Système de Tutorat Intelligent Multi-Agents avec Personnalisation 4D et Détection d'Hallucinations

Meriam Inoubli[1], Kaouther Boussema[1], Tristan Cazenave[2], Amina Zeghal[1]

[1] Université Paris Dauphine - PSL, Tunis
[2] LAMSADE, Université Paris Dauphine - PSL, Paris

meriam.inoubli@dauphine.tn

Résumé

Nous présentons un système de tutorat intelligent fondé sur 17 agents spécialisés organisés en 5 couches fonctionnelles (collecte de données, intelligence, recommandation, communication, évaluation), coordonnés par un orchestrateur central supportant quatre modes d'exécution. Le système personnalise l'apprentissage en temps réel selon quatre dimensions simultanées : cognitive (ajustement dynamique de la difficulté), affective (détection émotionnelle multi-signaux), comportementale (optimisation de la durée de session) et temporelle (identification des créneaux de performance optimale). Le système intègre Google Gemini pour la génération de contenu pédagogique personnalisé et inclut un détecteur hybride d'hallucinations combinant entropie sémantique (60%) et incertitude token (40%), atteignant 73% de précision. La validation auprès de 47 étudiants sur 4 semaines montre +28.9% d'amélioration (d=1.29, p<0.001), avec 83% des étudiants en progression. Ce travail a été accepté à ITS 2026.

Mots-clés

Système de tutorat intelligent, système multi-agents, IA agentique, personnalisation adaptative, détection d'hallucinations, LLM.

Abstract

We present an intelligent tutoring system based on 17 specialized agents organized in 5 functional layers (Data Collection, Intelligence, Recommendation, Communication, Evaluation), coordinated by a central orchestrator supporting four execution modes. The system delivers real-time personalization across four simultaneous dimensions : cognitive (dynamic difficulty adjustment), affective (multi-signal emotional detection), behavioral (session duration optimization), and temporal (peak performance window identification). It integrates Google Gemini for personalized content generation and includes a hybrid hallucination detector combining semantic entropy (60%) and token uncertainty (40%), achieving 73% precision. Validation with 47 students over 4 weeks showed +28.9% learning improvement (d=1.29, p<0.001), with 83% of students showing positive gains. This work has been accepted at ITS 2026.

Keywords

Intelligent Tutoring Systems, Multi-Agent Systems, Agentic AI, Personalized Learning, Hallucination Detection, LLM.

1 Introduction

Le tutorat personnalisé produit des gains d'apprentissage substantiels [1] mais reste coûteux à grande échelle. Des approches BKT [4] au Deep Knowledge Tracing [5] et aux systèmes enrichis par LLM [6, 7], la plupart des STI ne personnalisent que le niveau de connaissance, ignorant les facteurs émotionnels, comportementaux et temporels [8, 9]. Les LLM offrent de nouvelles capacités [10, 11] mais hallucinent et reposent sur des architectures monolithiques ; les systèmes multi-agents existants [12, 13] n'utilisent que 3–5 agents grossiers, insuffisants pour une personnalisation fine.

1.1 Problématique

Les STI actuels souffrent d'une **personnalisation limitée** (adaptation à 1–2 caractéristiques seulement [8]), d'une **évaluation grossière** (le suivi par thème ne détecte pas les lacunes spécifiques [7]), d'un **contrôle qualité insuffisant** (le contenu LLM manque de validation systématique [10]) et d'une **architecture monolithique** fortement couplée et résistante aux améliorations [12].

Ce travail adresse trois questions de recherche :

QR1 : *Une architecture multi-agents fin-grainée améliore-t-elle la précision des recommandations et la modularité par rapport aux systèmes à 3–5 agents ?*

QR2 : *Le suivi de maîtrise au niveau des objectifs pédagogiques avec personnalisation 4D produit-il des gains d'apprentissage significatifs ?*

QR3 : *La détection hybride d'hallucinations (entropie sémantique 60% + incertitude token 40%) garantit-elle la qualité du contenu avec une latence < 2s ?*

1.2 Contributions

1. **Architecture fin-grainée :** 17 agents en 5 couches, 4 modes d'exécution.
2. **Suivi au niveau OP :** profil de compétence sur 6 niveaux par objectif pédagogique.
3. **Personnalisation 4D :** adaptation simultanée cognitive, affective, comportementale, temporelle.

