



HAL
open science

Modélisation de l'activité gestuelle et sélection automatique de feedback pour des environnements interactifs d'apprentissage : application à la calligraphie

Rémy Frenoy

► **To cite this version:**

Rémy Frenoy. Modélisation de l'activité gestuelle et sélection automatique de feedback pour des environnements interactifs d'apprentissage : application à la calligraphie. Autre [cs.OH]. Université de Technologie de Compiègne, 2016. Français. NNT : 2016COMP2297 . tel-01557509

HAL Id: tel-01557509

<https://theses.hal.science/tel-01557509>

Submitted on 6 Jul 2017

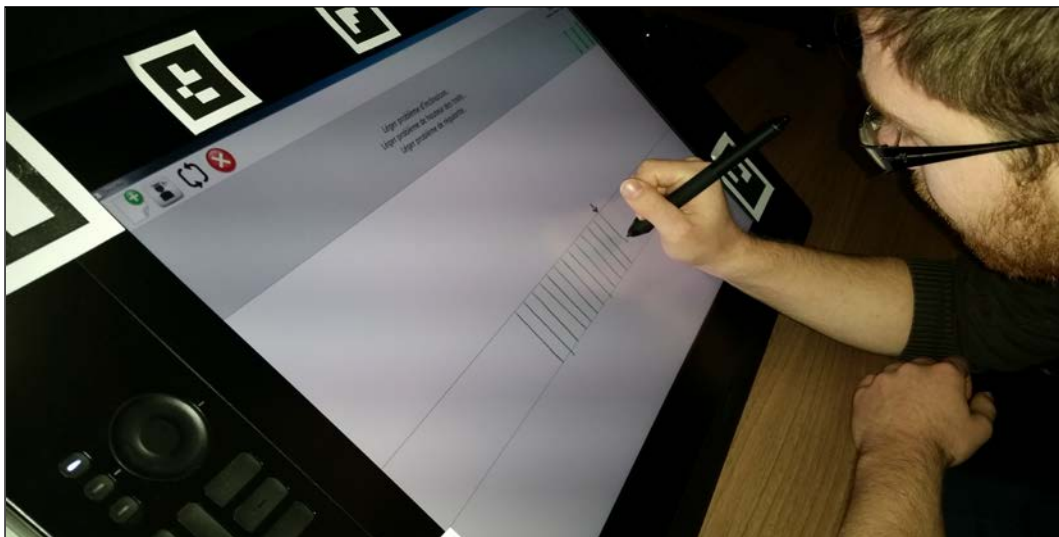
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Par **Rémy FRENOY**

Modélisation de l'activité gestuelle et sélection automatique de feedback pour des environnements interactifs d'apprentissage : application à la calligraphie

Thèse présentée
pour l'obtention du grade
de Docteur de l'UTC



Soutenue le 4 octobre 2016

Spécialité : Informatique : Unité de recherche Heudyasic (UMR-7253)

D2297



UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

**Modélisation de l'activité gestuelle et
sélection automatique de feedback pour
des environnements interactifs
d'apprentissage : application à la
calligraphie**

THÈSE DE DOCTORAT

Spécialité : Informatique

Unité de Recherche Heudyasic – (UMR-7253)

Présentée et soutenue publiquement le 4 octobre 2016 par

RÉMY FRENOY

devant un jury composé de :

Frédéric Bevilacqua	Directeur de recherche, IRCAM/Paris	(Rapporteur)
Pierre De Loor	Professeur des Universités, ENIB	(Rapporteur)
Olivier Gapenne	Professeur des Universités, UTC	(Directeur de thèse)
Yves Grandvalet	Directeur de recherche CNRS, UTC	(Président)
Vanda Luengo	Professeur des Universités, LIP6	(Examineur)
Pierre-Yves Oudeyer	Directeur de recherche INRIA, ENSTA	(Examineur)
Indira Thouvenin	Enseignant-chercheur HDR, UTC	(Directeur de thèse)

Table des matières

Table des matières	1
Table des figures	5
Liste des tableaux	7
1 Introduction	15
1.1 Contexte scientifique	15
1.2 Apprentissage de gestes	16
1.2.1 Approche cognitive	16
1.2.2 Approche dynamique	17
1.2.3 Conclusions	18
1.3 Environnements d'apprentissage de gestes	19
1.3.1 Feedback intrinsèque	19
1.3.2 Feedback extrinsèque	20
1.3.3 Dépendance à l'environnement	20
1.4 Environnements intelligents d'apprentissage	21
1.5 Problématiques	22
1.5.1 Caractérisation	23
1.5.2 Interprétation	23
1.5.3 Représentation	23
1.5.4 Adaptation	23
1.5.5 Choix et positionnement	24
1.6 Organisation du mémoire	25
2 Adaptation basée sur l'expertise : Modélisation de l'interaction	27
2.1 Introduction	27
2.2 Taxonomie des domaines d'apprentissage	28
2.3 Modélisation : domaines bien définis	31
2.3.1 Représentation de l'apprenant	31
2.3.2 Modélisation des capacités	36
2.3.3 Conclusions	43
2.4 Modélisation : domaines mal définis	44
2.4.1 Étude de cas	44
2.4.2 Construction automatique du domaine	45

Table des matières

2.4.3	Aide méta-cognitive	46
2.4.4	Approches constructivistes	47
2.4.5	Approches multi-paradigmes	48
2.4.6	Autres modèles d'apprentissage	49
2.4.7	Conclusions	50
2.5	Modélisation du couple apprenant/environnement d'apprentissage	50
2.5.1	Représentation de l'apprenant	51
2.5.2	Représentation de l'environnement d'apprentissage . . .	51
2.5.3	Relation entre apprenant et environnement	51
2.5.4	Adaptation basée sur une expertise	52
2.5.5	Caractérisation du geste	53
2.6	Conclusions	54
3	Adaptation basée sur l'expertise : Expérimentation	55
3.1	Introduction	55
3.2	Environnement CalliSmart (1)	57
3.2.1	Caractérisation du geste	57
3.2.2	Ensemble d'actions pédagogiques	58
3.3	Adaptation dans CalliSmart	61
3.3.1	Travaux relatifs à la reconnaissance de gestes	62
3.3.2	Intégration dans CalliSmart	67
3.4	Expérimentation	72
3.4.1	Hypothèses	72
3.4.2	Protocole	73
3.4.3	Participants	73
3.4.4	Évaluation	74
3.4.5	Résultats et discussion	74
3.5	Conclusions	78
4	Adaptation automatique : Modélisation de l'interaction	81
4.1	Motivations	81
4.2	Travaux similaires	82
4.2.1	Systèmes de recommandation	82
4.2.2	Approches par renforcement	85
4.2.3	Problème de bandit	87
4.2.4	Conclusions	92
4.3	Adaptation dynamique des règles de sélection	93
4.3.1	Rappels	93
4.3.2	Formalisation du problème	94
4.3.3	Sélection d'action	94
4.4	Conclusions	97
5	Adaptation automatique : Expérimentation	99
5.1	Introduction	99
5.2	Environnement CalliSmart (2)	100
5.2.1	Caractérisation du geste	100

5.2.2	Ensemble d'actions pédagogiques	101
5.2.3	Reconnaissance de gestes	104
5.3	Expérimentation	105
5.3.1	Protocole	105
5.3.2	Résultats et discussion	107
5.4	Conclusions	110
6	Conclusions et perspectives	111
6.1	Résumé et contributions	111
6.1.1	Caractérisation	111
6.1.2	Interprétation	112
6.1.3	Représentation	112
6.1.4	Adaptation	112
6.2	Limites et perspectives	114
6.2.1	Modélisation de l'activité	114
6.2.2	Ensemble des actions pédagogiques	114
6.2.3	Représentation des actions pédagogiques	114
6.2.4	Évolution de la plateforme CalliSmart et des modèles proposés	115
	Bibliographie	117

Table des figures

1.1	Exemples de ductus, modèles détaillés de la trace à produire	19
2.1	Schématisation de la taxonomie des domaines de Lynch et al. (Lynch et al., 2006).	29
2.2	Composantes possibles du modèle de l'apprenant (<i>student model</i>).	31
2.3	Représentation de l'apprenant selon un modèle par recouvrement.	34
2.4	Représentation de l'apprenant selon un modèle combinant les approches de représentation par recouvrement et par erreurs.	34
2.5	Deux règles de production pour la résolution du problème : trouver la valeur d'un des angles d'un triangle isocèle, adapté depuis (Mitrovic et al., 2003; Desmarais and Baker, 2012).	37
2.6	Modèle de Markov caché utilisé dans le cadre du Bayesian knowledge-tracing, adapté depuis (Beck and Sison, 2006)	38
2.7	Logique du tuteur par contrainte en fonction des conditions de pertinence et de satisfaction	41
2.8	Comparaison des règles de knowledge-tracing et des contraintes de tuteurs par contraintes, adapté depuis (Mitrovic et al., 2003; Desmarais and Baker, 2012)	42
3.1	Ensemble des opérations nécessaires à l'adaptation automatique de l'environnement	56
3.2	L'exercice de l'écriture du mot "minimum" allie contraintes morphocinétiques et topocinétiques	57
3.3	Environnement graphique sans feedback supplémentaire	59
3.4	Le feedback CR (connaissance du résultat) propose un score sur 100 sur chacune des dimensions d'erreur considérées	60
3.5	Feedback illustrant une erreur d'inclinaison d'un élément	60
3.6	Feedback illustrant la spatialisation des éléments	61
3.7	Feedback illustrant une erreur sur la taille d'un élément	61
3.8	Comportements schématiques possibles de modèles génératifs et discriminants (Soullard, 2013)	65
3.9	La plateforme CalliSmart, intégrée sur une table Wacom	68

Table des figures

3.10	Diagramme en boîtes illustrant la distance à la performance experte des participants des groupes contrôle (sans feedback) et expert (avec feedback et adaptation basée sur des règles expertes)	75
3.11	Diagramme en boîtes illustrant la distance à la performance experte des participants des groupes contrôle (sans feedback) et expert (avec feedback et adaptation basée sur des règles expertes) lors du post-test	76
3.12	Réponses des participants face à la proposition : “J’ai le sentiment d’avoir progressé”	77
3.13	Réponses des participants du groupe expert face à la proposition : “Le système reconnaît mes erreurs”	77
3.14	Réponses des participants du groupe expert face à la proposition : “Le système me propose des aides adaptées à mes erreurs”	78
5.1	Les pastilles de couleurs informent l’apprenant sur la qualité de chaque tracé	102
5.2	Le feedback “vélocité” indique des discontinuités dans la dynamique du tracé : en rouge les accélérations, en bleu les ralentissements . .	102
5.3	Le feedback “pression” indique par la largeur du trait ainsi qu’un gradient de couleur la pression appliquée par l’outil sur le plan . .	103
5.4	Le feedback “peau d’oignon” propose en grisé l’élément précédemment inscrit	103
5.5	Le feedback cible permet de focaliser vers la suite du geste	103
5.6	Illustration du scénario expérimental pour les deux groupes considérés	106
5.7	Diagramme barre illustrant le nombre moyen de sélection des différents feedbacks par apprenant	108
5.8	Réponses des participants face à la proposition : “J’ai le sentiment d’avoir progressé”	109
5.9	Réponses des participants face à la proposition : “Le système me propose des aides adaptées à mes erreurs”	109

Liste des tableaux

3.1	Caractéristiques calculées depuis l'activité de l'apprenant	68
3.2	Propriétés du jeu de données multilabel	69
3.3	Résultats de la classification multilabel sur le jeu de données (classe, classification) en validation croisée sur cinq sous-ensembles	70
3.4	Résultats de la classification multilabel sur le jeu de données (exemple, classification) en validation croisée sur cinq sous-ensembles	70
3.5	Aires sous la courbe ROC par motif d'erreur (classe, ranking) en validation croisée sur cinq sous-ensembles	70
3.6	Résultats de la classification multilabel sur le jeu de données (exemple, ranking) en validation croisée sur cinq sous-ensembles	70
3.7	Comparaison des aires sous la courbe ROC entre les modèles bayésiens naïfs (génératifs) et régressions logistiques (discriminants)	71
5.1	Caractéristiques calculées depuis les données extraites de l'activité de l'apprenant	104
5.2	Comparaison des aires sous la courbe ROC entre les modèles bayésiens naïfs (génératifs) et régressions logistiques (discriminants)	105
5.3	Comparaison entre les performances des deux groupes expérimentaux à chaque itération (tests t de Student), et entre les progressions globales (tests t de Student appariés)	107

Remerciements

Si la réalisation d'un doctorat est une épreuve parfois solitaire, j'ai eu la chance d'être accompagné tout au long de ces trois années par de nombreuses personnes. Ces personnes ont partagé les bons moments, et m'ont été d'un grand réconfort dans les périodes plus difficiles.

Je remercie tout d'abord mes directeurs de thèse, Indira Thouvenin et Olivier Gapenne, de m'avoir choisi pour réaliser cette thèse. Nos échanges au cours de ces trois ans ont permis, peu à peu, la construction des travaux présentés dans ce document. Je tiens également remercier Yann Soullard qui, en plus d'être un ami sur lequel on peut compter, a dépassé à de nombreuses reprises son rôle de post-doctorant pour m'encourager et m'aider. Sa participation et son influence sur les directions prises au cours de cette thèse, ainsi que sur la rédaction du présent document, est indéniable.

Je remercie bien sûr les amis "historiques", David, Benoît, Fabien, Samir, Loïc, Etienne, Maxime et Jérôme. Entre collocation improvisée et sorties au Thirsty, ils m'ont permis de trouver un équilibre précieux. Comment parler d'équilibre sans remercier Adé qui m'écoute tous les soirs raconter ma journée (et râler de temps en temps...) sans jamais cesser de m'encourager, et qui m'a permis de survivre au stress précédent la soutenance.

Je tiens à mentionner ici les différents membres de la "team RV" et plus largement de l'équipe ICI, Kévin & Kévin, Loïc à mon arrivée, puis Florian & Florian, Yann, Dorian, Sarah, Samba, Idir et Freddy, qui ont aussi contribué à la bonne ambiance qui a toujours régné dans nos bureaux du CI/GI. Les "collègues du train" également, Lucile, Juliette, Domitile et Anaïs, avec qui le trajet Paris-Compiègne paraissait plus court.

Au cours de mes études, j'ai eu l'occasion de suivre de nombreux cours donnés par de nombreux professeurs. Parmi ces derniers, quelques-uns m'ont particulièrement marqué, et m'ont (sûrement sans le savoir) donné envie de continuer. Je souhaiterais donc remercier Dominique Fettré, Christophe Leclerc, Virginie Camille et Djalil Kateb pour leur rigueur et leur pédagogie.

