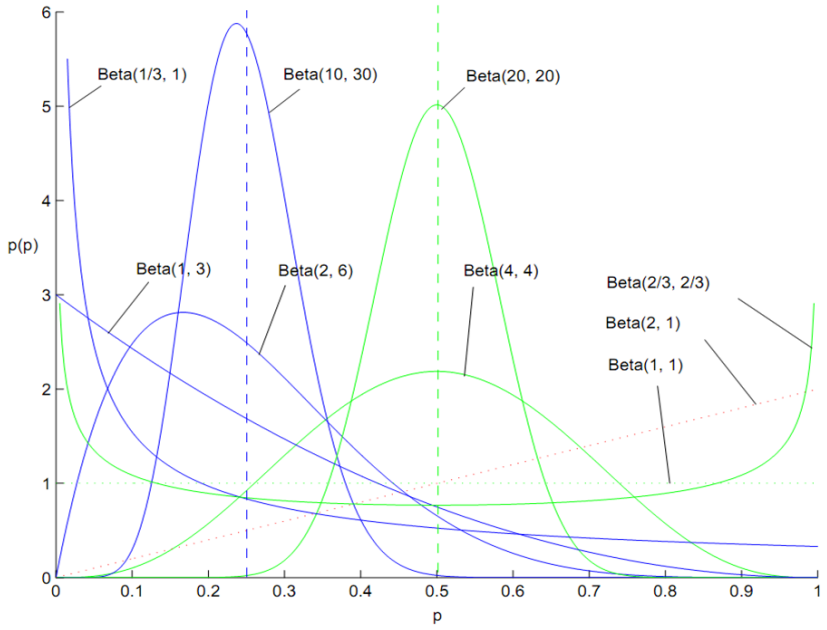

FIGURE 2.6 – Comportement de la distribution Beta avec différentes valeurs de paramètres $alpha$ et $beta$

$$B(x,\alpha,\beta) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du} & pour x \in [0,1] \\ 0 & sinon \end{cases} \tag{2.19}$$

Le chapitre suivant présente un état de l'art plus détaillé des EIAH et des systèmes de RàPC dédiés aux EIAH.

II

ÉTAT DE L'ART

3

Environnements Informatiques d'Apprentissage Humain

Dans ce chapitre sont mentionnés des travaux qui sont en rapport avec le travail de la thèse spécifiquement sur les EIAH, mais ils utilisent des approches différents, alors ici est faite une classification de ces travaux mais selon le thème principal qu'ils abordent.

3.1/ L'Intelligence Artificielle

L'intelligence artificielle a été appliquée avec succès au domaine de l'éducation dans plusieurs écoles de différents pays et pour l'apprentissage de l'informatique, les langues étrangères et les sciences comme le montre le travail de [Zhang. and Aslan, 2021] qui analyse des articles dans la période de 1993-2020, où il est importante de voir aussi que il n'y a pas une distinction en age, ni en niveau culturel, ni en niveau éducatif. En plus, avec les différentes techniques d'intelligence artificielle qui sont appliquées il est possible d'adapter la stratégie d'apprentissage pour n'importe qui et maximiser le rendement et acquisition des connaissances. Une caractéristique assez notable est que les représentations et algorithmes peuvent changer et évoluer. Aujourd'hui il y a encore beaucoup de potentiel pour développer encore les capacités de calcul et d'adaptation.

Un travail plus analytique de l'utilisation de l'intelligence artificielle dans l'éducation est [Chiu et al., 2023] qui présente une révision des travaux d'intelligence artificielle appliquée à l'éducation extraits des bases de données bibliographiques ERIC, WOS et Scopus dont la date de publication est dans la période 2012-2021, ce travail est focalisé sur les tendances et les outils employées pour aider dans le processus d'apprentissage. Après l'analyse de 92 travaux, une classification des contributions de l'IA est crée : apprentissage, enseignement, évaluation et administration. Dans l'apprentissage, est utilisée généralement l'IA pour attribuer des tâches en fonction des

compétences individuelles, fournir des conversations homme-machine, analyser le travail des étudiants pour obtenir des commentaires et accroître l'adaptabilité et l'interactivité dans les environnements numériques. Dans l'enseignement, peuvent être appliquées des stratégies d'enseignement adaptatives, améliorer la capacité des enseignants à enseigner et soutenir le développement professionnel des enseignants. Dans le domaine de l'évaluation, l'IA peut fournir une notation automatique et prédire les performances des élèves. L'IA aide dans le domaine de l'administration à améliorer la performance des plateformes de gestion, à soutenir la prise de décision pédagogique et à fournir des services personnalisés. Quelques-unes des lacunes identifiées de l'IA dans le domaine de l'éducation sont : les ressources recommandées par les plateformes d'apprentissage personnalisées sont trop homogènes, les données nécessaires pour les modèles d'IA sont très spécifiques, le lien entre les technologies et l'enseignement n'est pas bien clair, la plupart des travaux sont conçus pour un domaine ou un objectif très spécifique (peu de généralité), l'inégalité éducative car les technologies ne motivent pas tous les types d'élèves et parfois il y a des attitudes négatives envers l'IA, parce qu'elle est difficile à maîtriser et a intégrer dans les cours.

Les techniques d'IA peuvent aussi aider à prendre des décisions stratégiques qui visent des objectifs à long échéance comme le montre le travail de [Robertson and Watson, 2014] qui analyse plusieurs publications réalisées dans le domaine de l'utilisation de l'IA en jeux stratégiques. Les jeux stratégiques sont un très bon terrain d'entrainement parce que ils contient plusieurs règles, environnements pleins de contraintes, actions et réactions en temps réel, caractère aléatoire et informations cachées. Les techniques principales identifiées sont : l'apprentissage par renforcement, les modèles Bayésiens, la recherche en arbre, le raisonnement à partir de cas, les réseaux de neurones et les algorithmes évolutifs. Ici est notable aussi comme la combinaison de ces techniques permet dans certains cas d'améliorer le comportement global des algorithmes et d'obtenir meilleures réponses.