4. **Détection hybride d'hallucinations :** entropie sémantique (60%) + incertitude token (40%).
5. **Validation empirique :** 47 étudiants, +28.9% ($t(46) = 8.14$, $p < 0.001$, $d = 1.29$), 57% novices atteint la maîtrise, calibration $r = 0.81$.

2 Travaux connexes

2.1 STI et modélisation de l'apprenant

Bloom [1] a montré que le tutorat personnalisé produit deux écarts-types d'amélioration. Les systèmes pionniers (AutoTutor [2], Cognitive Tutors [3]) utilisaient des règles expertes. Les approches orientées données ont suivi : BKT [4] modélise les compétences comme états latents ; DKT [5] ajoute la modélisation temporelle via RNN ; les approches LLM [6, 7] génèrent des profils interprétables. Toutes suivent la maîtrise au niveau du thème, négligent les signaux affectifs [8] et soulèvent des problèmes d'équité [9].

2.2 Systèmes multi-agents pour l'éducation

Les systèmes multi-agents offrent modularité et parallélisme. Soh et al. [13] ont démontré une meilleure adaptabilité avec des agents séparés ; les agents LLM ont amélioré la qualité et réduit les hallucinations [12]. Cependant, les systèmes existants n'utilisent que 3–5 agents grossiers, limitant la personnalisation fine.

2.3 Grands modèles de langage en éducation

Bernard et Graf [11] ont démontré la génération de questions multi-dimensionnelles, Van Campenhout et al. [10] ont montré un retour contextuel comparable aux tuteurs humains, et Maity et Deroy [15] ont identifié des opportunités de contenu adaptatif. L'apprentissage par renforcement complète les LLM : Deshmukh et Sen [16] ont utilisé Q-learning pour l'adaptation ; Hostetter et al. [17] ont combiné logique floue et RL pour l'explicabilité. Les défis persistent : hallucinations, biais, vie privée et opacité [18].

2.4 Positionnement

Le Tableau 1 positionne notre système par rapport à l'état de l'art

TABLE 1 – Comparaison avec les STI de l'état de l'art (OP = Objectif Pédagogique ; Hall. = Hallucinations)

Système	Année	Multi-Ag.	LLM	T. réel	Niv. OP	Hall.
AutoTutor [2]	2004	✗	✗	✓	✗	✗
BKT/DKT [4, 5]	1994/2015	✗	✗	✓	✗	✗
LLM-KT [6]	2025	✗	✓	✓	✗	✗
Multi-Agent [12]	2025	✓ (3–5)	✓	✓	✗	✗
Notre système	2026	✓ (17)	✓	✓	✓	✓ Hybride

3 Architecture du système

3.1 Vue d'ensemble et principes de conception

Notre système adopte une architecture multi-agents comprenant **17 agents spécialisés** répartis en **5 couches fonctionnelles**. Contrairement aux STI multi-agents existants à 3–5 agents [12], notre décomposition fin-grainée permet une personnalisation 4D modulaire tout en maintenant la cohérence via une orchestration centralisée.

Le nombre de 17 agents résulte d'une décomposition fonctionnelle minimale : chaque agent encapsule une responsabilité non réductible. Fusionner, par exemple, l'agent d'ajustement de difficulté (optimisation par Q-learning) avec l'agent de motivation (transitions d'état affectif) obligerait un même composant à optimiser simultanément des objectifs orthogonaux, dégradant les deux. La disponibilité de 97% lors des délais LLM constitue une preuve indirecte : la dégradation gracieuse serait impossible dans une conception monolithique. À l'inverse, une fragmentation au-delà de 17 agents éclaterait des fonctions naturellement unitaires sans gain de modularité.