Last but not least, je remercie de tout mon coeur ma famille, et plus particulièrement mes parents et ma soeur Cécile, pour le soutien sans faille qu'ils m'ont toujours apporté.

Résumé

L'apprentissage de geste est un processus complexe, impliquant de nombreux processus sous-jacents (psychomoteurs, cognitifs, biophysiques). Cet apprentissage peut être divisé en plusieurs grandes phases, définies par la capacité de l'apprenant à produire et à percevoir son geste. À l'instar d'un formateur adaptant son discours et son attitude à l'apprenant, un environnement d'apprentissage doit pouvoir adapter les aides qu'il fournit, d'une part en analysant l'état de l'apprenant, et d'autre part en prédisant les aides qui bénéficieront à son apprentissage. Nous proposons une approche modélisant les interactions entre l'apprenant et l'environnement d'apprentissage ; c'est-à-dire l'évolution de l'état de l'apprenant en fonction des configurations successives de l'environnement ; permettant une adaptation automatique et dynamique de la sélection d'aides.

En s'appuyant sur des travaux portant sur la reconnaissance de gestes et les tuteurs intelligents, nous proposons un cadre formel permettant de représenter la qualité de l'activité gestuelle dans un espace métrique. Cette représentation interprète les résultats issus de modèles probabilistes comme des variables floues illustrant le niveau de l'apprenant sur chacune des dimensions du geste. Nous représentons l'environnement comme l'ensemble des actions possibles, chaque action étant représentée par un vecteur de paramètres estimant la situation dans laquelle l'action est la plus pertinente. Dans un premier temps, ces paramètres sont fixés par des experts, et nous considérons cette expertise comme parfaite. Dans un second temps, nous étudions une problématique omniprésente dans les domaines des environnements d'apprentissage, qui fait état de l'écart entre l'expertise et la réalité. Les domaines "mal définis" sont en effet répandus, du fait du coût et de la difficulté à réunir des experts, et de la complexité inhérente à la définition précise d'un domaine.

Ne considérant plus le seul avis des experts, le processus de sélection d'aides peut alors être vu comme une séquence de décisions dont l'objectif est de proposer à chaque itération l'action qui maximisera le gain en matière d'apprentissage. En s'appuyant sur des travaux récents portant sur les séquences de décisions, notre approche considère le processus de sélection d'aides comme un problème de bandit. Le problème de bandit vise à maximiser un gain lors d'une séquence de décisions, et modélise le compromis entre exploration (choisir une action dont l'influence est inconnue), et exploitation (choisir la meilleure action connue). Nous proposons dans ce cadre une extension des méthodes SoftMax.

L'implémentation de notre modèle sur une plateforme d'apprentissage de la calligraphie a été réalisée dans le cadre d'une collaboration avec des experts de ce domaine. Nous montrons, au travers de deux cas d'étude, l'intérêt de notre modèle pour l'apprentissage de la calligraphie. Dans le premier cas d'étude, l'adaptation est construite depuis notre représentation du geste, et un ensemble de règles fixées par des experts. Nous y montrons l'avantage apporté par l'apport d'une diversité d'aides. Dans le second cas d'étude, nous comparons deux types d'adaptation : une adaptation basée sur des règles, et une adaptation basée sur notre approche dynamique. Nous montrons les différences entre ces deux approches, et illustrons les avantages de l'approche dynamique lorsque les règles sont imprécises.

Abstract

Gesture learning is a complex and multi-step process where trainees are supposed to improve several psychomotor and cognitive skills. This process can be divided into phases depending on trainees' ability to perform and perceive their gestures. As human tutors adapt their behavior according to their perception and understanding of trainees learning situations' a learning environment should select an appropriate behavior from a representation of trainees' learning states and a prediction of the potential influence of every possible behavior. The work presented in this document describes an approach modeling the interactions between a trainee and a learning environment : it represents trainees' consecutive performances and the influence of the environment on these performances. This approach aims at permitting an adaptive selection of the pedagogical actions (i.e. behaviors) implemented in the environment.

Relying on related works in the domains of gesture recognition and intelligent tutoring systems, we propose to represent the gestural activity in a metric space. This representation interprets results from a probabilistic mode as fuzzy variables highlighting trainees' level on every aspect of the gesture. We represent the environment as the set of actions it can select, every action being represented by a feature vector describing the learning situation maximizing the action's influence. As a first step, these features are given by a set of experts, and we consider the rules provided as perfect. As a second step, we study an ubiquitous issue in the field of learning environments, which is the difference between the rules provided by experts and the reality of trainees' needs toward feedback. Ill-defined domains are indeed more and more common, as collecting expert knowledge is difficult and costly, and as studied learning domains are becoming more and more complex and difficult to define.

In this second step, the selection process does not rely on expert knowledge, and this process can be seen as a sequence of decisions. At each iteration, the goal is thus to select the action which would maximize the reward in terms of benefits for trainees' learning. The action selection process is represented as a multi-armed bandit problem, where the goal is to compromise between exploration of unknown actions and exploitation of known actions. We present an extension of SoftMax methods which handles multi-dimensional contextual rewards. Taking advantage of the collaboration with calligraphy experts, a calligraphy training platform was implemented as part of this work. Two studies, where participants train on this platform, show the benefits of the proposed

approach on calligraphy learning. In a first experiment, action selection is based on expert rules, and we show that providing a diversity of feedback improves skill acquisition. In a second experiment, we compare two configurations of the environment : a selection of actions based on expert rules, and a selection of actions based on the SoftMax method. We describe the difference between the influence of these two approaches on trainees' learning, and we point out the benefits of using dynamic rules.

Chapitre 1

Introduction

Nos travaux s’inscrivent dans le cadre du projet Descript, financé par la région Picardie. Ce projet, dont la vocation est de concevoir un environnement d’apprentissage du geste d’écriture, s’intéresse notamment à la calligraphie et à son apprentissage. Il mobilise des membres de l’Université de Technologie de Compiègne et de l’École Supérieure d’Art et de Design d’Amiens. Notre contribution au sein du projet Descript concerne la conception, l’implémentation et la validation d’un environnement intelligent d’apprentissage du geste calligraphique qui, à terme, servira dans le cadre des formations proposées à l’ESAD.

L’étude des environnements intelligents d’apprentissage fait partie des axes de recherche de l’équipe ICI (Information, Connaissance, Interaction) du laboratoire Heudiasyc. Plus précisément, nos travaux se situent dans la thématique “connaissance et interaction” de l’équipe, qui s’attache à l’étude des environnements interactifs dits *informés*. Selon la définition proposée par (Thouvenin, 2009), “un environnement informé est un environnement dont les comportements sont adoptés suite à l’interprétation de représentations dynamiques ou statiques de l’utilisateur”. Il s’agit donc d’une part de représenter l’utilisateur (son état, son activité) et d’autre part de définir des heuristiques permettant l’adoption d’un comportement adéquat.

Ces problématiques furent notamment étudiées ces dernières années au travers des thèses de Kévin Carpentier (Carpentier, 2015), qui étudie l’apport de la scénarisation dans les environnements virtuels de formation, et Loïc Fricoteaux (Fricoteaux, 2012) qui propose un environnement de formation à la navigation fluviale tenant compte de l’incertitude des données récupérées par les différents capteurs mis en oeuvre.

1.1 Contexte scientifique

La création d’environnements d’apprentissage, et d’apprentissages de gestes en particulier, pose de nombreux défis. La majorité des environnements portant spécifiquement sur l’apprentissage de gestes se focalisent sur la boucle de

rétroaction. Cette boucle est constituée de l'utilisateur "apprenant" qui agit, et de l'environnement qui, par rapport à ces actions, réagit en proposant des *feedbacks*. Ces *feedbacks* peuvent prendre différentes formes (e.g. une force poussant l'apprenant dans la bonne direction, une information visuelle, un signal sonore). La recherche portant sur les environnements d'apprentissage de gestes se focalise alors sur l'influence de chaque caractéristique du *feedback* sur les différentes dimensions du geste.

Les travaux s'inscrivant dans le cadre des environnements intelligents d'apprentissage s'attachent quant à eux à modéliser l'apprenant, c'est à dire à représenter son état à court terme (e.g. engagement dans la tâche, attention) et à long terme (connaissances maîtrisées, connaissances restant à maîtriser). Depuis une représentation de l'apprenant, les environnements intelligents d'apprentissage cherchent alors quel comportement adopter afin d'optimiser l'apprentissage de l'apprenant.

Nous présentons dans ce chapitre les principaux résultats issus d'études relatives au processus d'apprentissage de gestes (section 1.2). Nous présentons ensuite plusieurs travaux portant sur les environnements d'apprentissage de gestes (section 1.3) et les environnements intelligents d'apprentissage (section 1.4). La mise en perspective des problématiques liées à ces deux domaines par rapport au processus d'apprentissage de gestes motive la problématique générale de nos travaux, que nous précisons dans la section 1.5. Finalement, nous présentons l'organisation générale du mémoire dans la section 1.6.

1.2 Apprentissage de gestes

La littérature relative à l'étude du processus¹ d'apprentissage de gestes est très riche, et de nombreuses théories ont été proposées afin de décrire et d'expliquer cet apprentissage (Schmidt and Lee, 1988). Cette section ne vise donc pas à décrire de manière exhaustive l'état de la recherche dans ce domaine, mais à énoncer les principales conclusions qui, dans la perspective de la création d'un environnement d'apprentissage, nous intéressent.

1.2.1 Approche cognitive

Comme de nombreuses théories relatives au fonctionnement de l'organisme, les modèles cognitivistes reposent sur une métaphore machinique. Les expérimentations menées par (Hick, 1952) et (Fitts, 1954) montrent que le temps de réponse du sujet suite à un stimulus (un signal lumineux pour (Hick, 1952), une cible à pointer pour (Fitts, 1954)) est proportionnel à la quantité d'information que le sujet doit traiter avant d'agir. Ces travaux légitiment l'hypothèse cognitive du traitement de l'information, qui fait état d'un processus cognitif entre la perception du stimulus et l'établissement de la réponse. D'après le

1. Au sens propre, l'apprentissage de gestes est constitué de nombreux processus. Ainsi, lorsque l'on se réfère dans ce document au processus d'apprentissage de gestes, il faut y comprendre le processus global contenant tous les sous-processus inhérents à cet apprentissage.

modèle de Schmidt (Schmidt and Lee, 1988), le processus cognitif est une suite d'opérations correspondant à l'identification du stimulus, la décision relative à la réponse à déclencher, l'organisation de cette réponse, sa réalisation et son contrôle. Une modélisation complémentaire de celle de Schmidt considère le traitement de l'information comme un système de mémoires (Thomas, 1980) constitué de la mémoire sensorielle, la mémoire de travail (ou mémoire à court terme) et la mémoire à long terme. La mémoire sensorielle stocke pour un temps très court les informations perçues par le sujet. La mémoire de travail ne reçoit qu'une partie des informations disponibles en mémoire sensorielle (théorie d'un filtre perceptif ne laissant passer que les informations pertinentes). La mémoire de travail a une capacité limitée et une faible durée de stockage actif, de l'ordre de quelques secondes. Finalement, la mémoire à long terme possède une capacité considérée comme illimitée. Dans cette mémoire sont stockées toutes les expériences du sujet. Dans la logique de ce modèle, les processus cognitifs assurent le transit de l'information entre ces différentes mémoires.

L'apprentissage, dans le cadre cognitiviste, est un processus en trois phases décrites comme suit par (Fitts and Posner, 1967) :

- la phase cognitive est la première phase de l'apprentissage. Le sujet découvre la tâche à accomplir, et son premier problème est d'identifier le but à atteindre et les moyens à mettre en oeuvre pour y parvenir. Les mouvements sont saccadés, et la demande attentionnelle est très élevée et accapare le sujet. Ce stade est considéré comme achevé lorsque les premières réponses satisfaisantes sont produites ;
- la phase associative consiste à affiner progressivement le programme moteur. Cette phase est caractérisée par l'accroissement de la régularité des réponses ;
- la phase autonome est l'étape pendant laquelle l'apprenant automatise le processus, qui ne demande alors plus aucun contrôle attentionnel.

1.2.2 Approche dynamique

Dans le cadre des théories dynamiques, le comportement moteur est un phénomène émergent de contraintes liées à la tâche, à l'organisme, ou à l'environnement (Newell, 1986). Ces contraintes limitent les degrés de liberté du système, c'est à dire ses possibilités d'action. La métaphore convoquée ici est celle du pendule, dont on peut prédire la fréquence d'oscillation en fonction de sa longueur pour peu que l'amplitude de ses oscillations soit suffisamment faible. On dit alors que le comportement émerge spontanément des caractéristiques du pendule pour peu qu'un déséquilibre soit instauré. L'expérience de (Holt et al., 1991) montre qu'il est possible de prédire la fréquence naturelle de marche grâce à une équation pendulaire. Dans la théorie dynamique, le système est modélisé comme un ensemble d'oscillateurs qui peuvent fonctionner simultanément et s'influencer réciproquement. La coordination des oscillateurs constitutifs du système peut alors être capturée par des mesures appelées *variables collectives*. Ces variables mesurent par exemple la fréquence d'un oscillateur, ou le décalage de phase entre deux oscillateurs (ce décalage permet-

tant notamment de mesurer la coordination gestuelle). La théorie dynamique permet également de définir des comportements naturels, spontanés, appelés *attracteurs*. L'attracteur correspond à une situation où le système est stable (faible variabilité des variables collectives).

L'apprentissage de gestes dans le cadre de l'approche dynamique consiste à maîtriser des degrés de liberté. Le problème des degrés de liberté est selon (Bernstein, 1967) un problème de redondance : dans une activité complexe, le système présente davantage de degrés de liberté qu'il n'en est nécessaire pour réaliser la tâche. Maîtriser un geste revient alors à réduire le système à un système plus simple en réduisant les degrés de liberté aux seuls nécessaires à l'accomplissement du geste. L'apprentissage selon (Bernstein, 1967) se divise en trois phases :

- la phase de contrôle correspond à l'étape où l'apprenant novice réduit par lui-même le nombre de degrés de liberté de son geste, phénomène établi notamment dans le cadre d'apprentissage du tir au football (Arutyunyan et al., 1968) et de l'apprentissage du ski (Vereijken et al., 1992) ;
- la phase d'exploration succède à la phase de contrôle. Dans cette phase, l'apprenant ayant acquis un contrôle depuis un ensemble réduit de degrés de liberté relaxe cet ensemble. Dans le cadre du football (Arutyunyan et al., 1968) et du ski (Vereijken et al., 1992), les articulations initialement bloquées/rigidifiées par l'apprenant dans la phase de contrôle se relâchent et sont remplacées par une compensation depuis d'autres ensembles articulaires ;
- la phase d'exploitation est l'étape durant laquelle l'apprenant optimise son geste, qui gagne en efficacité. L'organisation gestuelle devient plus économique et l'apprenant sait maîtriser non seulement son geste mais également l'inscrire dans un environnement (e.g. exploitation du travail du poids lors de la réalisation par des experts de balanciers en suspension renversée aux barres parallèles (Delignières et al., 1998b)).