3.2/ Systèmes de Recommandation dans les EIAH

Les systèmes de recommandation dans les environnements d'apprentissage considèrent les exigences, les nécessites, le profil, les talents, les intérêts et l'évolution de l'apprenant pour s'adapter et recommander des ressources ou des exercices avec le but d'améliorer l'acquisition et maîtrise de concepts et des connaissances en général. L'adaptation de ces systèmes peut être de deux types, l'adaptation de la présentation qui montre aux apprenants les ressources en concordance avec leurs faiblesses et l'adaptation de la navigation qui change la structure du cours en fonction du niveau et style d'apprentissage

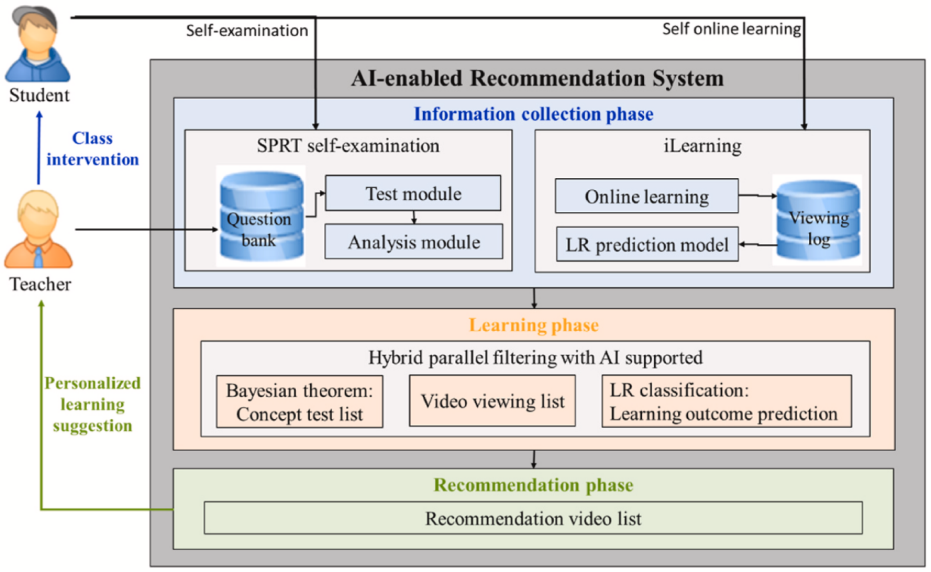

FIGURE 3.1 – Architecture du système de recommandation proposé dans [Huang et al., 2023]

de chaque apprenant [Muangprathub et al., 2020]. Ces systèmes montrent des effets positifs pour les apprenants comme le révèle le travail de [Huang et al., 2023], qui applique l'intelligence artificielle pour personnaliser des recommandations de ressources en vidéo et ainsi aider et motiver les apprenants, les effets ont été validés avec la différence entre des test préliminaires et des test après la finalisation du cours ainsi que un groupe de contrôle qui a suivi le même cours sans le système de recommandation. La figure 3.1 montre l'architecture du système où l'intelligence artificielle est utilisée dans différents étapes du processus avec les données de l'apprenant et la supervision et contrôle du professeur.

Le travail de [Seznec et al., 2020] propose un modèle générique de recommandation qui peut être utilisé dans les systèmes de recommandation de produits ou les systèmes d'apprentissage. Ce modèle est basé sur l'algorithme par renforcement UCB (Upper Confidence Bound) qui permet de trouver une solution approchée à une variante du problème bandit manchot non stationnaire (MAB), où les recompenses de chaque action diminuent chaque fois que elle est utilisée. Pour valider le modèle une comparaison avec trois autres algorithmes sur une base de données réelle a été réalisée, les métriques employées sont la moyenne accumulée du régret et la moyenne de la récompense accumulée. Avec les deux métriques, le modèle proposé est très performant.

Aussi [Ingkavara et al., 2022], met en évidence que les technologies et les systèmes de recommandation peuvent s'adapter à différents besoins et aspirations ainsi que

favoriser l'apprentissage auto-régulé. Ce type d'apprentissage aide aux apprenants à acquérir les habilités pour améliorer leur vitesse et performance ; parce que y trouvent des objectifs variables, un environnement structuré, des temps d'apprentissage variable et des ressources de support et de renforcement.

Comme vu précédemment, les techniques de l'intelligence artificielle sont largement utilisées dans l'apprentissage et l'enseignement ; l'apprentissage adaptatif pour suggérer des ressources d'étude est l'objectif du travail de [Lalitha and Sreeja, 2020] où est proposé un système avec un module de personnalisation qui collecte l'information de l'apprenant et ses exigences, un module de classification qui compare l'information avec d'autres profiles en utilisant l'algorithme KNN et un module de recommandation chargé de identifier les ressources communs entre les profiles et d'extraire information complémentaire d'Internet avec Random Forest pour améliorer le processus d'apprentissage du nouveau apprenant. Les techniques et les stratégies sont très variées mais l'objectif est de s'adapter aux apprenants et leurs besoins ; Dans [Zhao et al., 2023] les données des apprenants sont collectés et classifiés dans groupes avec des caractéristiques similaires, puis pour déterminer la performance de chaque apprenant est utilisée la méthode d'analyse par enveloppement des données (DEA) qui permet d'identifier les besoins spécifiques de chaque groupe et ainsi proposer un parcours d'apprentissage personnalisé.

Un aspect pertinent des systèmes de recommandation qu'il faut bien considérer est la représentation des données des enseignants, des apprenants et l'environnement ; parce que les multiples types de structures qu'on peut utiliser et son contenu peuvent conduire vers des résultats différents et changements dans la performance des algorithmes ou techniques employées. La proposition de [Su et al., 2022] est de stocker les données des apprenants dans deux graphes évolutifs qui contient les relations entre les questions et les réponses correctes ainsi que les réponses données par les apprenants, pour construire un graphe global de chaque apprenant avec son état cognitif et avec lequel est analysé et prédit la performance dans un contexte spécifique et ainsi pouvoir proposer un parcours personnalisé. Dans [Muangprathub et al., 2020] sont représentés les concepts comme trois ensembles : les objets (G), les attributs (M) et les relations entre G et M ; avec l'idée de les analyser avec l'approche FCA (Formal Context Analysis), Ce type de représentation permet de mettre en rapport les ressources, les questions et les sujets, ainsi si un apprenant étudie un sujet S_1 et il y a une règle qui relie S_1 avec le sujet S_4 ou la question Q_34, alors le système peut suggérer l'étude de ces ressources, l'algorithme complet explore la structure prédéfinie du cours pour recommander un parcours exhaustif d'apprentissage.