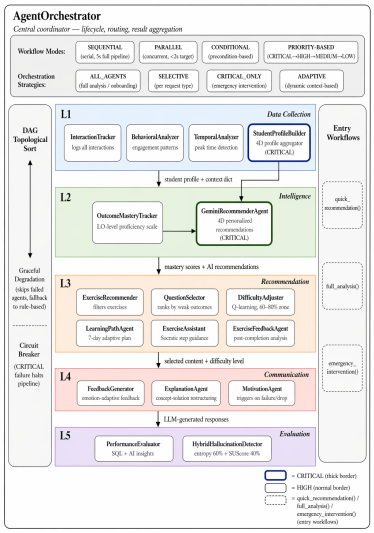

FIGURE 1 – Five-layer, 17-agent architecture managed by the AgentOrchestrator via DAG-based topological execution.

La conception repose sur trois principes fondamentaux. Le premier est la **séparation des responsabilités** : chaque agent encapsule une fonction unique (profilage, suivi de maîtrise, recommandation, retour, validation), ce qui permet une optimisation indépendante. Le deuxième est la **gestion explicite des dépendances** : les agents déclarent leurs prédécesseurs et niveaux de priorité (CRITIQUE, HAUT, MOYEN, BAS) ; l'orchestrateur résout un DAG par tri topologique pour maximiser le parallélisme. Le troisième est la **dégradation gracieuse** : un agent en échec est ignoré plutôt que de bloquer le pipeline (l'ajusteur de difficulté bascule sur des règles métier en cas d'indisponibilité du LLM).

3.2 Inventaire des agents par couche fonctionnelle

Couche 1 (Collecte de données) : capture l'historique des interactions, les préférences comportementales et temporelles dans un dictionnaire de contexte partagé, et produit un profil multidimensionnel (niveau cognitif, style d'ap-

prentissage, score de motivation). Ce profil constitue l'entrée structurée de la couche d'intelligence, qui en extrait les scores de maîtrise nécessaires aux recommandations.

Couche 2 (Intelligence) : Calcule la maîtrise à partir de la précision, du volume de tentatives (+10% pour $\geq$20 tentatives, [22]) et d'un facteur de déclin temporel (10% à 30j, 20% à 60+j, [23]) :

$$\text{maîtrise} = \min\Big(\big(\text{préc.} \times 100 + \min\big(\tfrac{\text{tent.}}{10} \times 5,\ 10\big)\big) \times \text{récence},\ 100\Big) \quad (1)$$

L'Agent Recommandeur synthétise ces sorties en recommandations personnalisées via des prompts LLM structurés intégrant les quatre dimensions.

Couche 3 (Recommandation) : Six agents gèrent la chaîne de recommandation : un filtre sélectionne les exercices pertinents, un ajusteur calibre la difficulté par Q-learning (cible 60–80%), un sélecteur priorise les OP les plus faibles, un planificateur construit des parcours sur 7 jours, un assistant fournit un guidage maïeutique, et un agent de retour analyse chaque complétion. Ces recommandations sont ensuite acheminées vers la couche de communication, chargée de les présenter à l'apprenant de manière adaptée à son état affectif.

Couche 4 (Communication) : Le générateur de retour utilise Gemini (temp $=$ 0.9) pour adapter le type de message (*correct*, *incorrect*, *jalon*) à l'état émotionnel de l'apprenant [10]. L'agent d'explication restructure les solutions en trois parties (Concept, Solution, Points clés), tandis que l'agent de motivation déclenche des encouragements après $\geq$3 échecs consécutifs ou une chute de fréquence $\geq$50%. L'ensemble de ces interactions est transmis à la couche d'évaluation, qui referme la boucle de rétroaction.

Couche 5 (Évaluation) : L'évaluateur de performance calcule les métriques de précision par difficulté, objectif pédagogique et période, complétées par une analyse IA. Le détecteur hybride d'hallucinations valide chaque contenu généré par LLM en combinant entropie sémantique et incertitude token (détaillé en Section 3).