1.2.3 Conclusions

L'apprentissage de gestes est un processus long et complexe dont la compréhension, malgré les théories de plus en plus précises qui le caractérisent, reste un problème ouvert. Les deux théories majoritaires que sont l'approche cognitive et l'approche dynamique partagent néanmoins un constat : cet apprentissage se fait en plusieurs phases, phases durant lesquels l'apprenant se concentre sur différents aspects de son activité. Pour optimiser l'apprentissage, il est alors nécessaire de confronter l'apprenant à une diversité de feedbacks qui lui permettront de prendre conscience et de développer des dimensions différentes de son activité (Fitts and Posner, 1967; Schmidt and Lee, 1988). Suite à ces conclusions, nous présentons dans la prochaine section les principales problématiques adressées par les environnements d'apprentissage de gestes.



FIGURE 1.1 – Exemples de ductus, modèles détaillés de la trace à produire

1.3 Environnements d'apprentissage de gestes

Nous nous concentrons ici sur les environnements dédiés à l'apprentissage de la calligraphie, ou plus généralement à l'apprentissage de l'écriture manuscrite. Le geste d'écriture est une activité complexe impliquant de profonds développements psychomoteurs et cognitifs (Van Galen, 1991). La production de l'écriture repose sur deux contraintes (Paillard, 1990) : les contraintes morphocinétiques et topocinétiques. La contrainte morphocinétique considère la forme du tracé (e.g. la forme d'une lettre), alors que la contrainte topocinétique considère la spatialisation du tracé dans l'espace graphique. Les ductus (figure 1.1) sont des modèles communément utilisés dans le cadre de l'apprentissage de la calligraphie, qui illustrent la dimension morphocinétique du tracé. D'un point de vue moteur, deux types de contrôle sont convoqués lors de la pratique de l'écriture : le contrôle *proactif*, qui correspond à la préparation et l'anticipation du mouvement, et le contrôle *rétroactif* qui correspond aux ajustements effectués au cours du geste.

1.3.1 Feedback intrinsèque

Les premiers environnements d'apprentissage de l'écriture manuscrite, développés dès le début des années 80, s'intéressent à la mobilisation des feedbacks intrinsèques (i.e. le feedback inhérent à la réponse motrice) lors du geste d'écriture. Les deux principaux feedbacks intrinsèques mobilisés dans l'activité d'écriture sont le feedback visuel et le feedback proprioceptif. Il apparaît que le feedback visuel joue un rôle crucial dans le contrôle du tracé (contrôle rétroactif), et permet au scripteur de correctement former les lettres et les lier entre elles pour former des mots (Danna and Velay, 2015). Le feedback proprioceptif permet quant à lui d'assurer la fluidité du geste (Danna and Velay, 2015).

1.3.2 Feedback extrinsèque

Depuis les premières études portant sur les feedback extrinsèques (Søvik and Teulings, 1983), le développement technologique (i.e. capteurs plus précis et moins coûteux, meilleure qualité des écrans, développement des interfaces haptiques) a facilité et enrichi les possibilités d'étude de ces feedbacks sur l'apprentissage. Les feedbacks extrinsèques sont des informations supplémentaires proposés pendant (feedback temps-réel) ou après (feedback a posteriori) la production du geste. Parmi l'ensemble des feedbacks extrinsèques, on distingue les feedback relatifs à la *connaissance du résultat*, qui renseignent "une information relative à la congruence du résultat par rapport à un but" (Simonet, 1985) ; des feedbacks relatifs à la *connaissance de la performance*, qui renseignent "une information relative à la nature des opérations mises en oeuvre" (Simonet, 1985). Tout feedback peut également être distingué par sa modalité (i.e. visuelle, sonore, haptique), ou son degré d'explicitation (e.g. intensité de la force appliquée dans le cadre d'un feedback haptique).

Les environnements d'apprentissage de gestes étudient l'influence de ces caractéristiques. Ces environnements proposent pour cela la répétition d'un même exercice au cours duquel un type spécifique de feedback est proposé. Un grand intérêt a été porté ces dernières années aux feedbacks haptiques (Feygin et al., 2002; Teo et al., 2002; Ben-Pazi et al., 2009; Palluel-Germain et al., 2007) qui permettent de corriger l'apprenant en appliquant des forces sur l'outil de l'apprenant, cet outil étant robotisé. Les feedbacks haptiques ont montré des résultats encourageant quant à la correction d'erreurs relatives à la dynamique du geste (Teo et al., 2002). Cette modalité est cependant critiquée du fait qu'elle nécessite l'enregistrement d'une performance experte qu'il faut ensuite reproduire (Danna and Velay, 2015).

L'utilisation de feedbacks sonores (Danna et al., 2015) informant l'apprenant à propos de la dynamique de son geste a récemment suscité un intérêt. Le feedback sonore semble en effet, dans le cadre de caractéristiques relatives à la dynamique du geste, informer de façon plus naturelle que les feedbacks visuels (Fitch and Kramer, 1994). (Ronsse et al., 2011) montrent également que la dépendance aux feedbacks sonores est moins forte que ne peut l'être la dépendance aux feedbacks visuels et haptiques. Les feedbacks sonores peuvent également être combinés aux autres types de feedbacks (Plimmer et al., 2011) (ce qui n'est pas nécessairement le cas des feedbacks visuels et haptiques). Un inconvénient des feedbacks sonores est par contre la difficulté pouvant être rencontrée à interpréter le sens du feedback (Sigrist et al., 2013).

1.3.3 Dépendance à l'environnement

L'un des principaux problèmes rencontrés par les environnements d'apprentissage de gestes réside dans l'incapacité des apprenants à retranscrire en conditions réelles (i.e. dans un environnement ne proposant pas de feedback extrinsèque) les progrès effectués lors de l'apprentissage (Schmidt and Lee, 1988; Ronsse et al., 2011). Cette incapacité s'explique par le fait que les apprenants

utilisent démesurément un contrôle rétroactif dont la perception est enrichie par les feedback intrinsèques délivrés par l'environnement. Une fois de retour en conditions réelles, le contrôle proactif nécessaire à la production gestuelle n'a pas progressé, et le contrôle rétroactif ne peut plus se reposer sur de quelconques feedbacks extrinsèques.

L'une des approches proposées pour lutter contre la dépendance à l'environnement consiste à diversifier les feedbacks proposés (Bayart et al., 2005). Dans cette approche, les apprenants répètent un exercice consistant à copier un symbole. Divers niveaux de feedback sont proposés lors de la séquence d'exercice : initialement l'apprenant n'a qu'à saisir l'outil qui "bouge tout seul"; dans une seconde étape une légère force guide l'outil de l'apprenant dans la bonne direction; dans une troisième étape les forces ne sont appliquées que lorsque l'apprenant s'éloigne de la trajectoire idéale; finalement dans la quatrième étape aucune force n'est appliquée. Cette diversification du feedback permet dans (Bayart et al., 2005) de réduire le phénomène de dépendance à l'environnement, en comparaison avec des groupes de participants pour lesquels le même type de feedback est proposé tout au long de la séquence d'exercices. La sélection du type de feedback à proposer est néanmoins indépendante de la performance de l'apprenant. Les problématiques relatives à la représentation de l'activité de l'apprenant, et à la sélection de comportements pertinents par rapport à cette représentation, n'ont pas été adressées dans le cadre de l'apprentissage de gestes. Ce sont cependant les principaux défis relevés dans le domaine des environnements intelligents d'apprentissage, que nous présentons maintenant.

1.4 Environnements intelligents d'apprentissage

Les environnements intelligents d'apprentissage étudient la manière dont l'environnement peut adapter son comportement en fonction de l'état de l'apprenant, afin d'optimiser son apprentissage. Depuis les premiers environnements proposés (Rich, 1979), de nombreux domaines d'apprentissage ont été étudiés, tels que l'arithmétique (Beal et al., 2010), l'algèbre (Koedinger et al., 1997), ou la programmation (Holland et al., 2009). L'adaptation de l'environnement à l'état de l'apprenant présente plusieurs enjeux : la modélisation de l'état de l'apprenant, et la stratégie permettant de sélectionner un comportement pertinent depuis une représentation de l'apprenant.

Le modèle de l'apprenant permet de situer ce dernier par rapport à l'ensemble des compétences qu'il doit acquérir. Dans des domaines d'apprentissage tels que l'arithmétique ou l'algèbre, de nombreuses approches ont été proposées permettant de construire des représentations de l'apprenant à court (i.e. représentation instantanée) et à long terme (i.e. mise en perspective par rapport à de précédentes représentations) (Mayo, 2001).

Les stratégies permettant à l'environnement d'adopter un comportement pertinent étant donnée une représentation de l'apprenant se basent sur des théories de l'apprentissage en psychologie. Les environnements se réfèrent alors

à un ensemble de règles définies par des experts et établissant des relations entre situations d'apprentissage, et comportement à adopter. Les experts s'appuient, lors de l'élaboration des règles, sur leur propre expérience, mais également sur des principes tels que la zone proximale de développement (Vygotsky, 1978; Murray and Arroyo, 2002; Clement et al., 2013). La zone proximale de développement (ZPD) est définie par l'auteur comme la zone comprise entre le niveau actuel et le niveau potentiel de l'apprenant. Le niveau actuel étant le niveau tel qu'il peut être mesuré lorsque l'apprenant répond à un problème, et le niveau potentiel correspondant au niveau qu'il peut atteindre lorsqu'il est aidé par des pairs initiés. Selon (Vygotsky, 1978), un comportement est adapté s'il est situé dans la ZPD, et un tel comportement favorisera la motivation et l'apprentissage de l'apprenant. Dans le cas où l'adoption d'un comportement consiste à sélectionner une nouvelle activité (e.g. un nouvel exercice d'algèbre) pour un apprenant, une activité pertinente est donc une activité plus complexe que celles ayant été faites avec succès, mais suffisamment simple pour que l'apprenant puisse la résoudre s'il est assisté par des pairs.

L'adaptation du comportement de l'environnement ne se résume pas à la sélection de la prochaine activité à proposer à l'apprenant. Plusieurs études (Kumar et al., 2007; Barnes and Stamper, 2008) portent notamment sur la faculté de l'environnement à prodiguer des conseils adaptés à l'état de l'apprenant (e.g. insister sur un aspect spécifique mal maîtrisé).

Contrairement aux environnements d'apprentissage de gestes qui se focalisent sur un type d'activité (un exercice à répéter, un type de feedback) afin d'en étudier l'influence, les environnements intelligents d'apprentissage permettent de proposer une diversité d'exercices (un ou plusieurs exercices, plusieurs types de feedback) pour favoriser l'apprentissage. Si l'on en suit la théorie du Schéma de Schmidt (Schmidt and Lee, 1988), un environnement d'apprentissage de gestes doit pouvoir proposer une diversité de feedbacks pour réellement profiter à l'apprentissage. Il apparaît donc intéressant d'étudier la possibilité de transposer les méthodes mobilisées dans les environnements intelligents d'apprentissage dans le cadre du geste. Toutefois, la transposition immédiate de ces méthodes n'est pas évidente. Nous présentons ainsi dans la prochaine section les problématiques majeures inhérentes à la création d'un environnement intelligent pour l'apprentissage de gestes.

1.5 Problématiques liées aux environnements intelligents d'apprentissage de gestes

La problématique générale de nos travaux consiste à créer un environnement capable d'adapter son comportement (i.e. sélectionner une activité pertinente) par rapport à une activité gestuelle. Nous divisons cette problématique en quatre thématiques : la caractérisation, l'interprétation et la représentation de l'activité gestuelle, et la stratégie d'adaptation de l'environnement par rapport à cette activité. Suite à la présentation de ces quatre problématiques, nous

précisons notre positionnement par rapport aux principales approches de la littérature.

1.5.1 Caractérisation

Le geste est le fruit de nombreux processus sous-jacents, et de nombreuses dimensions, ou aspects, se combinent lors de l'évaluation d'un geste (e.g. sa fluidité, sa précision, son amplitude). Une première étape consiste donc à caractériser l'activité gestuelle, c'est à dire établir un ensemble de caractéristiques suivant lesquelles le geste est considéré.

1.5.2 Interprétation

Les environnements intelligents ont été proposés dans le cadre de domaines dits de résolution de problème (nous avons cité précédemment des exemples appartenant à de tels domaines, comme l'algèbre ou l'arithmétique). S'il existe une incertitude quant à la signification des réponses proposées par l'apprenant face à un problème (une réponse correcte peut être due à un facteur chance, une réponse incorrecte à une erreur d'inattention), ces réponses sont malgré tout "interprétables" dans le sens où elles expriment une valeur définie qui peut être comparée à une valeur solution. La performance gestuelle n'est pas directement interprétable puisqu'une fois capturée par un ensemble de capteurs elle n'est représentée que par une séquence de positions et d'orientations dans l'espace. Un processus supplémentaire est donc nécessaire avant de pouvoir dériver du geste une valeur, ou un ensemble de valeurs, que l'environnement puisse interpréter.

1.5.3 Représentation

La notion d'acceptabilité de la réponse de l'apprenant diverge également selon que l'on se situe dans le domaine de la résolution de problème ou celui du geste. La bonne solution dans le cadre de la résolution de problème est celle qui répond à une loi, ou respecte les contraintes qui définissent le problème. La solution dans le cadre du geste est plus ambiguë, puisque la plupart des experts exerçant dans un domaine lié au geste critiquent la théorie suivant laquelle il puisse exister un unique geste idéal à reproduire. Ils estiment au contraire qu'il existe théoriquement une infinité de "bons gestes", le geste étant dépendant de caractéristiques individuelles telles que la morphologie. Ces différences se traduisent par le besoin de nouveaux modèles de l'apprenant qui prennent en compte la spécificité inhérente à la pratique gestuelle.