La recommandation personnalisé d'exercices est aussi une autre approche des systèmes de recommandation comme le fait [Zhou and Wang, 2021] de forme spécifique pour l'apprentissage de l'anglais, le système contient un module principal qui représente les apprenants comme des vecteurs selon le modèle DINA où sont stockes les points de connaissance acquise, le vecteur est n-dimensionnel $K = \{k_1, k_2, ..., k_n\}$ et chaque dimension correspond à un point de connaissance, ainsi si $k_1 = 1$ alors l'apprenant maîtrise le point de connaissance k_1, et si $k_2 = 0$ alors l'apprenant doit étudier le point de connaissance k_2 car il n'a pas acquise la maîtrise dans ce point-là. Sur ce type de représentation est basé la recommandation des questions proposées par le système, il y a aussi une librairie de ressources associes à chacune des possibles questions, avec lesquelles l'apprenant peut renforcer son apprentissage et avancer dans le programme du cours.

Certains travaux de recommandation et personnalisation considèrent des variables complémentaires aux notes comme dans [Ezaldeen et al., 2022], qui lie l'analyse du comportement de l'apprenant et l'analyse sémantique. La première étape consiste en collecter les données nécessaires pour créer un profile de l'apprenant, avec le profile complet, sont associés l'apprenant et un groupe de catégories prédéfinies d'apprentissage selon ses préférences et les données historiques, puis en ayant une valeur pour chaque catégorie sont cherchés les concepts associés pour les catégories et ainsi est générée une guide pour obtenir des ressources qu'on peut recommander sur le web. Le système est divisé en quatre niveaux comme montre la figure 3.2 où chacun des niveaux est chargé de une tache spécifique dont les résultats sont envoyés au niveau suivant jusqu'à arriver aux recommandations personnalisés pour l'apprenant.

Une limitation générale qui presentent les algorithmes d'IA dans les EIAH est que ce type d'algorithmes ne permettent pas d'oublier l'information avec laquelle ils ont été entrainés, une propiété nécessaire pour les EIAH dans certaines situations où peut se produire un disfonctionnement du système, la réévaluation d'un apprenant ou la reestructuration d'un cours.

Le tableau 3.1 montre un récapitulatif des articles analysés dans l'état de l'art des EIAH.

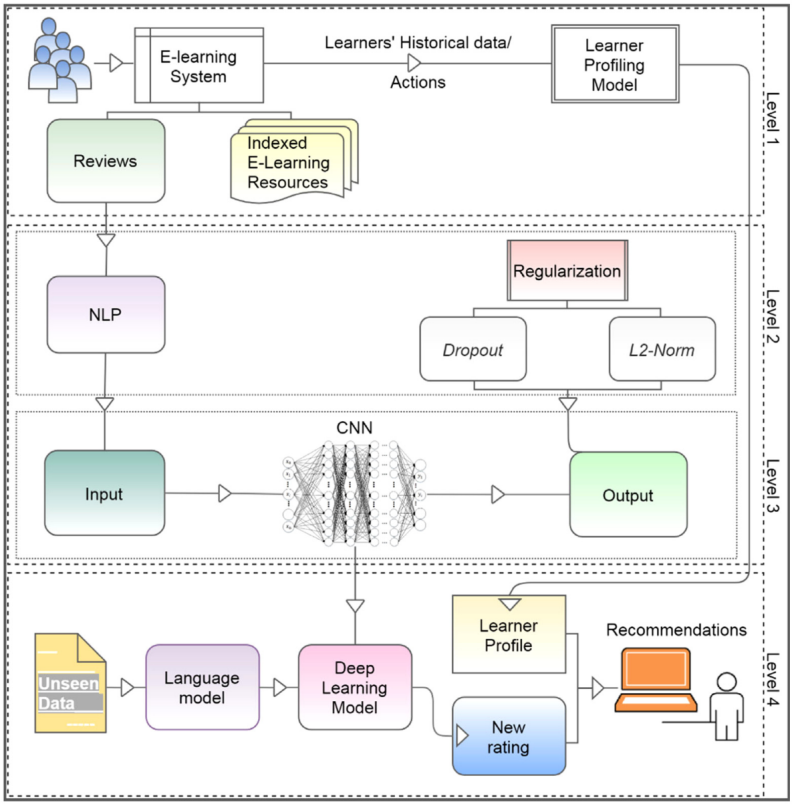

FIGURE 3.2 – Niveaux du système de recommandation dans [Ezaldeen et al., 2022]

Référence	Limites et faiblesses
[Zhang. and Aslan, 2021]	Niveau global d'application de EIAH. Analyse des effets de l'IA
[Chiu et al., 2023]	Aspects non cognitifs en IA. Motivation des apprenants
[Robertson and Watson, 2014]	Peu de test. Pas de standard d'évaluation
[Huang et al., 2023]	Recommandation en fonction de la motivation seulement. N'est pas général pour tous les apprenants
[Seznec et al., 2020]	Le modèle n'a pas été testé avec données d'étudiants. UCB prends beaucoup de temps pour explorer les alternatives
[Ingkavara et al., 2022]	Modèle très complexe. Définition de constantes subjectif et beaucoup de variables.
[Lalitha and Sreeja, 2020]	A besoin de beaucoup de données pour entrainement. Ne marche pas dans le cas 'cold-start', nést pas totalement automatisé
[Su et al., 2022]	Estimation de la connaissance de façon subjective. Il est nécessaire une étape d'entrenaiment
[Muangprathub et al., 2020]	Structure des règles complexe. Définition de la connaissance de base complexe
[Zhou and Wang, 2021]	Utilisation d'un filtre collaboratif sans stratification. Utilisation d'une seule metrique de distance
[Ezaldeen et al., 2022]	Il n y a pas de comparaison avec d'autres modèles différents de CNN. Beaucoup de variables à valeurs subjectives

TABLE 3.1 – Tableau de synthèse des articles analysés dans l'état de l'art des EIAH