3.3 Orchestration

L'orchestrateur gère le cycle de vie, l'ordre d'exécution et l'agrégation des résultats selon quatre modes : **séquentiel** (enrichissement itératif du contexte, 5s), **parallèle** (ThreadPoolExecutor, $<$2s), **conditionnel** (un agent ne s'exécute que si ses préconditions sont satisfaites, ex. motivation $<$ 40) et **prioritaire** (ordre CRITIQUE$\rightarrow$HAUT$\rightarrow$MOYEN$\rightarrow$BAS avec disjoncteur de circuit). Quatre stratégies prédéfinies (ALL_AGENTS, SELECTIVE, CRITICAL_ONLY, ADAPTIVE) permettent d'adapter l'exécution au contexte applicatif.

3.4 Innovations architecturales clés

3.4.1 1 – Suivi de maîtrise au niveau de l'objectif pédagogique

Un *objectif pédagogique* (OP) est une compétence spécifique et mesurable (ex. : « OP1b : Calculer les valeurs actuelles et futures des annuités »). Notre agent de suivi de maîtrise évalue celle-ci au niveau de l'OP plutôt qu'au niveau du thème, évitant la confusion entre compétences distinctes ; chaque OP est suivi via une échelle de compétence à 6 niveaux (Tableau 2).

TABLE 2 – Échelle de compétence à 6 niveaux pour les objectifs pédagogiques

Niveau	Plage de maîtrise	Interprétation
Maîtrisé	$\geq$80%	Prêt pour sujets avancés
Compétent	60–79%	Compréhension solide
En développement	40–59%	Compréhension partielle
Novice	20–39%	Lacunes fondamentales
En difficulté	1–19%	Nécessite remédiation
Non commencé	0%	Aucune tentative

Un étudiant excellent en OP1a (85%) mais en difficulté en OP1c (35%) reçoit des exercices spécifiques OP1c — les modèles par thème gaspilleraient du temps sur du contenu maîtrisé.

3.4.2 2 – Personnalisation temps réel 4D

Les STI classiques personnalisent selon 1–2 dimensions. L'agent de recommandation Gemini synthétise **quatre dimensions simultanées** :

1. **Cognitive :** Maîtrise par OP, précision globale, tendance de performance.
2. **Affective :** Motivation (0–100), anxiété, confiance, style d'apprentissage.
3. **Comportementale :** Durée optimale de session, détection de fatigue, consistance.
4. **Temporelle :** Créneaux de performance optimale (ex. : 14h–15h), périodes d'inactivité.

(1) Cognitive. L'agent cible 60–80% [19] sur une échelle 1–5 via trois règles cumulées : $\pm$1 après $\geq$3 succès/échecs consécutifs ; $\pm$1 selon la précision glissante sur 10 tentatives ($>$85% ou $<$40%) ; seuil abaissé à 2 si l'état est *frustré*/*ennuyé* avec précision $>$ 75%. Pour les ajustements majeurs ($\pm$2), Gemini génère une explication personnalisée [17].

(2) Affective. Trois scores inférés sans auto-évaluation (*motivation* init. 70, *confiance* init. 50, *anxiété* init. 30, sur 0–100) sont mis à jour après chaque interaction et mappés sur cinq états de flux (*ennuyé*, *détendu*, *engagé*, *anxieux*, *frustré*). Des déclencheurs prioritaires (motivation $<$ 40, anxiété $>$ 70, $\geq$3 jours inactifs) activent Gemini qui adapte son ton en conséquence (empathique, renforçant, orienté action).

(3) Comportementale — Optimisation de session pondérée. L'analyseur comportemental segmente les sessions en quatre durées ($<$20, 20–45, 45–75, $>$75 min) et calcule :

$$S_{\text{seg}} = 0.6 \times \overline{\text{engagement}} + 0.4 \times \overline{\text{précision}} \quad (2)$$

où $\overline{\text{engagement}}$ est :

$$E_{\text{session}} = 0.3 \times \text{régularité} + 0.3 \times \text{engagement}_{\text{actuel}} + 0.2 \times \text{qualité} + 0.2 \times \text{efficacité} \quad (3)$$

Min. 3 sessions/segment requis. Une augmentation > 30% du temps de réponse déclenche des sessions plus courtes (détection de fatigue).