1.5.4 Adaptation

L'adaptation du comportement de l'environnement par rapport à une représentation de l'apprenant nécessite également de prendre en compte la spécificité des processus sous-jacents à la pratique du geste. Si les environnements

intelligents reposent sur une expertise, mobilisant potentiellement des théories telles que la ZPD, il n'est ainsi pas certain que ces théories s'appliquent à l'apprentissage de gestes. Jean Piaget a notamment mis en doute la pertinence d'utiliser les théories qu'il a développées dans le cadre de l'apprentissage de capacités motrices (Delignières et al., 1998a). Une problématique de première importance réside donc en l'établissement et la comparaison des stratégies que pourraient suivre un environnement d'apprentissage de gestes.

1.5.5 Choix et positionnement

Le projet Descript ayant pour but de créer un environnement intelligent d'apprentissage qui puisse être utilisé lors de certaines phases de la formation à la calligraphie dispensée à l'ESAD, nous nous concentrons dans la majeure partie de nos travaux sur des périodes d'apprentissage courtes, de l'ordre de trente minutes à une heure. Ce choix a une incidence sur le choix du nombre et des types des feedbacks implémentés au sein de l'environnement, ainsi que sur les stratégies pédagogiques adoptées. Travailler sur de courtes périodes d'apprentissage nous permet en effet de réduire le nombre de feedbacks à implémenter. Les problématiques liées au contexte d'apprentissage sont également réduites lors de périodes courtes d'apprentissage. Nous proposerons tout de même, dans le dernier chapitre de ce mémoire, une extension de notre modèle permettant la prise en compte du contexte.

Bien que les environnements usuels d'apprentissage de gestes se focalisent sur l'influence de la nature du feedback sur l'apprentissage, nos travaux ne s'inscrivent pas dans cette problématique. Les feedbacks implémentés dans le cadre de l'environnement que nous proposons sont identifiés comme pertinents par des experts en calligraphie, et l'évaluation de ces types de feedback à titre individuel sur l'apprentissage n'est pas le sujet de nos travaux. Nous nous attacherons par contre à proposer des méthodes permettant la sélection de feedbacks pertinents par rapport à la situation de l'apprenant, et ce sont les influences de ces stratégies d'adaptation que nous évaluerons.

Certains environnements permettent à plusieurs apprenants de collaborer lors de la réalisation d'une tâche, ou enrichissent l'interaction entre un apprenant et un tuteur humain. Si ces environnements offrent des perspectives intéressantes, encourageant notamment les participants à s'entre-évaluer et ainsi à questionner leurs performances (Slavin, 1990), la problématique pédagogique est alors tout autre : il s'agit davantage de favoriser les échanges entre les apprenants (e.g. en proposant des interfaces de chat, en permettant à plusieurs utilisateurs de mutualiser leurs connaissances) que de guider/informer l'apprenant lors de sa performance. Dans l'optique d'un environnement d'apprentissage de gestes utilisable à l'ESAD, qui s'adresse en priorité à des utilisateurs apprenant seuls sur la plateforme développée, nous ne nous positionnons pas dans le cadre collaboratif et concentrons nos travaux sur une interaction entre un apprenant et un environnement d'apprentissage.

Précédemment dans ce chapitre, nous avons précisé que, dans le cadre de l'interaction entre l'apprenant et l'environnement, la réaction de l'environne-

ment face aux actions de l'utilisateur consistait à sélectionner un feedback pertinent. La faculté de l'environnement à sélectionner un feedback par rapport à sa pertinence eu égard à la situation d'apprentissage constitue notre définition de l'adaptation de l'environnement. Deux grands types d'adaptation sont référencés dans la littérature relative à l'adaptation des logiciels : les logiciels sont dits adaptatifs² lorsqu'ils modifient leur comportement de manière autonome, sans contrôle laissé à l'utilisateur ; au contraire les logiciels sont dits adaptables³ s'ils peuvent adopter plusieurs comportements, mais que la sélection de ces comportements est laissée à l'utilisateur (Oppermann and Rasher, 1997). Dans le cadre du geste, il est difficile pour un apprenant, d'autant plus s'il est novice, de mesurer l'apport d'un type de feedback sur son apprentissage, c'est pourquoi nous dirigeons nos recherches vers les environnements adaptatifs.

1.6 Organisation du mémoire

Nous nous intéressons dans la première partie de ce manuscrit aux problématiques de caractérisation et de représentation du geste. Pour cela, nous détaillons dans le chapitre 2 la littérature relative à la représentation de l'apprenant dans les environnements intelligents. Outre les approches proposées dans le cadre des domaines liés à la résolution de problèmes, nous présentons des approches récentes proposées dans le cadre de domaines dits "mal définis", c'est à dire dont la caractérisation et l'évaluation posent de nouveaux défis. Nous précisons en quoi, par rapport aux différentes taxonomies des domaines d'apprentissage proposées, l'apprentissage de gestes entre dans cette catégorie des domaines mal définis. Suite à cet état de l'art, nous présentons notre approche. Dans cette approche, nous proposons une reformulation générale du problème de représentation du couple apprenant/environnement d'apprentissage. En représentant ces deux entités dans un même espace de représentation métrique, nous montrons qu'il est possible de représenter à la fois l'apprenant, les feedbacks, et les règles de sélection de comportements. Nous motivons également dans ce chapitre l'adoption d'une caractérisation du geste selon des dimensions d'erreur.

Le chapitre 3 étudie la question de l'interprétation du geste. Si nous caractérisons le geste par des dimensions (ou classes) d'erreur, encore faut-il pouvoir évaluer une performance relativement à ces dimensions (i.e. afin de positionner le geste dans l'espace de représentation métrique). Nous proposons d'interpréter le geste par la probabilité que l'erreur ait été commise lors de la performance. La reconnaissance et l'interprétation de gestes sont des problématiques étudiées dans le domaine de l'apprentissage automatique. Nous présentons dans ce chapitre les principales méthodes de la littérature. Du fait que le geste soit caractérisé par plusieurs dimensions d'erreur, et plusieurs erreurs pouvant être détectées pour un même geste, le problème de classification devient multilabel. Nous comparons alors, dans le cadre d'une expérimentation utilisant un jeu

2. Adaptive software.

3. Adaptable software.

d'observations que nous avons collecté, deux des principales méthodes proposées dans le cadre multilabel : les méthodes de *binary relevance* de *label power-set*. Suite à cette comparaison, nous présentons une première expérimentation mesurant l'influence sur l'apprentissage d'un environnement qui, depuis nos approches de caractérisation, de représentation et d'interprétation du geste, sélectionne son comportement depuis un ensemble de règles définies par des experts.

Dans la seconde partie de ce document, nous nous intéressons particulièrement à la stratégie de sélection de comportements. Une stratégie se basant uniquement sur un ensemble de règles ne permet pas d'illustrer le caractère individuel du besoin de l'apprenant en terme de feedback, et nous nous intéressons donc à des méthodes permettant d'affiner de manière automatique les règles de sélection en fonction des précédentes interactions entre l'environnement et l'apprenant. Nous étudions dans le chapitre 4 la pertinence de l'utilisation d'approches issues des systèmes de recommandation et de l'apprentissage par renforcement. Nous proposons finalement notre approche représentant le problème de sélection comme un problème de bandit, où il s'agit de trouver le compromis entre exploration de nouveaux feedbacks et exploitation des feedbacks connus pour maximiser un gain lors d'une séquence de décisions. Nous présentons dans le chapitre 5 une seconde expérimentation où nous comparons l'adaptation basée sur l'expertise et l'adaptation dynamique. Le chapitre 6 résume les travaux présentés, identifie leurs limites, et propose plusieurs perspectives d'évolution.

Chapitre 2

Adaptation basée sur l'expertise : Modélisation de l'interaction

2.1 Introduction

L'interaction entre un utilisateur humain et un environnement informatisé est un processus bilatéral où l'utilisateur accomplit un certain nombre d'actions face auxquelles l'environnement sélectionne, parmi l'ensemble des comportements qu'il peut adopter, celui ou ceux qui lui paraissent les plus pertinents.

Un aspect fondamental lorsqu'il s'agit de permettre à un environnement informatisé d'interagir avec un utilisateur, indépendamment du but de cette interaction, consiste à modéliser l'utilisateur. Dans le cadre d'environnements d'apprentissage, l'utilisateur "apprenant" peut être caractérisé par de nombreuses propriétés, ces propriétés pouvant représenter sa capacité à accomplir la tâche demandée (e.g. erreurs commises, temps nécessaire à l'accomplissement de la tâche), son état cognitif (e.g. attention portée à la tâche, motivation), ou ses caractéristiques propres (e.g. genre, âge). Outre ces caractéristiques instantanées (qui représentent l'état de l'apprenant à un instant précis), l'étude de l'évolution de ces caractéristiques au fil de l'apprentissage peut également revêtir un intérêt primordial dans l'établissement d'une représentation de l'apprenant.

La définition d'un *modèle de l'apprenant* consiste à réunir un ensemble de caractéristiques permettant de modéliser de manière fidèle l'apprenant eu égard aux objectifs d'apprentissage. Suivant le domaine d'apprentissage, la définition de cet ensemble de propriétés mesurables permettant de modéliser l'apprenant peut s'avérer plus ou moins complexe, d'où l'établissement de taxonomies des domaines d'apprentissage regroupant ces derniers suivants qu'ils soient *bien* ou *mal* définis.

Bien que la modélisation de l'apprenant permette à l'environnement d'avoir une représentation de l'apprenant, cette représentation seule n'est pas suffisante pour que l'environnement puisse interagir. Il est pour cela nécessaire de définir une heuristique, ou un ensemble d'heuristiques, qui permettront à l'environ-

nement d'adopter un ou des comportements¹ en fonction de la représentation qu'il a de l'apprenant. Dans le cadre de l'apprentissage, ces heuristiques sont majoritairement basées sur des règles. Ces règles, fournies par des experts, illustrent le ou les comportements qu'ils auraient adoptés face à chaque situation. En effet l'apprentissage en général, et l'apprentissage de gestes en particulier, est un processus long et complexe. Tout au long de cet apprentissage, l'apprenant évolue, acquiert de nouvelles capacités, prend du recul par rapport à son activité. Cette évolution amène l'apprenant à se focaliser sur différentes composantes de son activité. Le formateur, humain ou synthétique, doit savoir répondre à cette évolution par des comportements adaptés à la situation : il doit analyser l'état de l'apprenant, et en fonction de cette analyse proposer l'aide ou le conseil qui lui permettra de continuer sa progression (une approche *orientée apprenant* telle que présentée par (Woolf, 2009), section 2.4).

Nous nous intéressons dans ce chapitre aux différents types de modèles permettant de caractériser l'utilisateur dans le cadre d'environnements d'apprentissage. Nous précisons dans un premier temps les principales taxonomies permettant de catégoriser les domaines d'apprentissage, en discutant notamment ce que sont des domaines bien et mal définis, et la place de l'apprentissage de gestes par rapport à ces catégories. Nous présentons ensuite les principaux modèles de l'apprenant présents dans la littérature. Nous décrivons également comment, depuis ces modèles, des heuristiques de sélection de comportements peuvent être formulées. Nous motivons et présentons finalement notre approche permettant la sélection d'actions pédagogiques depuis une activité gestuelle.

2.2 Taxonomie des domaines d'apprentissage

Le *domaine* représente l'ensemble des connaissances déclaratives (savoir) ou procédurales (savoir-faire) qui, une fois acquises, fondent l'expertise (Fitts and Posner, 1967; Anderson, 1983). L'objectif d'un environnement d'apprentissage est alors de transmettre tout ou partie de cet ensemble de connaissances. C'est pourquoi la caractérisation du domaine est indissociable, et constitue la première étape, de la modélisation de l'apprenant.

Les domaines d'apprentissage sont potentiellement infinis, et leurs natures sont très hétérogènes. Les compétences que l'apprenant doit acquérir puis savoir mobiliser, ainsi que l'évaluation de la performance de cet apprenant, ont des caractères bien différents suivant que ces connaissances soient liées aux mathématiques, à un geste, à un art. Une taxonomie proposée pour classer ces domaines les différencie en fonction de la complexité inhérente à leur modélisation. La taxonomie de Lynch (Lynch et al., 2006) classe les domaines suivant deux axes orthogonaux, le premier axe représentant la définition du domaine (depuis *bien défini* vers *mal défini*), et le second axe représentant la complexité

1. Plutôt que *comportement*, nous utiliserons plutôt, dans le cadre de l'apprentissage, le terme d'*action pédagogique*. Une action pédagogique peut consister en la sélection d'une activité ou, dans le cadre de nos travaux, la sélection de retours, ou *feedback*, informant l'apprenant sur son activité.

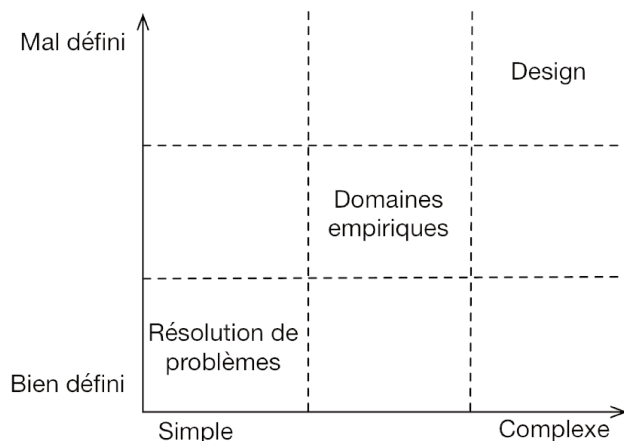


FIGURE 2.1 – Schématisation de la taxonomie des domaines de Lynch et al. (Lynch et al., 2006).

de l'évaluation d'une performance dans ledit domaine (depuis *simple* vers *complexe*, voir Figure 2.1). (Lynch et al., 2006) puis (Woolf, 2009) considèrent qu'un domaine situé dans cet espace de représentation appartient à l'une des trois catégories suivantes :

Résolution de problèmes Les domaines liés à la résolution de problèmes (i.e. problèmes de mathématiques ou de physique Newtonienne) sont les plus simples, puisqu'ils sont bien définis, c'est-à-dire que l'évaluation de la performance (ici le résultat) fournie par l'apprenant peut être effectuée à partir d'un nombre fini de règles.

Domaines empiriques² Sont considérés dans cette catégorie les problèmes tels que l'éthique ou la loi, pour lesquels il est impossible d'évaluer directement la performance sans vérification empirique. Une logique basée sur des règles est ici insuffisante. Pourtant, une expertise peut permettre une évaluation précise et robuste de la performance. Les auteurs situent ces domaines au milieu des deux axes de l'espace.