4

ÉTAT DE L'ART (RAISSONEMENT À PARTIR DE CAS)

4.1/ RAISONNEMENT À PARTIR DE CAS (RÀPC)

Le raisonnement à partir de cas est un paradigme très générique, et a été utilisé dans différents domaines, en médecine a été appliqué pour explorer et aider la prise de décisions des médecins dans les traitements qui requièrent administration de doses précises selon les cas et les patients, [Petrovic et al., 2016] propose un algorithme de raisonnement à partir de cas qui indique quel est la dose correcte de radiation pour le traitement du cancer. Pour administrer une dose de radiothérapie il est nécessaire de connaître avec précision deux paramètres : le nombre de faisceaux et l'angle de chacun d'eux. L'algorithme proposé essaie de trouver la valeur optimal pour la combinaison des deux paramètres en utilisant les réseaux de neurones et l'adaptation des cas connus pour changer le nombre de faisceaux et adapter le angle de chaque faisceau. La validation de l'algorithme est évaluée avec une base de 80 cas réels de cancer du cerveau extraits de l'hôpital de Nottingham City, le nombre de neurones et couches a été changé de façon empirique. Les résultats montrent que l'utilisation des cas historiques et la construction des solutions à partir des solutions déjà connues atteint une amélioration de 12% dans les cas du nombre de faisceaux et 29% dans les cas de l'angle des faisceaux.

Plusieurs travaux ont obtenu de bons résultats en appliquant le RàPC avec des modifications dans chacune des phases ou en combinant différents algorithmes au design de produits. [Roldan Reyes et al., 2015] propose comme le montre la figure 4.1 un algorithme pour produire propylène glycol dans un réacteur chimique, dans ce cas la phase de réutilisation du RàPC est mélangée avec la recherche des états qui satisfassent le nouveau problème (Constraint satisfaction problems CSP) en utilisant l'information des cas de base déjà résolus. Les solutions trouvées sont évaluées selon le nombre de

changements faits aux solutions déjà connues (Parcimonie), le nombre de possibles solutions trouvées (Accuracy), l'évaluation de commentaires faits par des experts et la complexité des transformations réalisées. Il n'est pas obligatoire de modifier tout le cycle du RàPC pour l'adapter à une problématique spécifique, comme [Jung et al., 2009] qui développe un système basé sur l'hybridation du RàPC avec des réseaux de neurones pour concevoir des produits, le système se focalise uniquement dans les phases de rechercher et réutiliser du RàPC où sont exécutes les algorithmes implémentés. Le système détermine de façon automatique les valeurs pour les paramètres nécessaires à la conception de un produit particulier en suivant le cycle traditionnel du RàPC. Avec l'algorithme de k-moyennes est extrait un cas représentatif de la base de cas et l'adaptation des solutions des voisins trouvées est faite avec le réseau de neurones RBFN. Avec 830 test, les nouveaux cas très similaires aux cas de la base initiale est obtenu un bon design en 99.4% des cas, pour les cas dissemblables le pourcentage est de 94.8%. Le RàPC peut être utilisé aussi comme outil secondaire pour aider d'autres algorithmes à trouver la bonne configuration, dans [Butdee and Tichkiewitch, 2011] est implémenté un réseau de neurones classique pour designer la géométrie d'une matrice pour l'extrusion de l'aluminium, car actuellement c'est un processus qui se fait manuellement par essai et erreur. Le RàPC ici est utilisé pour aider à déterminer les valeurs optimales des paramètres du réseau.

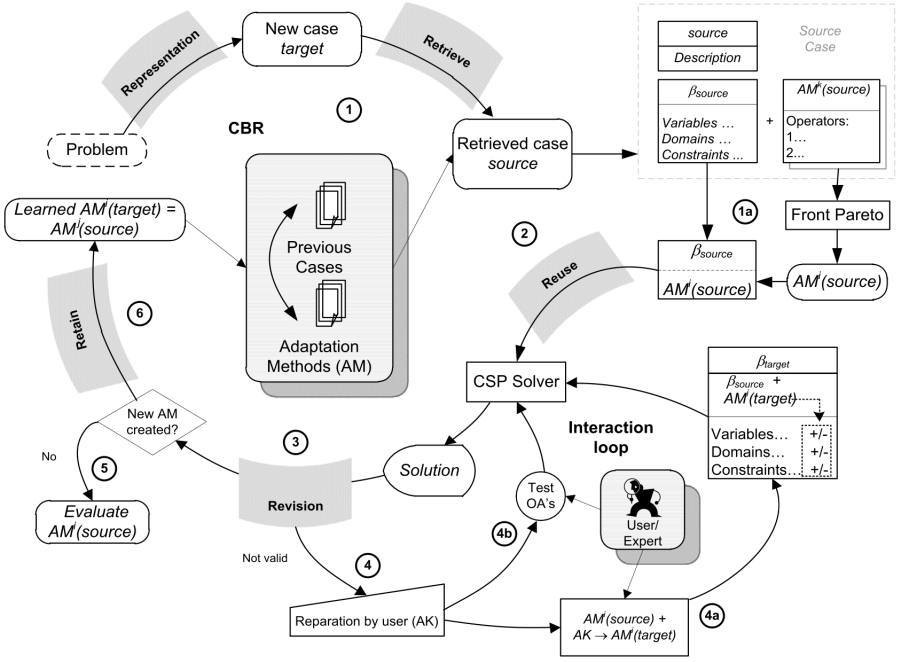

FIGURE 4.1 – Cycle du RàPC modifié. ([Roldan Reyes et al., 2015])

L'une des problématiques la plus représentative et la plus fondamentale de l'utilisation du