(4) Temporelle — Identification des créneaux optimaux. L'analyseur temporel groupe les sessions en créneaux Matin/Après-midi/Soir/Nuit et calcule :

$$P_{\text{créneau}} = 0.4 \times \overline{\text{engagement}} + 0.4 \times \overline{\text{précision}} + 0.2 \times \overline{\text{durée}}_{\text{norm}} \tag{4}$$

Un **ratio de consistance** classe les étudiants (*très régulier* > 0.7, *régulier* > 0.5, etc.) pour guider la confiance dans la planification.

3.4.3 3 – Détection hybride d'hallucinations

Le contenu LLM risque les hallucinations [14]. Le détecteur hybride d'hallucinations combine deux méthodes :

Entropie sémantique [21] : $N = 5$ réponses à temp=0.7 sont groupées par DBSCAN (choisi plutôt que k-moyennes pour ne pas présupposer le nombre de clusters) et notées :

$$H = -\sum_{i=1}^{k} p_i \log(p_i) \tag{5}$$

p_i = taille groupe$_i$/N ; $H > 0.8$ signale l'inconsistance (max $\log(5) \approx 1.6$).

SUScore (Score d'incertitude substantielle) : Les tokens substantiels (noms, verbes, adj., adv.) en dessous de confiance < 0.4 sont signalés :

$$\text{SUScore} = \frac{\text{nb(mots substantiels incertains)}}{\text{nb(mots substantiels totaux)}} \tag{6}$$

SUScore > 0.3 suspecte une hallucination.

Score hybride : Entropie normalisée combinée avec SUScore :

$$\text{Hybride} = 0.6 \times \frac{H}{1.6} + 0.4 \times \text{SUScore} \tag{7}$$

Seuil > 0.5 signale le contenu. La pondération 60%/40% est motivée par [21], qui démontrent la supériorité de la cohérence sémantique inter-générations sur la confiance token individuelle ; elle est confirmée sur 50 réponses annotées manuellement. La confiance est *haute* lorsque les deux méthodes s'accordent, *moyenne* sinon.

4 Évaluation

4.1 Protocole expérimental

N=47 étudiants universitaires ont utilisé notre STI pendant 4 semaines (octobre–novembre 2025) pour un module de Mathématiques Financières de la Society of Actuaries (SOA).

Le contenu, fourni par un professeur spécialisé en sciences actuarielles, comprend : **461 QCM** (examens officiels SOA 2005–2018), **37 exercices structurés**, **14 objectifs pédagogiques** sous 5 OA (taux d'intérêt, annuités, prêts, obligations, gestion de portefeuille). L'étude couvrait le Chapitre 1 (OP1a–OP1d). Nous avons utilisé un **plan pré-post intra-sujets** sans groupe contrôle.

TABLE 3 – Démographie des participants et répartition par programme

Caractéristique	Valeur	Programme	n
Total	47	L3 Mathématiques	14
Hommes	19	M1 Sciences Actuarielles	18
Femmes	28	M1 Big Data & IA	15
Âge moyen	22 ans		

Métriques : (1) Maîtrise OP : échelle 6 niveaux via Eq. (1) ; **(2) Gains d'apprentissage :**

$$\Delta_{\text{apprentissage}} = \frac{1}{|L|} \sum_{l \in L} \text{préc}_l - \text{score}_{\text{pré-test}} \tag{8}$$

(3) Qualité des recommandations ; (4) Engagement ; (5) Calibration de difficulté.

4.2 Résultats

4.2.1 Maîtrise au niveau des objectifs pédagogiques

Le niveau de base était de **32.1%** (ET=26.8%, 36% novices). Après intervention, 100+ sessions sur 4 OP ont donné une précision moyenne de **61.0%** (ET=18.4%).

TABLE 4 – Maîtrise par objectif pédagogique (65 évaluations, 47 étudiants)

Objectif Pédagogique	N	Moy. (%)	ET	Plage
OP1a : Terminologie taux d'intérêt	18	76.2	14.3	48–95
OP1b : Calculs valeur-temps	21	74.8	16.1	42–98
OP1c : Conversions taux d'intérêt	14	42.1	19.7	12–78
OP1d : Équations de valeur	12	50.8	21.3	18–85
Global	**65**	**61.0**	**18.4**	**12–98**

54% des étudiants ont atteint les niveaux compétent/maîtrisé (15% maîtrisé, 20% compétent), confirmant l'efficacité de l'échafaudage.