Design Cette catégorie regroupe des domaines tels que l'architecture, la composition musicale, et plus généralement tous les domaines relatifs à la création. Ces domaines, les plus éloignés du domaine le plus simple de résolution de problèmes, sont situés aux extrémités des deux axes. Les performances ne peuvent être évaluées à partir d'un ensemble de règles, et l'analyse empirique est hautement sujette à la subjectivité, et ne peut par conséquent pas être considérée comme précise, ni comme robuste.

Les domaines *bien définis*³, d'une nature plus exacte, se prêtent plus facilement à une modélisation informatique. Il "suffit" alors de définir une repré-

3. "Well-defined domains", par opposition à "ill-defined domains", que l'on traduira ici par "domaines mal définis".

sentation formelle du domaine (e.g. si le triangle est rectangle, que le problème consiste à calculer la longueur de l’hypoténuse et que les longueurs des deux autres côtés sont connus, alors le résultat est égal à la racine carrée de la somme des carrés des deux autres côtés). La grande majorité des travaux des tuteurs intelligents se positionnent dans le cadre des domaines bien définis.

La modélisation des domaines mal définis est plus problématique, puisqu’il s’agit de représenter un ensemble de règles subjectives (e.g. en guitare classique, le pincement d’une corde doit se faire avec la pulpe du doigt, de manière à ce que le son résultat soit “clair” et “rond”). La définition de domaines “mal définis” est elle-même imprécise tant elle regroupe de domaines différents. Outre la taxonomie de Lynch précédemment énoncée, d’autres classifications ont été proposées. (Simon, 1973) identifie les domaines mal définis depuis la comparaison entre la résolution de problèmes chez les humains et les solveurs artificiels, et propose les conditions suivantes :

- un domaine est mal défini si les instructions ou informations nécessaires à sa résolution sont incomplètes ou vagues ;
- un domaine est mal défini si les critères déterminant l’atteinte du but sont complexes et imprécis ;
- un domaine est mal défini s’il n’y a pas de stratégie claire permettant la résolution.

Les dernières taxonomies proposées font état d’un continuum entre les domaines bien et mal définis plutôt que d’une frontière claire (Mitrovic et al., 2009; Le et al., 2013). Nous adopterons plus simplement la définition proposée par Fournier-Viger (Fournier-Viger et al., 2010b) : “Un domaine mal défini est un domaine qui pose de nouveaux défis où les approches classiques ne peuvent être appliquées.”

Par rapport à toutes ces définitions, nous plaçons l’apprentissage de gestes parmi les domaines mal définis. En effet, il n’existe pas de geste “idéal” que tout apprenant devrait copier, et même si tel était le cas, les différences morphologiques existant entre chaque individu amèneraient des différences non négligeables entre deux gestes en apparence identiques. En outre, si l’évaluation d’un geste devait se tenir à un ensemble de règles ou de contraintes, celles-ci seraient au mieux complexes, au pire subjectives. Si l’on s’en tient à la définition de Fournier-Viger énoncée précédemment, on peut donc affirmer que la modélisation, et a fortiori l’adaptation, dans le cadre d’une activité gestuelle amène de nombreux défis, d’où la catégorisation de l’apprentissage de gestes comme un domaine mal défini.

Bien que nous nous positionnions dans un domaine d’apprentissage mal défini, il est intéressant d’étudier la littérature portant sur la modélisation d’apprenant dans le cadre de domaines bien définis. En effet, ces derniers furent les premiers à être étudiés, et constituent encore aujourd’hui un domaine de recherche très actif. Les problématiques inhérentes à ces travaux se sont enrichies au fur et à mesure des avancées scientifiques, si bien que de nombreuses approches proposées dans le cadre des domaines mal définis reposent sur des approches initialement conçues pour des domaines d’apprentissage bien définis. Nous présentons donc dans la suite de ce document les principales approches

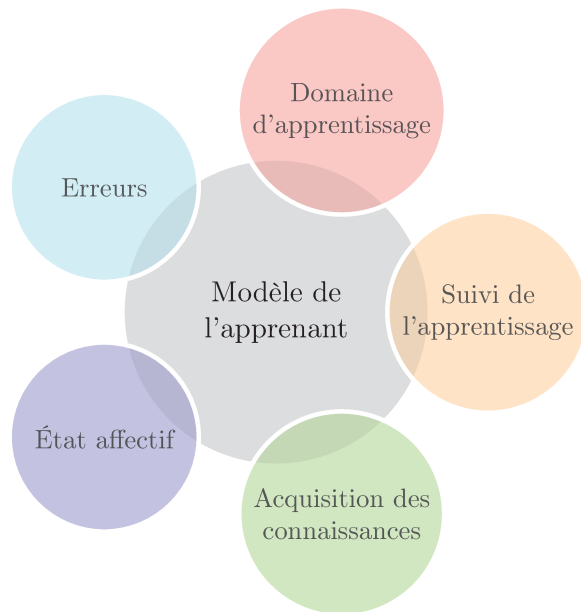


FIGURE 2.2 – Composantes possibles du modèle de l'apprenant (*student model*).

modélisant l'apprenant dans ces deux cadres, suite à quoi nous présentons notre approche.

2.3 Modélisation de l'apprentissage : domaines bien définis

La majorité des tuteurs intelligents étudient des domaines bien définis. Nous décrivons dans cette section les principales approches résultant de ces travaux. Pour chaque approche, nous étudions la possibilité de transposer la méthode utilisée dans le cadre de domaines mal définis. La première partie de cette section porte sur les modélisations de l'apprenant, c'est à dire la représentation haut niveau permettant d'analyser son état général. Dans une seconde partie, nous décrivons la modélisation des capacités, c'est à dire les stratégies utilisées par les tuteurs intelligents pour étudier le niveau de l'apprenant sur chacune des capacités faisant partie du domaine d'apprentissage.

2.3.1 Représentation de l'apprenant

2.3.1.1 Approche orientée stéréotypes

La représentation par stéréotypes, proposée pour la première fois par (Rich, 1979), considère l'apprenant comme un ensemble de propriétés. Il existe deux

types de représentation par stéréotypes : l'approche *fixe* et l'approche *par défaut*.

L'approche fixe repose sur l'hypothèse générale que l'ensemble des apprenants appartenant au même ensemble de propriétés (en d'autres mots, les apprenants ayant les mêmes stéréotypes), vont agir et apprendre de la même manière. Les propriétés peuvent être liées au niveau de l'apprenant : dans (Wheeler and Regian, 1999), l'environnement propose différents niveaux ordonnés par difficulté. La propriété modélisant l'apprenant correspond à son niveau, qui augmente lorsqu'il réussit à répondre à deux exercices consécutifs sans intervention de l'environnement. Les propriétés peuvent également être liées à la personnalité de l'apprenant (e.g. introverti/extraverti) ou à ses caractéristiques propres (âge, genre, classe sociale) (Narciss et al., 2014).

L'approche par défaut utilise les stéréotypes afin d'initialiser le système. Il permet à l'environnement de ne pas partir de zéro et de proposer un premier niveau d'adaptation, avant de préciser le modèle de l'apprenant au fur et à mesure que celui-ci interagit avec l'environnement en utilisant une autre approche comme un modèle par recouvrement (Section 2.3.1.2).

Avantages de l'approche L'approche fixe peut être utile dans le cadre de domaines simples, ou lorsque certaines caractéristiques différencient très fortement les groupes d'apprenants. Elle peut également être utile dans le cadre de domaines mal définis, lorsqu'il est impossible de fournir une modélisation précise de l'activité de l'apprenant. L'approche par défaut permet quant à elle de proposer une solution au problème du *cold start*, lorsque l'environnement, face à un nouvel apprenant avec qui il n'a encore jamais interagi, ne peut que décider aléatoirement du comportement à adopter.

Limites de l'approche L'hypothèse sur laquelle se fonde l'approche fixe est critiquable puisqu'elle considère que l'ensemble des apprenants possédant les mêmes caractéristiques agissent et apprennent de la même manière. Elle ne doit donc être utilisée que dans des domaines très spécifiques, ou conjointement avec d'autres types d'approche.

2.3.1.2 Approche orientée domaine

L'approche orientée domaine vise à représenter l'ensemble du domaine, c'est à dire l'ensemble des solutions d'un problème donné. Soit \mathcal{P} l'ensemble qui regroupe l'ensemble des actions possibles de l'apprenant, et $\mathcal{D} \subset \mathcal{P}$ l'ensemble des actions relevant du domaine (ce sont les actions acceptables). L'évaluation d'une action comme acceptable dépend du contexte (i.e. la solution $x = 4$ est acceptable comme résultat du problème $x + 2 = 6$, elle ne l'est pas comme résultat du problème $x + 1 = 6$). Si on note $\mathcal{D}_x \subset \mathcal{D}$ l'ensemble des actions solutions dans le contexte (ou la tâche) x , l'approche orientée domaine va alors chercher à énumérer les règles définissant les actions de \mathcal{D}_x pour un contexte x donné. Ces règles peuvent être issues d'une analyse de l'activité, permettant d'illustrer l'opération, ou l'ensemble des opérations, nécessaires à l'accomplissement de la tâche (e.g. lors de l'addition de deux nombres à deux chiffres, il est

nécessaire de sommer les chiffres des unités, puis d'ajouter la retenue si cette somme est supérieure à neuf, puis de sommer les chiffres des dizaines ainsi que l'éventuelle retenue).

Cette approche est notamment utilisée dans le cadre des modèles par recouvrement⁴ qui représentent l'apprenant comme un sous-ensemble du domaine d'expertise (Figure 2.3). Les modèles par recouvrement font partie des principaux modèles utilisés. Initialement proposés par (Carr and Goldstein, 1977), ces modèles ont depuis beaucoup évolué. Représentant initialement l'acquisition de connaissances par de simples valeurs booléennes (l'apprenant sait ou ne sait pas), ils permettent désormais de représenter de manière qualitative le niveau de connaissance (dans quelle mesure l'apprenant sait, ou sait faire). L'approche proposée par (Brusilovsky and Millán, 2007) évalue pour cela l'apprenant par la probabilité qu'il ait acquis la maîtrise de la connaissance.

L'approche visant à représenter l'apprenant comme un sous-ensemble du domaine d'expertise permet d'identifier aisément les manques de l'apprenant : si l'on considère l'ensemble \mathcal{E} des compétences acquises par l'apprenant, et pour rappel \mathcal{D}_x représente l'ensemble des connaissances formant l'expertise, alors l'ensemble des connaissances qui restent à acquérir est l'ensemble $\{a | a \in \mathcal{D}_x \setminus \mathcal{E}\}$. (Brusilovsky and Millán, 2007) considèrent cette différence comme le niveau d'expertise de l'apprenant.

Le modèle par recouvrement différentiel (Holt et al., 1994) est une extension du modèle par recouvrement. Alors que le modèle classique ne propose que deux couches (l'ensemble définissant le domaine et l'ensemble définissant l'apprenant), le modèle différentiel intègre également un ensemble constitué des connaissances attendues. Cette extension permet de n'analyser qu'une partie de l'apprentissage. Dans le cas où un exercice se focalise sur une notion x et ne considère pas la notion y , l'environnement pourra évaluer l'apprenant uniquement par rapport à x , y ne faisant alors pas partie des connaissances attendues à ce stade de l'apprentissage (Brown and Burton, 1978).

Avantages de l'approche Cette représentation permet de définir formellement le modèle. Depuis la performance de l'apprenant, il est possible de le modéliser en inférant sur les règles définissant le domaine. Cette approche permet de ne modéliser qu'une partie du domaine.

Limites de l'approche Si cette approche se prête particulièrement aux domaines bien définis, il est plus complexe de la transposer dans le cadre de domaines mal définis, puisqu'il est par essence complexe, voire impossible, de décomposer ces domaines sous forme d'items d'apprentissage dont on pourrait mesurer la maîtrise. Un autre inconvénient de cette approche est qu'elle est totalement dépendante du cas d'étude : un environnement basé sur une telle approche ne peut être réutilisé dans un autre domaine. Finalement, cette approche ne considère que les actions faisant partie du domaine. Les erreurs, ou les idées fausses⁵, c'est à dire l'ensemble $\{a | a \in \mathcal{P} \cap \overline{\mathcal{D}}\}$, n'y est ainsi pas représenté.

4. Traduction de "overlay models".

5. Traduction de "misconception".

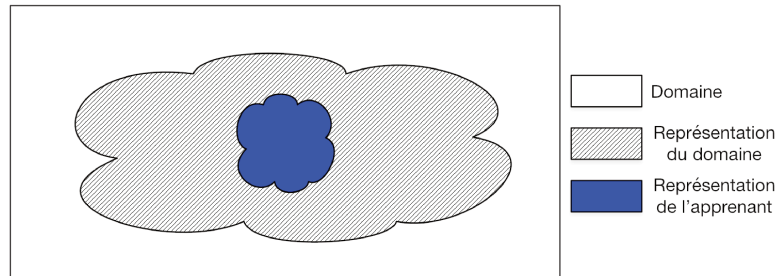


FIGURE 2.3 – Représentation de l'apprenant selon un modèle par recouvrement.

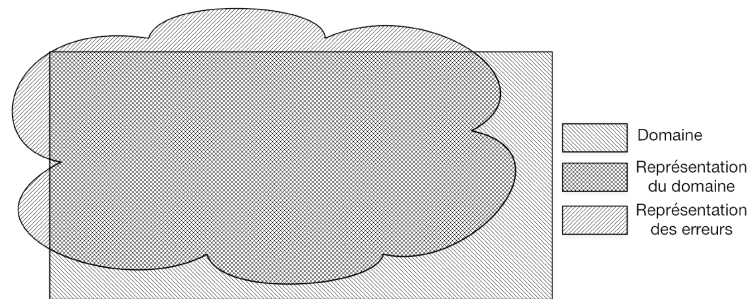


FIGURE 2.4 – Représentation de l'apprenant selon un modèle combinant les approches de représentation par recouvrement et par erreurs.

2.3.1.3 Approche orientée erreurs

La modélisation des erreurs⁶ permet de pallier l'un des inconvénients des modèles par recouvrement en représentant les erreurs ainsi que les idées fausses que peut développer l'apprenant (Figure 2.4).