RàPC est la création de recettes de cuisine à partir de une base de ingrédients et d'un ensemble de recettes déjà connues, mais malgré l'apparente simplicité du problème il y a beaucoup de variations et modèles, dans [Grace et al., 2016] est abordée cette problématique, mais en additionnant un nouveau cycle au cycle traditionnel du RàPC. Le premier cycle génère des descriptions abstraites du problème avec un réseau de neurones et algorithmes génétiques ; le seconde cycle prend les descriptions abstraites comme nouveaux cas, cherche les cas similaires et adapte les solutions rencontrées. L'exécution des deux cycles considère certains critères prédéfinis par l'utilisateur. En comparant le même problème avec le cycle traditionnel du RàPC il y a effectivement une amélioration de la qualité des recettes proposées et qui vont plus en accord avec les critères définis. Sur la même problématique, mais plus focalisé sur la génération des recettes novatrices, [Maher and Grace, 2017], utilise un modèle probabiliste qui considère les préférences des utilisateurs, un indicateur de curiosité et un composant qui synthétise tout le design ; une fois que toutes les exigences sont remplies, la recette est créée en combinant les ingrédients avec un réseau d'apprentissage profond. Les changements dans la représentation des cas est aussi une source d'amélioration des résultats parce que parfois d'un côté il y a plus d'information et de l'autre certain techniques ou algorithmes fonctionnent mieux, c'est le cas que montre [Müller and Bergmann, 2015] où les recettes sont codées comme un processus de transformation et mélange des ingrédients en suivant une suite de pas ordonnés, pour créer innovation dans les recettes une mesure de distance est employée entre les ingrédients avec laquelle il est possible de trouver une recette et changer les ingrédients pour les plus proches, de la même façon il est possible de créer recettes plus proches des exigences des utilisateurs, les étapes de transformation appelées aussi opérateurs sont stockes et catégorisées les unes avec les autres avec une métrique qui permet de les échanger pour obtenir une nouvelle recette.

Une autre type d'application du RàPC très notable est le génération, analyse et correction du texte écrit ; pour la réalisation de ces tâches parfois il est nécessaire de transformer le texte en représentation numérique ou établir une fonction ou mesure qui va indiquer quels mots ont une proximité sémantique. Le travail de [Ontañón et al., 2015] utilise le RàPC pour générer des histoires en utilisant le texte d'autres histoires, en explorant les possibles transformations de forme que la nouvelle histoire ne soit pas très similaire aux histoire déjà connues mais que elle soit cohérent. Le travail est plus focalisé sur la phase de révision, car les résultats de la révision déterminent si la histoire générée corresponde avec les critères spécifiés ou si le cycle d'adaptation doit recommencer ; l'adaptation se fait en cherchant dans la base des cas, les personnages, les contextes, les objets et en unifiant tout avec des actions ; la plus part des histoires générées sont cohérents mais sont des paragraphes très courts. Dans [Lepage et al., 2020] les phrases écrites

en langue française sont corrigées, ce travail n'utilise pas la transformation numérique des phrases, ni connaissances linguistiques, mais seulement retrouve les phrases similaires en utilisant l'algorithme LCS (Longest Common Subsequence) et en calculant la distance parmi toutes les phrases pour savoir si elle est bien écrite ou pas, sinon le système peut proposer une correction en changeant certains mots selon le contexte et calculer à nouveau les distances qui déterminent la cohérence et pertinence de la phrase.

Le sport fait partie aussi des domaines explorés avec le RàPC, qui est généralement employé pour suggérer des routines d'entraînement, prédire des temps de course, ainsi comme évaluer et améliorer la performance des athlètes. La prédiction du temps de course pour un athlète est l'objectif du travail de [Smyth and Cunningham, 2018] où les coureurs ont été suivis et analysés pour la prédictions des temps de finalisation de une course, ici est implémenté un algorithme pour ceux qui font de la marathon en utilisant l'algorithme KNN pour rechercher les cas similaires et en faisant la prédiction avec la moyenne pondérée des meilleurs temps d'arrivée des cas similaires retrouvés dans la première étape du RàPC. Tandis que [Smyth and Willemsen, 2020] essaie de prédire le meilleur temps personnel pour des patineurs en se basant sur l'analogie de que les patineur avec des caractéristiques et histoire de course similaires auront des temps similaires aussi, mais parfois calculer une moyenne des temps similaires trouvés, ne suffit pas, car il y a des variables qui peuvent changer beaucoup la performance du patineur comme : le type de course, une piste spécifique, la distance de course, etc ; l'algorithme a été testé avec une base de données qui contiens l'information de 21 courses de 500m, 700m, 1000m, 1500m, 3km, 5km and 10km entre Septembre 2015 et Janvier 2020. Un système multifonctionnel apparait dans [Feely et al., 2020] car il permet de obtenir une prédiction du temps de course, suggérer un plan du rythme de la course et il recommande aussi un plan d'entraînement pour une course donnée ; les trois fonctionnalités sont implémentées avec RàPC et en utilisant 'information des coureurs qui présentent un parcours historique et des caractéristiques physiques similaires, les plans d'entraînement ont été définis génériquement sur 16 semaines avant le début de la marathon parce que c'est le temps usuel pour la préparation de tous les coureurs. Le système a été évalué avec une base de données de 21000 coureurs dans la période de 2014 à 2017 qui ont couru les marathons de Dublin, Londres ou New-York.

Les systèmes de recommandation et le RàPC peuvent aussi être combinés comme le fait [Obeid et al., 2022], dont le objectif est d'arriver à minimiser les limitations des systèmes traditionnels ainsi comme avoir la capacité d'analyser des données hétérogènes et dans des grands espaces dimensionnels, dans ce travail est recommandé aux élèves du secondaire le parcours de carrière et les universités/collèges qui correspondent à leurs intérêts, ce travail montre aussi une taxonomie des techniques algorithmiques

généralement utilisées dans les systèmes de recommandation pour les EIAH (figure 4.2). La utilisation du RàPC dans les environnements informatiques pour l'apprentissage humain (EIAH) et la personnalisation montrent des résultats positifs comme l'indique le travail de [Supic, 2018] dont le modèle suit le cycle traditionnel du RàPC en combinant les modèles d'apprentissage traditionnel et digital, les principales contributions sont : la représentation des cas et la recommandation des parcours d'apprentissage personnalisés selon l'information des autres apprenants. Pour démontrer l'effectivité du modèle est crée une base de cas initiales avec laquelle est recommandé le parcours à 120 apprenants, les résultats ont été obtenus avec l'aide des examens avant et après avoir suivi le parcours recommandé par le modèle.