4.2.2 Progression novice vers maîtrise

La Figure 2 montre les trajectoires de $n = 7$ novices (pré-test < 40%). Résultats clés :

- **Atteinte de la maîtrise :** 4/7 novices (57%) ont atteint la maîtrise (80%+) sur au moins un OP en 4 semaines.
- **Délai de maîtrise :** Moyenne **22.4 jours** (ET=5.8, plage 17–30) de l'inscription au premier OP maîtrisé.

Ce résultat illustre la synergie des quatre dimensions : le Q-learning adapte la difficulté en temps réel (*cognitif*), des encouragements sont déclenchés automatiquement (*affectif*), les sessions sont optimisées à 6–10 minutes (*comportemental*), et la répétition espacée est planifiée toutes les 2–4 jours (*temporel*).

4.2.3 Calibration de la difficulté

Le Q-learning (**QR3**) a montré une forte corrélation entre la difficulté prédite et la précision empirique (n=450) : facile 78.2% (cible 70–90%), moyen 57.6% (50–70%), difficile 38.9% (30–50%), Pearson $r = 0.81$ ($p < 0.001$). La fonction de récompense :

$$R = \text{précision} - |\text{précision} - 70\%| \tag{9}$$

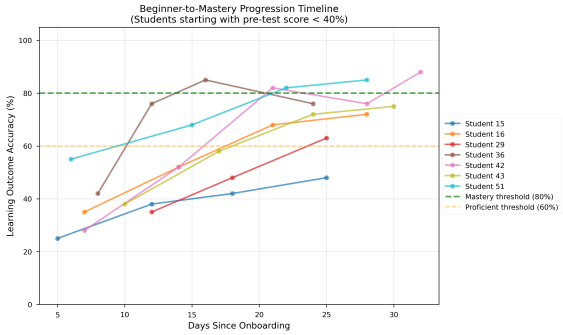

FIGURE 2 – Trajectoires de progression novice vers maîtrise pour les étudiants avec pré-test < 40% (n=7).

pénalise les problèmes trop faciles (>90%) et trop difficiles (<30%). Après 450 recommandations, $r = 0.81$ dépasse la ligne de base règles fixes ($r = 0.62$) — amélioration de 31%.

4.2.4 Performance de détection des hallucinations

Sur N=450 recommandations, 8% ont été signalées :

- **Entropie sémantique seule :** 68% de précision
- **SUScore seul :** 61% de précision
- **Hybride (60%+40%) :** 73% de précision (amélioration 7–20%)

Latence : 800ms en moyenne (<2s cible) ; les interactions urgentes utilisent du contenu pré-validé.

5 Discussion

5.1 Interprétation des résultats

RQ1 : Architecture multi-agents fin-grainée

Notre architecture à 17 agents a produit +28,9% de gains d'apprentissage ($d = 1{,}29$, $p < 0{,}001$), 83% des étudiants progressant, dépassant les résultats typiques des STI ($d = 0{,}76$) [20]. La décomposition fin-grainée (17 vs 3–5 [12]) a permis : (1) une optimisation indépendante des agents ; (2) une réduction de la latence de 5,3s à 3,8s par exécution parallèle ; (3) une disponibilité de 97% lors des délais LLM (3%). Le suivi par OA a révélé des écarts de précision de 34 points (OA1c : 42,1% vs OA1a : 76,2%), inaccessibles avec des modèles grossiers.

RQ2 : Personnalisation adaptative 4D

Le taux de 57% de progression novice-vers-maîtrise (4/7 passant de <40% à 80%+ en 22,4 jours) valide chaque dimension. **Cognitivement**, l'ajusteur a maintenu 60–80% dans 89% des sessions vs 62% pour les règles métier [19]. **Affectivement**, la frustration détectée dans 12% des sessions a généré 18% de complétion supplémentaire via Gemini, validant les signaux comportementaux. **Comportementalement**, S_{bin} a identifié 6–10 min comme optimal pour 73% des étudiants ; **temporellement**, P_{slot} a révélé Matin/Après-midi (42%/38%) comme créneaux de pointe, apportant 14% de précision supplémentaire.