L'apparition de ce modèle fut motivée par un intérêt pédagogique : certains motifs d'erreur peuvent illustrer chez l'apprenant un raisonnement erroné ; l'étude de telles erreurs peut permettre de préciser le modèle de l'apprenant et d'apporter une aide adaptée. Prenons le problème du développement du terme $(a+b)^2$: un enseignant ne considérera pas de la même manière la réponse $a+b^2$ de la réponse a^2+b^2 . L'idée est de modéliser un ensemble d'actions $\mathcal{F} \subset \overline{\mathcal{D}}$ qui correspond à des actions n'appartenant pas au domaine de compétence. Cette modélisation, qui s'ajoute à la modélisation de l'ensemble d'actions appartenant au domaine \mathcal{D} permet alors une modélisation plus riche de l'ensemble \mathcal{P} des actions possibles de l'apprenant.

La modélisation des erreurs peut provenir de deux sources : une énumération et une classification empirique des erreurs par un expert (principe utilisé par

6. On trouve également dans la littérature les termes de "bug library" ou "Perturbation model".

le tuteur Buggy (Brown and Burton, 1978), ou une approche générative qui considère les erreurs comme des extrapolations des concepts étudiés (Mayo, 2001).

Les modèles d'erreur sont communément utilisés, en conjonction avec d'autres types de modèles. On peut notamment citer (Baschera and Gross, 2010) qui utilisent une telle approche afin de fournir des aides pertinentes dans le cadre de l'apprentissage de l'orthographe. Dans un exercice demandant à l'apprenant d'écrire le mot "Unmut"⁷, l'environnement identifie différents types d'erreur, pouvant être liés par exemple à une erreur de typographie⁸, à une erreur de dyslexie ou de dysphonie. L'environnement peut alors adopter le comportement correspondant au type d'erreur identifié.

Avantages de l'approche Utilisée en parallèle d'une représentation du domaine, la représentation des erreurs permet de modéliser plus précisément l'activité de l'apprenant, en prenant en compte non seulement ses compétences vis-à-vis du domaine, mais également ses erreurs. Ce type de modèle peut également permettre d'identifier les idées fausses de l'apprenant, et de fournir alors des aides plus pertinentes.

Limites de l'approche Qu'elles proviennent de l'énumération d'un expert ou d'une approche générative, la formalisation des erreurs est un processus long, et il est compliqué d'aboutir à une énumération exhaustive des erreurs. C'est pourquoi cette approche n'est utilisée que dans le cadre de domaines d'apprentissage simples et bien définis.

2.3.1.4 Autres approches

Des revues étendues des différents modèles proposés dans le domaine des tuteurs intelligents sont disponibles dans les états de l'art établis par (Kobsa, 2001) pour la période 1978-2000, et (Chrysafiadi and Virvou, 2013) pour la période 2000-2013. Ce dernier article discute notamment les travaux portant sur la modélisation des aspects affectifs. Ces aspects sont de plus en plus étudiés, mobilisant notamment les avancées technologiques permettant la reconnaissance des expressions faciales (D'Mello and Graesser, 2010) et le suivi du regard (D'Mello et al., 2012) afin d'en déduire le degré d'attention et/ou de motivation de l'apprenant. Étant néanmoins plus éloignés des travaux que nous présentons ici, nous renvoyons le lecteur intéressé par ces aspects vers les lectures suscitées. L'ouvrage (Woolf, 2009) fournit également une description détaillée des différentes approches utilisées pour la création d'environnements d'apprentissage intelligents.

7. "Unmut" est un mot allemand signifiant mécontentement, ou colère.

8. À noter que contrairement à la disposition des touches sur les claviers français (disposition AZERTY), les touches "M" et "N" sont adjacentes sur les claviers allemands (disposition QWERTY), ce qui tend à favoriser les erreurs de typographie lors de l'écriture du mot "unmut".

2.3.2 Modélisation des capacités

Dans la section 2.3.1, nous avons énuméré les premières approches permettant de représenter l'apprenant, en fonction de l'ensemble des connaissances à acquérir (domaine), de ces erreurs ou fausses idées, ou en fonction de ses caractéristiques propres. Si ces considérations haut niveau permettent de situer la façon dont l'apprenant est considéré, et permettent, en inférant sur les différences établies, de fournir des retours adaptés à l'apprenant (en somme, ces techniques permettent de représenter une vision à long terme de l'apprenant (Mayo, 2001)), de nombreuses études travaillent sur des méthodes permettant de préciser cette représentation de l'apprenant en modélisant la façon dont chaque compétence peut être acquise (une représentation à court terme selon (Mayo, 2001)). Nous décrivons dans cette partie trois approches distinctes, bien qu'ayant de nombreuses similitudes, et des résultats comparables (Mitrovic et al., 2003) lorsqu'ils peuvent être mis en oeuvre : les tuteurs intégrant un système expert, les tuteurs cognitifs, et les tuteurs par contraintes.

2.3.2.1 Systèmes experts

L'une des premières approches permettant le diagnostic des connaissances consiste à formuler un ensemble de règles intégrées au sein d'un système expert. Le système doit alors être capable d'admettre en entrée la performance de l'apprenant, et d'inférer sur les règles pour établir un diagnostic. Exemple : soit un triangle de côté adjacent a , de côté opposé b et d'hypoténuse c , on demande la longueur du côté a connaissant celles des côtés b et c . Si la réponse est $\sqrt{c^2 - b^2}$ alors l'apprenant sait identifier les différents côtés d'un triangle rectangle, et sait appliquer le théorème de Pythagore. Si par contre la réponse est $\sqrt{c^2 + b^2}$, alors l'apprenant connaît la formule de Pythagore mais ne sait pas identifier les côtés adjacents, opposés et l'hypoténuse.

Le système expert peut être utilisé pour générer des solutions expertes. Ces solutions sont comparées avec les solutions de l'apprenant. Ces solutions expertes sont alors utilisées à des fins de démonstration, ou pour suggérer des étapes de résolution. C'est dans cette optique que l'approche par système expert est utilisée dans le cadre du tuteur Guidon (Clancey, 1984). Il permet de comparer les solutions proposées par l'apprenant avec les solutions "idéales" dans les tuteurs AutoTutor (Graesser et al., 2000) et DesignFirst-ITS (Moritz and Blank, 2008).

Avantages Les systèmes experts s'intègrent alors dans des modèles par recouvrement, et permettent de fournir des aides plus riches que celles fournies par d'autres approches.

Inconvénients En contrepartie, l'utilisation d'un système expert n'est pas générique et ne peut être transféré dans d'autres domaines. Le développement d'un tel système est très coûteux et n'est possible que dans des domaines bien définis. Les inférences faites par le système sont en outre parfois peu compréhensibles, et alors inutilisables dans un but pédagogique.

<p>Si le but est de trouver un angle d'un triangle isocèle ABC avec $AB=AC$,</p> <p> Si l'angle A est inconnu,</p> <p> Alors donner à l'angle A la valeur de l'angle B.</p> <p>Si le but est de trouver un angle d'un triangle isocèle ABC avec $AB=AC$,</p> <p> Si les angles A et B sont connus,</p> <p> Alors donner à l'angle C la valeur $180-A-B$.</p>

FIGURE 2.5 – Deux règles de production pour la résolution du problème : trouver la valeur d'un des angles d'un triangle isocèle, adapté depuis (Mitrovic et al., 2003; Desmarais and Baker, 2012).

2.3.2.2 Tuteurs cognitifs

Les tuteurs cognitifs sont basés sur la théorie ACT-R (Anderson and Lebiere, 1998) qui spécifie l'organisation et le fonctionnement de la mémoire et de l'apprentissage. La mémoire y est divisée en deux unités distinctes : la mémoire déclarative, et la mémoire procédurale. La mémoire déclarative se présente sous la forme d'un ensemble de connaissances factuelles ou explicites, tandis que la mémoire procédurale intègre des processus ou schèmes de raisonnements implicites ou tacites (Lallé, 2013). Selon l'auteur, l'apprentissage de nouvelles connaissances est un processus nécessitant plusieurs étapes : la première consiste en l'acquisition de connaissances déclaratives (savoir), qui se traduit par la suite par l'acquisition de connaissances procédurales (savoir-faire). Dans une dernière phase, les connaissances procédurales s'intègrent à une mécanique et l'apprenant devient expert. Cette théorie se rapproche de la théorie de l'apprentissage moteur décrite dans le chapitre 1.

Les tuteurs cognitifs se fondent sur deux composants : le *model-tracing* et le *knowledge-tracing*. Le *model-tracing* est un modèle générique de l'apprenant, qui cherche à situer l'apprenant par rapport au domaine grâce à des règles de production (e.g. figure 2.5, approches présentées dans la section 2.3.1).

Le *knowledge-tracing* est un modèle spécifique à l'apprenant, qui se concentre sur les performances successives de celui-ci afin d'évaluer son niveau d'acquisition de connaissance. Les modules de *model-tracing* et de *knowledge-tracing* permettent au tuteur cognitif d'interpréter le comportement de l'apprenant. Le *knowledge-tracing* est des sujets les plus étudiés, et de nombreuses évolutions ont été apportées depuis le premier modèle (Corbett and Anderson, 1994). L'une des approches ayant reçu le plus d'attention, le *Bayesian knowledge-tracing* (BKT) (Käser et al., 2014) utilise des modèles de Markov à états cachés (HMM) (Rabiner, 1989) pour modéliser le passage d'un état novice (l'apprenant ne maîtrise pas) à l'état expert (l'apprenant maîtrise).

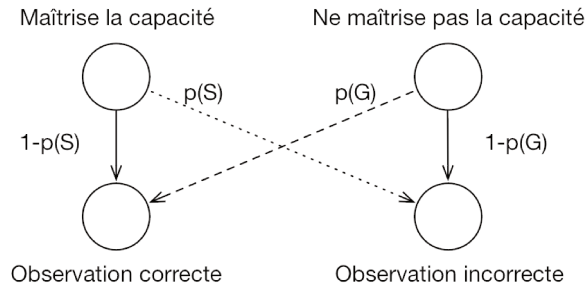


FIGURE 2.6 – Modèle de Markov caché utilisé dans le cadre du Bayesian knowledge-tracing, adapté depuis (Beck and Sison, 2006)

2.3.2.2.1 Bayesian knowledge-tracing

Les modèles de Markov à états cachés sont une extension des chaînes de Markov. Les chaînes de Markov correspondent à un simple graphe d'états, doté d'une fonction de transition probabiliste. A chaque pas de temps, le modèle subit une transition qui va potentiellement modifier son état. Cette transition permet donc au système modélisé d'évoluer selon une loi de probabilité connue. Dans un modèle de Markov à états cachés, l'état du système n'est pas directement observable, et est caché par un processus d'observation.

Les HMM comportent trois types de paramètres : les probabilités initiales, les probabilités d'émission et les probabilités de transition. Dans le cadre du BKT, les probabilités d'émission sont définies par la probabilité de commettre une erreur à propos d'un concept pourtant connu (*slip probability* $p(S)$) et la probabilité de réussir une capacité pourtant inconnue (*guess probability* $p(G)$). Les probabilités de transitions sont définies par la probabilité d'acquérir une capacité (*learning probability* $p(L)$), et la probabilité d'oublier, de perdre une capacité préalablement acquise (*forgetting probability* $p(F)$). Cette dernière probabilité est souvent établie comme constante et égale à zéro dans le cadre des BKT. Le dernier paramètre permettant de définir le modèle est la probabilité de maîtriser une capacité a priori ($p(L_0)$). La figure 2.6 illustre le modèle. La probabilité initiale que l'apprenant u maîtrise la capacité j est donnée par l'équation (2.1a). En fonction de la capacité de l'apprenant à appliquer de manière satisfaisante la capacité j (on note $x = 1$ si le résultat est satisfaisant, $x = 0$ dans le cas contraire), la probabilité conditionnelle exprimant la probabilité que la capacité ait été précédemment acquise est mise à jour par les équations (2.1b) et (2.1c). La probabilité d'acquisition de la capacité est mise à jour via l'équation (2.1d). Enfin, la probabilité $p(C_t)$ que l'apprenant maîtrise

la compétence C à l'instant t est calculée par l'équation (2.1e).

$$p(L_1)_{j,u} = p(L_0)_j \quad (2.1a)$$

$$p(L_{t-1}|x=1)_{j,u} = \frac{p(L_{t-1})_{j,u} \cdot (1 - p(S)_j)}{p(L_{t-1})_{j,u} \cdot (1 - p(S)_j) + (1 - p(L_{t-1})_{j,u}) \cdot p(G)_j} \quad (2.1b)$$

$$p(L_{t-1}|x=0)_{j,u} = \frac{p(L_{t-1})_{j,u} \cdot p(S)_j}{p(L_{t-1})_{j,u} \cdot p(S)_j + (1 - p(L_{t-1})_{j,u}) \cdot p(G)_j} \quad (2.1c)$$

$$p(L_t)_{j,u} = p(L_{t-1}|x)_{j,u} + (1 - p(L_{t-1}|x)_{j,u}) \cdot p(T)_j \quad (2.1d)$$

$$p(C_t)_{j,u} = p(L_t)_{j,u} \cdot (1 - p(S)_j) + (1 - p(L_t)_{j,u}) \cdot p(G)_j \quad (2.1e)$$

Dans le modèle standard du BKT, chaque capacité est représentée par un modèle de Markov à états cachés tel que présenté ci-dessus. Les paramètres du modèle peuvent être choisis par des experts, mais sont généralement optimisés en utilisant des données d'apprenants et des méthodes telles que l'espérance-maximisation, le gradient conjugué (Corbett and Anderson, 1994) ou le brute-force discret (Baker et al., 2008). L'utilisation d'un tel modèle peut s'inclure dans une approche par recouvrement (Section 2.3.1.2) et permet de représenter plus finement l'apprenant.

2.3.2.2.2 Analyse des facteurs d'apprentissage

Une seconde approche permettant le knowledge-tracing, également basée sur la théorie de l'ACT-R, est l'analyse des facteurs d'apprentissage (*LFA*⁹) (Cen et al., 2006). Cette théorie se base également sur les résultats de (Newell and Rosenbloom, 1981) qui montrent une relation entre le taux d'erreur et le nombre d'essais (Équation (2.2)).

$$\epsilon = \beta X^\gamma \quad (2.2)$$

avec :

- ϵ le taux d'erreur ;
- X le nombre d'essais ;
- β le taux d'erreur sur le premier essai qui décrit la difficulté intrinsèque de la capacité ;
- γ le paramètre d'apprentissage qui illustre la difficulté que représente l'acquisition de la capacité.