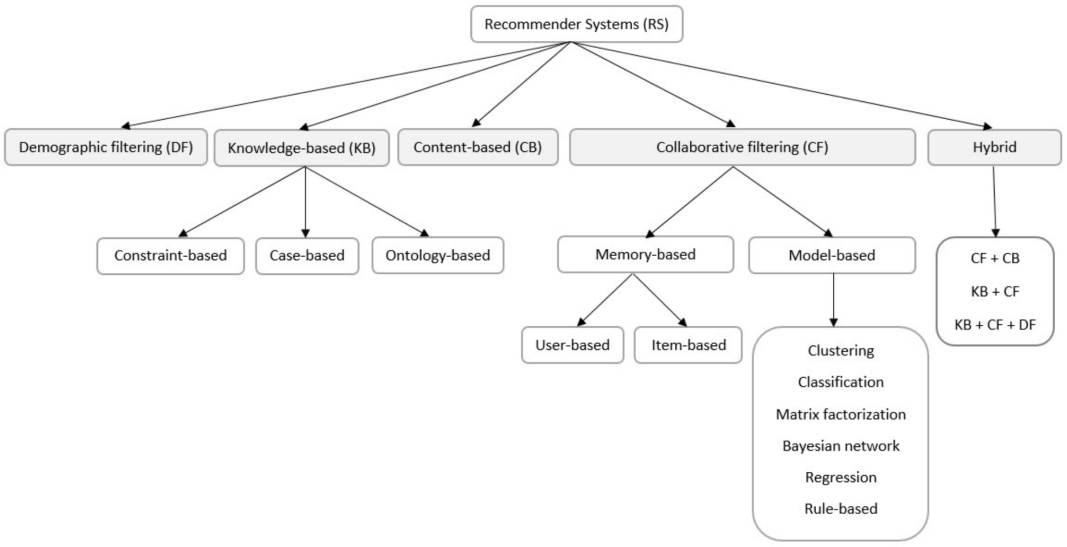

FIGURE 4.2 – Taxonomie des techniques algorithmiques employées pour des modules de recommandation dans les EIAH ([Obeid et al., 2022])

Certains travaux ont appliqué le raisonnement à partir de cas à un problème spécifique en proposant des représentations des cas et des solutions différents et d'autres ont modifié le cycle conceptuel comme est mobtré dans la figure 4.3 extraite de [Grace et al., 2016] où est additionné un autre cycle complémentaire avec l'idée d'améliorer le résultat du processus du RàPC en utilisant le *Deep Learning* ; dans [Butdee and Tichkiewitch, 2011] la phase de stockage est modifiée en retenant les cas dont les résultats n'ont pas eu de succès, pour guider le processus dans la fabrications de nouvelles pièces ; [Robertson and Watson, 2014] additionne pour chaque cas, une valeur d'utilité espérée selon chaque possible action avec laquelle est possible d'obtenir une prédiction probabiliste de l'application des actions à un état initiale donné.

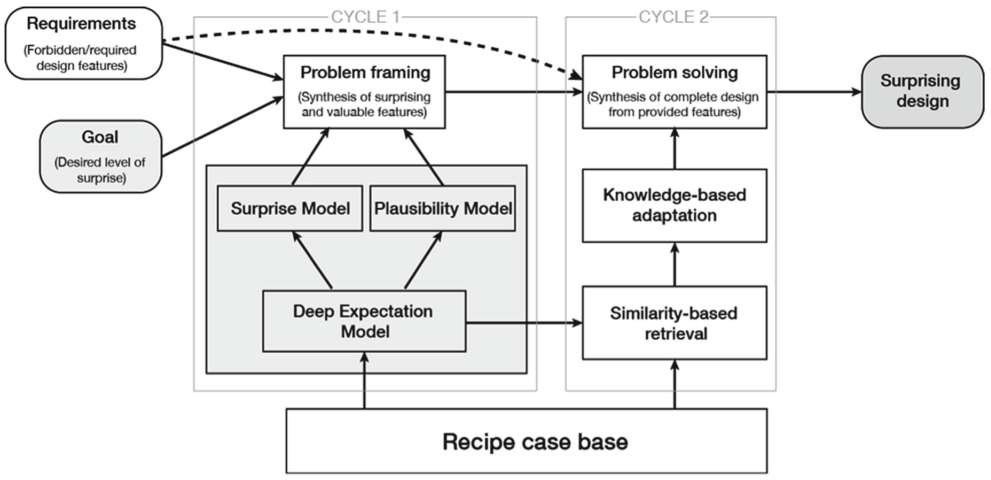

FIGURE 4.3 – Addition d'un cycle complémentaire avec *Deep Learning* au RàPC ([Grace et al., 2016])

Le tableau 4.1 montre un récapitulatif des articles analysés dans l'état de l'art du RàPC.

Deux problèmes très communes des systèmes de recommandation sont : le *cold-start* qui occurs au début d'une recommandation lorsque il n'y a pas suffisament de données pour calculer ou inférir une recommandation appropiée et le gray-sheep que se produit lorsque un utilisateur présente un comportement très différent des utilisateurs qui sont stockés dans la base de données et le système ne peut pas générer des recommandations en se basant sur l'information disponible.

Ref	Limites et faiblesses
[Petrovic et al., 2016]	Grand quantité de données pour que l'algorithme marche bien. Recommandation limitée à un problème très spécifique
[Roldan Reyes et al., 2015]	Un seul méthode d'adaptation est utilisé. Le modèle n'exploite pas les cas qui présentent une erreur
[Jung et al., 2009]	Le modèle d'adaptation proposé marche bien seulement avec des cas très proches. L'apprentissage dans le RàCP se limite à stocker les nouveaux cas
[Butdee and Tichkiewitch, 2011]	Le RàPC n'est pas modifié ou amélioré. La révision dans le RàPC n'est pas automatique
[Grace et al., 2016]	Beaucoup de données nécessaires pour entrainer le modèle de 'Deep Learning'. Les solution générées n'ont pas été validées
[Maher and Grace, 2017]	L'évaluation des solution générées n'est pas automatique. Les solutions sont produites avec un seul point de vue
[Müller and Bergmann, 2015]	Un seul approche pour adapter les cas dans le RàPC. La révision dans le RàPC n'est pas automatique
[Ontañón et al., 2015]	Les solutions générées peuvent présenter des incohérences. Il n y a pas une métrique objective pour mesurer la qualité des réponses
[Lepage et al., 2020]	Les résultats ne sont pas très bonnes. Le modèle n'a pas de connaissances linguistiques
[Smyth and Cunningham, 2018]	Les prédictions n'ont pas été testées dans le monde réel. Les données sont très spécifiques et peu variées
[Smyth and Willemsen, 2020]	Les prédictions sont pour des cas limités. Le modèle est entrainé pour un type d'utilisateur spécifique
[Feely et al., 2020]	Seulement une technique de sélection de solutions. Les données nécessaires pour le modèle sont complexes
[Obeid et al., 2022]	L'ontologie peut être limitée pour évaluer plusieurs cas inconnus ou imprévus. Il est nécessaire d'avoir une forte connaissance dans le domaine
[Supic, 2018]	Le modèle a été validé avec peu de données. Les recommandation se basent seulement dans l'information des autres apprenants