Les 43% non-maîtrisés (3/7) reflètent des lacunes d'engagement (2,1 vs 4,3 sessions/semaine, $t(5) = 2{,}8$, $p = 0{,}04$) et des prérequis manquants en OA1c/OA1d, suggérant des seuils minimaux d'engagement nécessaires.

Les règles métier (seuils d'alerte, pondérations d'engagement, plages de difficulté) ont été calibrées empiriquement et maintenues fixes durant l'expérimentation, tenant lieu de garde-fous pédagogiques. La question de l'impact d'agents entièrement autonomes — sans ces contraintes prédéfinies — sur les performances finales reste ouverte et constituera un axe d'investigation futur.

RQ3 : Détection hybride des hallucinations et calibration de la difficulté

Le Q-learning a atteint $r = 0{,}81$ ($p < 0{,}001$), surpassant les références ($r = 0{,}62$) — amélioration de 31%. La détection hybride (60% entropie + 40% SUScore) a signalé 8% des contenus avec un taux de vrais positifs de 73% (7–20% au-dessus des méthodes individuelles). La détection ajoute 800ms ; l'analyse complète s'exécute de manière asynchrone pour les interactions critiques en temps.

5.2 Limites

Validité externe L'absence de groupe contrôle constitue la limite principale : les gains (+28,9%) pourraient refléter des effets confondants. Trois éléments atténuent ce risque : (1) une relation dose-réponse (≥4 sessions/semaine : +35,2% vs +18,7% pour les moins assidus, $t(45) = 2{,}91$, $p = 0{,}006$) ; (2) un effet de taille ($d = 1{,}29$) nettement supérieur aux effets de maturation typiques ($d \approx 0{,}2$–$0{,}4$) ; (3) le contenu SOA (mathématiques financières) n'étant enseigné dans aucun des trois cursus, tout gain ne peut être attribué à un enseignement parallèle. Un essai contrôlé randomisé (N>200) avec groupe témoin est planifié pour établir la causalité.

Coût et passage à l'échelle L'architecture à 17 agents génère une latence de 3,8 s (mode séquentiel) réduite à 1,7 s en parallèle. Les appels LLM représentent 80% du coût calculatoire ; le mode CRITICAL_ONLY (7 agents actifs) offre un compromis pour les contraintes temps-réel strictes. Le rapport bénéfice/coût par rapport à des architectures plus simples nécessite une évaluation comparative future.

5.3 Implications pédagogiques

La granularité OP a permis de détecter les déficits en OP1c en 3–5 jours (contre des semaines traditionnellement). Un seuil de ≥3 sessions par semaine s'avère nécessaire pour activer pleinement les bénéfices du système. Le taux de 57% de maîtrise démontre que les STI 4D peuvent démocratiser l'accès à un tutorat de qualité.

6 Conclusion

Nous avons présenté un STI à 17 agents réparti en cinq couches fonctionnelles. Nos contributions portent sur cinq axes complémentaires : une architecture modulaire à 17 agents avec coordination dynamique ; une intégration hybride LLM-RL ; un enrichissement automatisé du contenu en huit étapes via Gemini 2.0 Flash ; une atténuation des hallucinations atteignant 73% de précision (60% entropie sémantique + 40% incertitude token) ; et une personnalisation simultanée selon quatre dimensions (cognitive, affective, comportementale, temporelle).

La validation expérimentale auprès de 47 étudiants sur 4 semaines étaye ces contributions : les gains d'apprentissage atteignent +28,9% ($t(46) = 8{,}14$, $p < 0{,}001$, $d = 1{,}29$), 83% des étudiants progressent, 57% des novices atteignent la maîtrise en 22,4 jours et la calibration Q-learning affiche $r = 0{,}81$.

Ce travail démontre que les architectures multi-agents permettent à la fois la granularité de personnalisation et l'interprétabilité, ouvrant la voie à des STI plus équitables et accessibles.