La courbe d'apprentissage représentée par l'équation (2.2) est étendue dans l'analyse des facteurs d'apprentissage afin de modéliser l'apprentissage de grands ensembles d'apprenants. L'apprentissage est alors modélisé par une régression logistique illustrée par l'équation (2.3), qui se base sur les hypothèses suivantes (Cen et al., 2006) :

- Les apprenants ont des connaissances différentes ;

9. Learning Factors Analysis

- Le paramètre d'apprentissage dépend principalement de la nature de la compétence. Les différences entre les apprenants sont négligeables.

$$\ln\left(\frac{p_{t,u}}{1-p_{t,u}}\right) = \theta_u + \sum \beta_j Q_{jt} + \sum Q_{jt}(\gamma_j N_{ju}) \quad (2.3)$$

avec

- u l'apprenant ;
- t une étape du problème ;
- j une connaissance ;
- $p_{t,u}$ la probabilité que l'apprenant performera correctement lors de l'étape t ;
- Q une matrice avec Q_{jt} une case de la matrice indiquant si la connaissance j est associée à l'étape t ;
- N_{ju} le nombre d'étapes de résolution de problèmes passés où l'apprenant i a dû mobiliser la connaissance j ;
- β_j un coefficient représentant la difficulté de la capacité j ;
- γ_j un coefficient représentant l'importance du temps nécessaire pour apprendre la capacité j ;
- θ_u un coefficient représentant la facilité d'apprentissage de l'apprenant.

L'analyse de facteurs d'apprentissage connaît un succès important, et de nombreuses variantes sont proposées. Parmi celles-ci, on peut notamment citer l'analyse des facteurs de performance (*PFA*¹⁰) (Pavlik Jr et al., 2009) qui raffine l'équation (2.3) pour prendre en compte l'historique du nombre d'erreurs faites par l'apprenant par rapport à la compétence j (Équation (2.4)).

$$\ln\left(\frac{p_{t,u}}{1-p_{t,u}}\right) = \theta_u + \sum \beta_j Q_{jt} + \sum Q_{jt}(\mu_j S_{ju} + \rho_j F_{tu}) \quad (2.4)$$

où la notation reprend celle de l'équation (2.3) et

- S_{ju} correspond au nombre d'étapes de résolutions de problèmes précédemment effectuées par l'apprenant de façon correcte ;
- F_{tu} représente le nombre d'étapes de résolutions de problèmes précédemment effectuées par l'apprenant de façon incorrecte ;
- μ est un coefficient sur le bénéfice des précédents succès lors de l'apprentissage ;
- ρ est un coefficient sur le bénéfice des précédents échecs lors de l'apprentissage.

Comme pour le BKT, les paramètres sont optimisés en maximisant la log-vraisemblance (Pavlik Jr et al., 2009).

Avantages Les tuteurs cognitifs permettent une analyse très précise de l'état d'apprentissage de l'apprenant. Ils permettent d'étudier la probabilité que chacune des capacités soit acquise. Les tuteurs cognitifs ont également la faculté de pouvoir proposer des retours immédiats, puisqu'ils suivent chaque étape de l'apprentissage de l'apprenant.

10. Performance Factors Analysis

<p>Si la condition de pertinence est vérifiée, Alors Si la condition de satisfaction n'est pas vérifiée, Alors il y a une erreur.</p>
--

FIGURE 2.7 – Logique du tuteur par contrainte en fonction des conditions de pertinence et de satisfaction

Inconvénients Les tuteurs cognitifs nécessitent une définition exhaustive et précise du domaine d'apprentissage en règles de production et en énumération d'erreurs possibles. Ils ne sont donc applicables que dans un nombre restreint de domaines bien définis.

2.3.2.3 Tuteurs par contraintes

Les tuteurs par contraintes (*CBM*¹¹) ont été proposés par Ohlsson (Ohlsson, 1994) pour faire face à l'inefficacité¹² des solutions permettant la représentation de l'apprenant. En effet, les approches et stratégies précédemment décrites s'avèrent très complexes à mettre en oeuvre puisqu'il est nécessaire, comme précisé dans les différentes sections précédentes, de pouvoir décrire chaque composante de l'apprentissage. En outre, une fois le domaine décrit, les inférences nécessaires au suivi de l'apprenant peuvent devenir coûteuses. Ohlsson propose alors de réduire l'étendue du domaine à modéliser aux états ayant un intérêt pédagogique. Alors qu'une performance est considérée comme correcte par les tuteurs cognitifs lorsqu'elle correspond à la règle énoncée, le tuteur par contrainte considère comme correcte une performance qui ne violera aucune contrainte. Les CBM sont constitués de deux types de condition :

- les **conditions de pertinence**¹³, contraignant le cadre du problème étudié ;
- les **conditions de satisfaction**¹⁴, contraignant la réponse au problème défini par la contrainte de pertinence.

Le modèle du domaine est alors un ensemble de couples $\langle C_r, C_s \rangle$ de contraintes de pertinence et de contraintes de satisfaction qui vérifient la logique présentée dans la figure 2.7.

De nombreux tuteurs intelligents sont basés sur une approche par contraintes. Parmi les plus connus, on peut citer SQL tutor (Mitrovic and Ohlsson, 1999) qui enseigne le langage de requête de bases de données SQL ; CAPIT (Mayo, 2001), qui enseigne la ponctuation ; KERMIT (Suraweera and Mitrovic, 2002) puis son extension EER-Tutor (Weerasinghe and Mitrovic, 2011), qui aident à la création de diagrammes de type entité-association ; J-LATTE (Holland

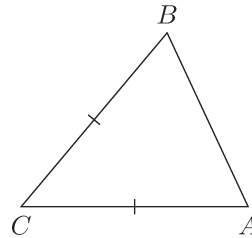
11. Constraint-based models

12. Comme traduction de "intractable"

13. Relevance condition

14. Satisfaction condition

L'angle A mesure 65°
Quelle est la valeur de l'angle C ?



Règles de Knowledge Tracing

Deux règles de production

Si le but est de trouver un angle d'un triangle isocèle ABC avec $AB = AC$ et si l'angle A est inconnu,

Alors donner à l'angle A la valeur de l'angle B .

Si le but est de trouver un angle d'un triangle isocèle ABC avec $AB = AC$, et que les angles A et B sont connus,

Alors donner à l'angle C la valeur $180 - A - B$.

Production erronée

Si le but est de trouver l'angle d'un triangle isocèle ABC avec $AB = AC$, si les angles A et C sont situés sur le bas du triangle et si A est connu,

Alors donner à l'angle C la valeur de l'angle A .

Contraintes

C_{r1} L'un des angles de base d'un triangle isocèle est connu (θ_1), et l'apprenant doit calculer le second angle de base (θ_2).

C_{s1} La valeur de l'angle θ_1 est égale à la valeur de l'angle θ_2 .

C_{r2} L'un des angles de base d'un triangle isocèle est connu (θ_1), et l'apprenant a calculé la valeur d'un autre angle (θ_2), et $\theta_2 = \theta_1$.

C_{s2} θ_2 est un angle de base.

C_{r3} Deux angles d'un triangle isocèle sont connus (θ_1 et θ_2), et l'apprenant a calculé la valeur du troisième angle θ_3 .

C_{s3} La valeur de θ_3 est égale à $180 - \theta_1 - \theta_2$.

FIGURE 2.8 – Comparaison des règles de knowledge-tracing et des contraintes de tuteurs par contraintes, adapté depuis (Mitrovic et al., 2003; Desmarais and Baker, 2012)

et al., 2009), qui enseigne le langage de programmation Java. La figure 2.8 illustre les différences entre règles de knowledge-tracing et contraintes des tuteurs par contraintes dans le cadre de la recherche de la valeur d'un angle d'un triangle isocèle.

Avantages La création de tuteurs par contraintes est moins coûteuse en terme de temps, et moins difficile, que celle d'un tuteur cognitif. Cette différence s'explique par la nature même de la représentation : lorsqu'un tuteur cognitif demande l'établissement de règles de production et d'erreurs définissant totalement le domaine, les tuteurs cognitifs ne le représentent qu'à l'aide de contraintes représentant sa frontière.

Inconvénients La contrepartie de cette représentation plus légère est une modélisation moins précise de l'état de l'apprenant, et donc une capacité moindre à pouvoir fournir des aides appropriées.

Les différences entre tuteurs cognitifs et par contraintes sont néanmoins à relativiser par le temps passé à la définition du domaine. Il est établi que dans le cas de tuteurs cognitifs, 200 à 300 heures de travail sont nécessaires pour définir une heure d'activité de l'apprenant. Il est alors possible, pour réduire cette durée, de réduire le nombre de règles de production/d'erreurs. Mais la précision du modèle en est alors affectée. À l'opposé, il est possible d'affiner la granularité des conditions de pertinence et de satisfaction qui constituent le modèle par contraintes. Le modèle y gagne alors en précision, mais le coût nécessaire à la définition du domaine, et le coût relatif aux inférences nécessaires à la compréhension de l'activité de l'utilisateur peuvent alors devenir prohibitifs.

2.3.3 Conclusions

Dans cette partie, nous avons décrit les principales approches utilisées pour effectuer le diagnostic de l'état de l'apprenant dans le cadre de domaines bien définis. L'approche par stéréotypes modélise les caractéristiques propres de l'apprenant, et peut donc aussi bien être utilisée dans le cadre de domaines mal définis. Elle n'offre par contre pas une précision fine, puisqu'elle considère de la même manière tous les apprenants ayant les mêmes caractéristiques, indépendamment de leur activité. Les modèles par recouvrement permettent de modéliser l'apprentissage de l'apprenant sur le long terme, en le considérant comme un sous-ensemble du domaine d'expertise. Les modèles par erreurs permettent d'enrichir le modèle par recouvrement en lui permettant d'étendre ses considérations aux actions n'appartenant pas au domaine. Enfin, nous avons décrit plusieurs stratégies permettant de modéliser l'acquisition de la compétence : les systèmes experts et les tuteurs cognitifs, qui peuvent être très coûteux à mettre en place mais qui permettent une modélisation fine et des retours immédiats ; et les tuteurs par contraintes qui constituent une alternative aux tuteurs cognitifs, et qui modélisent la frontière entre ce qui est correct et ce qui ne l'est pas. Dans l'ensemble de ces approches, les aides ou retours proposés à l'apprenant se basent sur la modélisation puis sur une expertise : depuis sa connaissance

du domaine et des théories de l'apprentissage, l'expert établit le lien entre une situation du modèle et une aide que l'environnement devra proposer.

Dans la prochaine section, nous nous intéressons aux modèles permettant le diagnostic et l'adaptation dans les domaines mal définis.

2.4 Modélisation de l'apprenant : domaines mal définis

Les stratégies décrites dans la section 2.3 ne sont pas directement applicables dans les environnements se positionnant dans les domaines mal définis, du fait de la difficulté à définir et/ou à mesurer les caractéristiques constituant le modèle de l'apprenant et les heuristiques de sélection d'actions pédagogiques. Pourtant, plusieurs méthodes de modélisation de l'apprenant et de sélection d'actions pédagogiques ont été proposées ces quinze dernières années dans le cadre de domaines mal définis. Ces méthodes peuvent être divisées en deux catégories (Fournier-Viger et al., 2010b, 2013) : les tuteurs s'inspirant des stratégies adoptées par les tuteurs humains afin d'adapter leur comportement malgré une représentation dégradée de l'apprenant ; et les tuteurs réduisant le domaine mal défini en un sous-ensemble suffisamment bien défini pour que les techniques présentées dans la section 2.3 puissent être utilisées. Toutes ces approches peuvent s'inscrire dans la stratégie dite "multi-paradigmes" qui vise à profiter des avantages de plusieurs stratégies en les combinant.

2.4.1 Étude de cas

Le principe de l'approche basant la sélection d'actions pédagogiques sur les études de cas (*CBR*¹⁵) est simple : si l'on a accès à un historique de situations face auxquelles un comportement a été adopté, et si l'influence de ces comportements sur l'apprentissage de l'apprenant a été évaluée, alors il est possible de réutiliser ces expériences face à une nouvelle situation. Cette stratégie repose sur l'hypothèse qu'un même feedback proposé dans deux situations similaires provoquera dans chacune de ces situations une influence similaire. Ce paradigme reproduit l'expérience d'un tuteur humain, qui au fur et à mesure qu'il pratique distingue les conseils ou retours à prodiguer à l'apprenant en fonction de son état (e.g. son niveau de pratique, sa motivation, son niveau de fatigue). L'étude de cas est un processus cyclique composé de quatre étapes (Aamodt and Plaza, 1994) :

1. la **récupération** : il s'agit de récupérer les cas précédents qui sont similaires au cas en cours d'étude. Des méthodes inductives peuvent être utilisées pour permettre de retrouver des cas similaires en fonction de critères définis ;
2. la **réutilisation** : les solutions des cas précédents ayant la plus grande similarité avec le cas courant sont réutilisées pour construire la solution. Il est parfois nécessaire de transformer les solutions dans le but de les

15. Case-based Reasoning.

généraliser, ou au contraire de les spécialiser, avant qu'elles ne répondent au problème courant ;

3. la **révision** : la solution donnée pour le cas courant est évaluée, et l'information déterminant si cette solution est acceptable ou non est récupérée ;
4. la **sauvegarde** : le problème courant et la solution qui lui a été apportée sont sauvegardés, et appartiennent désormais à l'historique des cas précédents.

Cette approche est notamment utilisée dans le tuteur CATO (Aleven, 2003). L'étude de cas permet de valoriser les connaissances produites depuis les interactions passées, qui permet au système d'affiner sa décision au fur et à mesure que la base de cas grandit. Cette méthode repose néanmoins sur la précision de la mesure de similarité utilisée, et nécessite donc de pouvoir dégager les critères permettant de mesurer à quelle point deux situations sont, ou ne sont pas, similaires. Une synthèse complète des travaux portant sur les études de cas est proposée par (López, 2013).

2.4.2 Construction automatique du domaine

Les méthodes relatives au data mining et au machine learning sont évidemment présentes dans de nombreuses (sinon toutes) les approches présentées dans ce chapitre. On se réfère néanmoins ici aux méthodes utilisant le data mining pour construire de manière automatique depuis les solutions de l'apprenant des sous-ensembles du domaine. Cette approche est notamment utilisée dans (Barnes and Stamper, 2008), et considère qu'il peut être suffisant de réduire un domaine mal défini à une partie de cet ensemble que l'on arrive à modéliser. Dans (Barnes and Stamper, 2008), l'apprenant doit construire un raisonnement logique pour construire une démonstration. Le raisonnement logique de l'apprenant est modélisé par un processus de décision Markovien (MDP^{16}) associé à un ensemble d'états S , un ensemble d'actions A , des probabilités de transition p , et une fonction de gain μ . Lorsqu'une action a est effectuée dans un état s , la probabilité de transition de l'état s à l'état s' est modélisée par la probabilité $p(s'|s, a)$ et le gain associé vaut $\mu(s'|s, a)$. À chaque itération du raisonnement de l'apprenant, l'état s correspond aux précédentes itérations de la démonstration, et l'action a à l'étape en cours de son raisonnement. Les démonstrations peuvent alors être modélisées par un réseau dont les états sont liés par des actions. Le conseil généré par le tuteur correspond alors à la solution optimale du MDP. Cette solution optimale est calculée par un algorithme d'apprentissage par renforcement. En appliquant un gain important sur l'état correspondant à la fin de la démonstration (gain de 100), et en appliquant des pénalités aux états incorrects (pénalité de 10) et à chaque état parcouru (pénalité de 1), la valeur V de chaque état s est calculée depuis l'équation (2.5).