TABLE 4.1 – Tableau de synthèse des articles analysés dans l'état de l'art du RàPC

III

CONTRIBUTIONS

BIBLIOGRAPHIE

[Auer et al., 2021] Auer, F., Lenarduzzi, V., Felderer, M., and Taibi, D. (2021). From monolithic systems to microservices : An assessment framework. *Information and Software Technology*, 137 :106600.

[Butdee and Tichkiewitch, 2011] Butdee, S. and Tichkiewitch, S. (2011). Case-based reasoning for adaptive aluminum extrusion die design together with parameters by neural networks. In Bernard, A., editor, *Global Product Development*, pages 491–496, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Chiu et al., 2023] Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., and Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education : Artificial Intelligence*, 4 :100118.

[Cunningham and Delany, 2021] Cunningham, P. and Delany, S. J. (2021). K-nearest neighbour classifiers - a tutorial. *ACM Comput. Surv.*, 54(6).

[Ezaldeen et al., 2022] Ezaldeen, H., Misra, R., Bisoy, S. K., Alatrash, R., and Priyadarshini, R. (2022). A hybrid e-learning recommendation integrating adaptive profiling and sentiment analysis. *Journal of Web Semantics*, 72 :100700.

[Feely et al., 2020] Feely, C., Caulfield, B., Lawlor, A., and Smyth, B. (2020). Using case-based reasoning to predict marathon performance and recommend tailored training plans. In Watson, I. and Weber, R., editors, *Case-Based Reasoning Research and Development*, pages 67–81, Cham. Springer International Publishing.

[Grace et al., 2016] Grace, K., Maher, M. L., Wilson, D. C., and Najjar, N. A. (2016). Combining cbr and deep learning to generate surprising recipe designs. In Goel, A., Díaz-Agudo, M. B., and Roth-Berghofer, T., editors, *Case-Based Reasoning Research and Development*, pages 154–169, Cham. Springer International Publishing.

[Gupta et al., 2021] Gupta, S., Chaudhari, S., Joshi, G., and Yağan, O. (2021). Multi-armed bandits with correlated arms. *IEEE Transactions on Information Theory*, 67(10) :6711–6732.

[Hajduk et al., 2019] Hajduk, M., Sukop, M., and Haun, M. (2019). *Cognitive Multi-agent Systems : Structures, Strategies and Applications to Mobile Robotics and Robosoccer*. Studies in Systems, Decision and Control. Springer International Publishing.

[Henriet et al., 2017] Henriet, J., Christophe, L., and Laurent, P. (2017). Artificial intelligence-virtual trainer : An educative system based on artificial intelligence and

designed to produce varied and consistent training lessons. *Proceedings of the Institution of Mechanical Engineers, Part P : Journal of Sports Engineering and Technology*, 231(2) :110–124.

[Hoang, 2018] Hoang, L. (2018). *La formule du savoir. Une philosophie unifiée du savoir fondée sur le théorème de Bayes*. EDP Sciences.

[Huang et al., 2023] Huang, A. Y., Lu, O. H., and Yang, S. J. (2023). Effects of artificial intelligence–enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers and Education*, 194 :104684.

[Ingkavara et al., 2022] Ingkavara, T., Panjaburee, P., Srisawasdi, N., and Sajjapanroj, S. (2022). The use of a personalized learning approach to implementing self-regulated online learning. *Computers and Education : Artificial Intelligence*, 3 :100086.

[Jung et al., 2009] Jung, S., Lim, T., and Kim, D. (2009). Integrating radial basis function networks with case-based reasoning for product design. *Expert Systems with Applications*, 36(3, Part 1) :5695–5701.

[Lalitha and Sreeja, 2020] Lalitha, T. B. and Sreeja, P. S. (2020). Personalised self-directed learning recommendation system. *Procedia Computer Science*, 171 :583–592. Third International Conference on Computing and Network Communications (CoCoNet'19).

[Leikola et al., 2018] Leikola, M., Sauer, C., Rintala, L., Aromaa, J., and Lundström, M. (2018). Assessing the similarity of cyanide-free gold leaching processes : A case-based reasoning application. *Minerals*, 8(10).

[Lepage et al., 2020] Lepage, Y., Lieber, J., Mornard, I., Nauer, E., Romary, J., and Sies, R. (2020). The french correction : When retrieval is harder to specify than adaptation. In Watson, I. and Weber, R., editors, *Case-Based Reasoning Research and Development*, pages 309–324, Cham. Springer International Publishing.

[Lin, 2022] Lin, B. (2022). Evolutionary multi-armed bandits with genetic thompson sampling. In *2022 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8.

[Maher and Grace, 2017] Maher, M. L. and Grace, K. (2017). Encouraging curiosity in case-based reasoning and recommender systems. In Aha, D. W. and Lieber, J., editors, *Case-Based Reasoning Research and Development*, pages 3–15, Cham. Springer International Publishing.

[Muangprathub et al., 2020] Muangprathub, J., Boonjing, V., and Chamnongthai, K. (2020). Learning recommendation with formal concept analysis for intelligent tutoring system. *Heliyon*, 6(10) :e05227.

[Müller and Bergmann, 2015] Müller, G. and Bergmann, R. (2015). Cookingcake : A framework for the adaptation of cooking recipes represented as workflows. In *International Conference on Case-Based Reasoning*.