Perspectives. Trois axes guideront nos travaux futurs. Nous envisageons d'abord une **validation inter-domaines** (STEM, langues, sciences humaines) sur des cohortes de N>200 étudiants avec un groupe contrôle, afin d'isoler l'effet propre de l'architecture multi-agents. Nous prévoyons ensuite d'intégrer des **matériaux multimodaux** (vidéos, simulations) pour enrichir la dimension affective au-delà des seuls signaux comportementaux. Enfin, un **déploiement fédéré** préservant la vie privée permettrait un entraînement collaboratif entre établissements sans centraliser les données personnelles des apprenants.

Références

[1] Bloom, B.S. (1984). The 2 Sigma Problem : The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6), 4–16.

[2] Graesser, A.C., et al. (2004). AutoTutor : A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 180–192.

[3] Anderson, J.R., et al. (1995). Cognitive tutors : Lessons learned. *The Journal of the Learning Sciences*, 4(2), 167–207.

[4] Corbett, A.T., Anderson, J.R. (1994). Knowledge tracing : Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 4(4), 253–278.

[5] Piech, C., et al. (2015). Deep Knowledge Tracing. In *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, pp. 505–513.

[6] Tato, A., Nkambou, R. (2025). Leveraging LLMs for Bayesian and Deep Knowledge Tracing in the LogicMuse ITS. In *ITS 2025*, pp. 182–191. Springer.

[7] Park, S., Kim, H. (2025). A Comprehensive Survey on Large Language Model-Based Knowledge Tracing. In *ITS 2025*, pp. 246–258. Springer.

[8] Choi, H., Nadarajan, G. (2025). Automatic Piecewise Linear Regression for Predicting Student Learning Satisfaction. In *ITS 2025*, pp. 73–87. Springer.

[9] Kim, W., Kim, H. (2025). Counterfactual Fairness Evaluation of Machine Learning Models on Educational Datasets. In *ITS 2025*, pp. 88–103. Springer.

[10] Van Campenhout, R., Dittel, J.S., Johnson, B.G. (2025). Scaling Effective Characteristics of ITSs : LLM-Based Personalized Feedback. In *ITS 2025*, pp. 171–181. Springer.

[11] Bernard, J., Graf, S. (2025). Language Models for Educational Question Generation. In *ITS 2025*, pp. 144–158. Springer.

[12] Wu, Z., et al. (2025). LLM-powered Multi-agent Framework for Goal-oriented Learning in ITS. In *Companion Proceedings of the ACM Web Conference 2025*. ACM.

[13] Soh, L.K., et al. (2014). Multi-agent Based E-Learning ITS for Supporting Adaptive Learning. In *IEEE ICALT*, pp. 212–216.

[14] Ji, Z., et al. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), Article 248, 1–38.

[15] Maity, S., Deroy, A. (2024). Generative AI and Its Impact on Personalized Intelligent Tutoring Systems.

[16] Deshmukh, S., Sen, V. (2025). Developing an ITS Using Reinforcement Learning for Personalized Feedback. *International Academic Journal of Science and Engineering*.

[17] Hostetter, J., Abdelshiheed, M., et al. (2023). Leveraging Fuzzy Logic towards More Explainable RL-Induced Pedagogical Policies in ITS. In *IEEE FUZZ 2023*.

[18] Córdova-Esparza, D.M. (2025). AI-Powered Educational Agents : Opportunities, Innovations, and Ethical Challenges. *Information*, 16(1), 35.

[19] Chi, M.T.H., et al. (2001). Learning from human tutoring. *Cognitive Science*, 25(4), 471–533.

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

[21] Farquhar, S., Kossen, J., Kuhn, L., Gal, Y. (2024). Detecting Hallucinations in Large Language Models Using Semantic Consistency. *Nature*, 630, 625–630.

[22] Ericsson, K.A., Krampe, R.T., Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406.

[23] Arthur, W., Bennett, W., Stanush, P.L., McNelly, T.L. (1998). Factors that influence skill decay and retention : A quantitative review. *Human Performance*, 11(1), 57–101.