[Nkambou et al., 2010] Nkambou, R., Bourdeau, J., and Mizoguchi, R. (2010). *Advances in Intelligent Tutoring Systems*. Springer Berlin, Heidelberg, 1 edition.

[Obeid et al., 2022] Obeid, C., Lahoud, C., Khoury, H. E., and Champin, P. (2022). A novel hybrid recommender system approach for student academic advising named cohrs, supported by case-based reasoning and ontology. *Computer Science and Information Systems*, 19(2) :979–1005.

[Ontañón et al., 2015] Ontañón, S., Plaza, E., and Zhu, J. (2015). Argument-based case revision in cbr for story generation. In Hüllermeier, E. and Minor, M., editors, *Case-Based Reasoning Research and Development*, pages 290–305, Cham. Springer International Publishing.

[Petrovic et al., 2016] Petrovic, S., Khussainova, G., and Jagannathan, R. (2016). Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning. *Artificial Intelligence in Medicine*, 68 :17–28.

[Richter and Weber, 2013] Richter, M. and Weber, R. (2013). *Case-Based Reasoning (A Textbook)*. Springer-Verlag GmbH.

[Richter, 2009] Richter, M. M. (2009). The search for knowledge, contexts, and case-based reasoning. *Engineering Applications of Artificial Intelligence*, 22(1) :3–9.

[Robertson and Watson, 2014] Robertson, G. and Watson, I. D. (2014). A review of real-time strategy game ai. *AI Mag.*, 35 :75–104.

[Roldan Reyes et al., 2015] Roldan Reyes, E., Negny, S., Cortes Robles, G., and Le Lann, J. (2015). Improvement of online adaptation knowledge acquisition and reuse in case-based reasoning : Application to process engineering design. *Engineering Applications of Artificial Intelligence*, 41 :1–16.

[Seznec et al., 2020] Seznec, J., Menard, P., Lazaric, A., and Valko, M. (2020). A single algorithm for both restless and rested rotting bandits. In Chiappa, S. and Calandra, R., editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 3784–3794. PMLR.

[Sinaga and Yang, 2020] Sinaga, K. P. and Yang, M.-S. (2020). Unsupervised k-means clustering algorithm. *IEEE Access*, 8 :80716–80727.

[Smyth and Cunningham, 2018] Smyth, B. and Cunningham, P. (2018). An analysis of case representations for marathon race prediction and planning. In Cox, M. T., Funk, P., and Begum, S., editors, *Case-Based Reasoning Research and Development*, pages 369–384, Cham. Springer International Publishing.

[Smyth and Willemsen, 2020] Smyth, B. and Willemsen, M. C. (2020). Predicting the personal-best times of speed skaters using case-based reasoning. In Watson, I. and Weber, R., editors, *Case-Based Reasoning Research and Development*, pages 112–126, Cham. Springer International Publishing.

[Soto-Forero et al., 2024] Soto-Forero, D., Ackermann, S., Betbeder, M.-L., and Henriet, J. (2024). The intelligent tutoring system ai-vt with case-based reasoning and real time recommender models. In Recio-Garcia, J. A., Orozco-del Castillo, M. G., and Bridge, D., editors, *Case-Based Reasoning Research and Development*, pages 191–205, Cham. Springer Nature Switzerland.

[Su et al., 2022] Su, Y., Cheng, Z., Wu, J., Dong, Y., Huang, Z., Wu, L., Chen, E., Wang, S., and Xie, F. (2022). Graph-based cognitive diagnosis for intelligent tutoring systems. *Knowledge-Based Systems*, 253 :109547.

[Supic, 2018] Supic, H. (2018). Case-based reasoning model for personalized learning path recommendation in example-based learning activities. In *2018 IEEE 27th International Conference on Enabling Technologies : Infrastructure for Collaborative Enterprises (WETICE)*, pages 175–178.

[Wang et al., 2021] Wang, F., Liao, F., Li, Y., and Wang, H. (2021). A new prediction strategy for dynamic multi-objective optimization using gaussian mixture model. *Information Sciences*, 580 :331–351.

[Xu et al., 2021] Xu, S., Cai, W., Xia, H., Liu, B., and Xu, J. (2021). Dynamic metric accelerated method for fuzzy clustering. *IEEE Access*, 9 :166838–166854.

[Zhang. and Aslan, 2021] Zhang., K. and Aslan, A. B. (2021). Ai technologies for education : Recent research and future directions. *Computers and Education : Artificial Intelligence*, 2 :100025.

[Zhao et al., 2023] Zhao, L.-T., Wang, D.-S., Liang, F.-Y., and Chen, J. (2023). A recommendation system for effective learning strategies : An integrated approach using context-dependent dea. *Expert Systems with Applications*, 211 :118535.

[Zhou and Wang, 2021] Zhou, L. and Wang, C. (2021). Research on recommendation of personalized exercises in english learning based on data mining. *Scientific Programming*, 2021 :5042286.

[Zuluaga et al., 2022] Zuluaga, C. A., Aristizábal, L. M., Rúa, S., Franco, D. A., Osorio, D. A., and Vásquez, R. E. (2022). Development of a modular software architecture for underwater vehicles using systems engineering. *Journal of Marine Science and Engineering*, 10(4).

Document réalisé avec LaTeX et :
le style LaTeX pour Thèse de Doctorat créé par S. Galland — https://github.com/gallandarakhneorg/tex-templates
la collection de paquets `tex-upmethodology` — http://www.arakhne.org/tex-upmethodology/

école doctorale **sciences pour l'ingénieur et microtechniques**

Titre : Adaptation en temps réel d'une séance d'entraînement par intelligence artificielle

Mots-clés : Mot-clé 1, Mot-clé 2

Résumé :

L'objectif de ce travail de thèse est la prise en compte en temps réel du travail d'un apprenant pour une meilleure personnalisation d'AI-VT (Artificial Intelligence Virtual Trainer).

Title: Adaptation en temps réel d'une séance d'entraînement par intelligence artificielle

Keywords: Keyword 1, Keyword 2

Abstract:

This is the abstract in English.

Université Bourgogne-Franche-Comté
32, avenue de l'Observatoire
25000 Besançon, France