16. Markovian Decision Process

$$V(s) = \mu(s) + \max_{a \in A} \sum_{s' \in S} p(s, s') V(s') \quad (2.5)$$

La solution optimale depuis un état s peut alors être trouvée depuis un algorithme du plus court chemin.

L'avantage de cette approche est qu'elle ne requière pas de connaissance spécifique ni d'expertise du domaine. Le système peut également apprendre depuis les nouvelles solutions fournies par les apprenants. L'inconvénient est l'impossibilité de fournir un quelconque retour à l'apprenant lorsqu'il se situe sur une partie du réseau inexplorée jusque là (ce qui représente 20% des situations dans le cadre des travaux de (Barnes and Stamper, 2008)).

2.4.3 Aide méta-cognitive

Une approche utilisée notamment dans (Walker et al., 2008) consiste à diviser l'activité en un nombre fini de dimensions, et à fournir un début d'évaluation ainsi que des conseils généraux par rapport à chacune de ces dimensions. Le but n'est pas de fournir une information précise concernant la connaissance à acquérir, mais de définir les principaux défauts de l'utilisateur et l'inciter à les travailler en proposant une piste d'amélioration. L'étude de (Walker et al., 2008) porte sur l'apprentissage de connaissances liées à l'interculturalité. Les apprenants visionnent le film *Monsieur Ibrahim* qui porte sur l'immigration en France. Suite au film, ils doivent disserter sur des questions relatives à l'immigration, aux stéréotypes et discriminations dont peuvent être victimes les immigrés. Les dimensions permettant d'évaluer la performance sont les suivantes :

- la dissertation porte-t-elle sur le sujet énoncé ?
- la dissertation contient-elle un bon argumentaire ?
- la dissertation présente-t-elle plusieurs points de vue ?
- la dissertation ajoute-t-elle des éléments extérieurs (non-présents dans la vidéo) ?
- la dissertation contient-elle des références à des éléments historiques ou culturels ?

L'évaluation de la performance des apprenants sur chacune de ces dimensions est effectuée grâce à une analyse sémantique des dissertations. Plusieurs ensembles de retours sont affectés à chacune des évaluations de chaque dimension. En fonction de l'évaluation sur chacune des dimensions, le tuteur propose alors une évaluation qualitative et des pistes telles que : "Rappelez-vous, il est toujours important d'appuyer ses idées sur des faits issus du film ou de votre connaissance de la culture française", qui sont tirées aléatoirement parmi l'ensemble de retours jugés pertinents par rapport à l'évaluation. Cette pertinence est définie par des règles expertes (e.g. proposer le retour cité précédemment si la dissertation ne s'appuie sur aucune référence culturelle). L'avantage pédagogique de cette approche est de laisser à l'apprenant une partie de l'évaluation de sa performance, en ne lui proposant que des conseils relatifs à la méthode.

L'apprenant, en quelque sorte, apprend à apprendre. Du point de vue de la modélisation, cette approche à l'avantage de ne nécessiter qu'un modèle assez succinct, où il n'est pas nécessaire de représenter la performance de l'apprenant de manière très précise. En contrepartie, cette approche montre des limites lorsque l'apprenant a acquis un certain niveau, et que les connaissances qui lui restent à acquérir deviennent plus spécifiques.

2.4.4 Approches constructivistes

Dans le cadre d'environnements d'apprentissage, les approches constructivistes consistent à amener l'apprenant à prendre conscience de ses stratégies d'apprentissage, afin qu'il puisse construire son savoir. Les activités proposées dans le cadre de l'apprentissage réflexif cherchent notamment à inciter les apprenants à s'investir dans la tâche, en favorisant la réflexion et le traitement des notions afin de favoriser leur compréhension et leur application future. Au sein de cette approche, il n'existe alors pas de méthode "correcte" d'investigation, seulement des constructions individuelles qui s'avèrent satisfaisantes si elles aboutissent à l'appropriation du savoir et du savoir-faire. Plusieurs résultats ont montré les avantages d'une telle approche sur l'acquisition de connaissances, en particulier grâce à l'interaction favorisée dans ce cadre d'apprentissage (White and Frederiksen, 1995). L'implication des apprenants dans la tâche s'en trouve également encouragée (Shute and Glaser, 1990).

Si ce paradigme d'apprentissage peut être aussi bien utilisé dans le cadre de domaines bien définis que mal définis, c'est dans ce dernier cadre que la majorité des environnements de formation ont été développés. Les environnements Belvedere (Suthers et al., 2001) et Rashi (Dragon et al., 2006) mettent notamment en oeuvre des théories constructivistes dans des environnements d'apprentissage. Dans (Suthers et al., 2001), les apprenants sont amenés à réfléchir à des théories scientifiques telles que les raisons de la disparition des dinosaures. Dans (Dragon et al., 2006), les apprenants peuvent tester l'impact à long terme de catastrophes (incendies, pluies acides) sur l'état d'une forêt.

Cet attrait pour les théories constructivistes peut s'expliquer de deux manières : premièrement par les résultats positifs énoncés précédemment concernant l'amélioration de l'implication des apprenants et de l'acquisition du savoir, mais aussi par le fait qu'il ne soit pas primordial de modéliser exhaustivement le domaine dans ce type d'approche. Cela ne signifie pas qu'il n'existe pas de défi lors de la création de ces environnements. Les enjeux sont cependant différents. Il s'agit premièrement de pouvoir suivre le développement de l'apprenant alors même que, si l'on suit les théories cognitivistes, ce développement est propre à chaque apprenant. Il s'agit ensuite de pouvoir intervenir de façon appropriée pour stimuler l'apprenant, ou lui suggérer de s'intéresser à une composante particulière de son activité. L'environnement Rashi a par exemple la particularité de pouvoir offrir une diversité de retours à l'apprenant (images interactives, vidéos, cartes dynamiques), et l'une des principales problématiques soulevées par (Dragon et al., 2006) concerne ce dernier point : alors qu'il est nécessaire, comme nous l'avons nous-mêmes décrit dans le précédent chapitre, de fournir

une diversité de types de retours, comment choisir le type de retour à fournir à chaque instant de l'apprentissage ? Toujours au sein de l'environnement Rashi, les hypothèses formulées sont simplement comparées à une base de connaissances expertes pour en déduire une évaluation. Depuis cette évaluation, le choix de tel ou tel type de retour s'effectue en suivant une base de règles. Cette base de règles lie la situation (contexte plus évaluation) à un type de retour à fournir.

2.4.5 Approches multi-paradigmes

L'approche multi-paradigmes émet l'hypothèse que l'ensemble des approches décrites dans la section 2.3 présente des avantages et des inconvénients dans le cadre de domaines mal définis. À partir de cette hypothèse, l'approche consiste à combiner plusieurs paradigmes sur des parties précises de l'environnement au sein desquelles ils peuvent être implémentés, afin de combiner leurs avantages et de minimiser leurs inconvénients. L'approche multi-paradigmes est notamment utilisée dans le cadre de l'environnement CanadarmTutor (Fournier-Viger et al., 2013) qui vise à entraîner les astronautes à la manipulation d'un robot de la station spatiale internationale. Plusieurs types de retour sont proposés par CanadarmTutor :

- La description a posteriori de la performance de l'apprenant ;
- L'apport de conseils et d'aides proactives permettant de guider l'apprenant ;
- La génération de démonstrations montrant à l'apprenant comment manipuler le robot ;
- Le détail des capacités acquises par l'apprenant ;
- La génération d'exercices personnalisés en fonction des capacités de l'apprenant.

Suivant les critères de (Simon, 1973), la manipulation du robot est une tâche mal définie pour plusieurs raisons :

1. Il n'existe pas de stratégie claire que l'apprenant doit suivre. (Fournier-Viger et al., 2013) expliquent notamment que, si des stratégies pourraient pragmatiquement être définies (considérant par exemple que la minimisation des mouvements de l'apprenant pourrait constituer une stratégie optimale), il est en réalité impossible de définir une telle stratégie du fait des trop nombreuses variables en jeu (contexte du mouvement, familiarité de l'apprenant avec une gestuelle particulière).
2. Le domaine inclut des concepts abstraits partiellement indéterminés ou n'ayant pas de définition absolue. Des composantes liées à la manipulation du robot, telles que la sécurité, l'effort fourni, ou la difficulté du geste répondent à ce critère.
3. La tâche ne peut être divisée en sous-tâches puisque le geste de manipulation doit être considéré dans son intégralité.
4. Il existe un grand nombre (sinon une infinité) de solutions acceptables.

Afin de fournir les différents types de retours énoncés précédemment, (Fournier-Viger et al., 2013) convoquent les paradigmes suivants :

Système expert (section 2.3.2.1) Le paradigme du système expert est utilisé pour générer le chemin correspondant au geste à suivre pour manipuler le robot. L’expertise consiste à donner des degrés de désirabilité à plusieurs zones de l’environnement. Un algorithme de planification de chemin utilise alors ces degrés pour générer une solution considérée comme correcte. Cette solution peut alors être utilisée comme référence et être comparée à la solution de l’apprenant. L’un des inconvénients observés lors de l’utilisation de cette méthode porte sur les solutions générées qui ne se trouvent pas toujours être réalistes. Un autre inconvénient provient du fait que cette méthode ne s’appuie pas sur des données réelles.

Tuteur cognitif (section 2.3.2.2) Le modèle cognitif se base sur une description de la tâche de manipulation sous la forme d’un ensemble de règles. Ces règles représentent le savoir-faire (connaissances procédurales) que doit acquérir l’apprenant. En inférant sur ces règles, le modèle cognitif permet à CanadarmTutor d’évaluer partiellement la performance accomplie, et de proposer des retours immédiats par rapport à cette performance. Le domaine étant mal défini, l’inconvénient de cette approche est qu’elle ne modélise que le sous-ensemble du domaine qu’il a été possible de décrire sous forme de règles.

Data mining (section 2.4.2) L’approche numérique utilisée ici permet de contrer les inconvénients des deux premiers modèles. Les sous-tâches qu’il s’avère impossible de modéliser dans le cadre de l’approche cognitive sont ici analysées à l’aide de données réelles, issues d’exercices antérieurs. L’implémentation se fait, dans le cadre de CanadarmTutor, en trois étapes : une phase de récupération et d’évaluation de données depuis un exercice type dans lequel de nombreuses actions distinctes sont demandées à l’apprenant ; une phase d’extraction sur les données récupérées utilisant un algorithme d’extraction de motifs séquentiels (Fournier-Viger et al., 2010a) ; une phase de reconnaissance qui renvoie pour chaque séquence d’actions d’un apprenant le motif qui lui correspond le plus, ainsi que son évaluation.

Les premiers résultats expérimentaux montrent la capacité de l’approche multi-paradigmes à couvrir un sous-ensemble plus important du domaine, et donc de pouvoir réagir à un plus grand nombre de situations d’apprentissage. L’un des défis notés par (Fournier-Viger et al., 2013) reste cependant la difficulté à sélectionner l’action pédagogique la plus pertinente dans chaque situation. Cette problématique résonne avec celles énoncées dans les approches constructivistes présentées dans la section 2.4.4.

2.4.6 Autres modèles d’apprentissage

D’autres modèles d’apprentissage sont utilisés dans le cadre de domaines mal définis. Plusieurs travaux portent sur les avantages liés à la collaboration

entre humains (Walker et al., 2008; Suthers et al., 2001), ou avec un agent virtuel (Veletsianos and Russell, 2014). Les environnements collaboratifs offrent des espaces d'information et de réflexion partagés qui favorisent les interactions entre apprenants, et par la même occasion l'apprentissage. Les interactions promues par les environnements collaboratifs peuvent en partie réduire le besoin de modélisation du domaine (Walker et al., 2008), d'où l'intérêt de ce paradigme dans le cadre de l'apprentissage dans des domaines mal définis. La planification, et particulièrement les scénarios narratifs, constituent également une piste envisagée (Hodhod and Kudenko, 2008; Carpentier et al., 2013).

2.4.7 Conclusions

Ces dernières années, plusieurs environnements ont été proposés dans le cadre de domaines d'apprentissage mal définis. Ces environnements s'inspirent et viennent s'ajouter aux nombreux environnements d'apprentissage présentés dans la section 2.3. La majorité de ces environnements se concentrent sur un sous-ensemble bien défini du domaine. D'autres environnements se concentrent sur des types différents d'aide ne nécessitant pas une modélisation aussi précise de l'apprenant (aides méta-cognitives). Nous constatons que les paradigmes de représentation de l'apprenant, ainsi que le format des heuristiques de sélection d'actions pédagogiques, se trouvent être très différents suivant l'approche adoptée. Ces fortes différences rendent complexes la compréhension de ces environnements et des règles de sélection qui les régissent. Les règles de sélection d'actions pédagogiques sont en outre fortement couplées à la modélisation de l'apprenant, ce qui complexifie la maintenance (i.e. ajout/mise à jour/suppression d'actions pédagogiques) et l'évaluation (i.e. pertinence des règles de sélection) de ces environnements.

Suite à ces constats, nous proposons dans la section suivante une reformulation générale du problème de sélection d'actions pédagogiques dans le cadre d'environnements interactifs d'apprentissage, et nous proposons un méta-modèle représentant l'apprenant, l'environnement d'apprentissage, et les règles de sélection d'actions pédagogiques comme des entités distinctes réunies dans un même espace de représentation.

2.5 Modélisation du couple apprenant/environnement d'apprentissage

Nous proposons ici une reformulation du problème relatif à la sélection d'actions pédagogiques au sein d'un environnement d'apprentissage. Nous y distinguons clairement les trois entités que sont l'apprenant, les actions pédagogiques et les règles de sélection d'actions. Cette distinction permet de modéliser l'apprenant d'une part, et l'ensemble d'actions pédagogiques d'autres part, de manière indépendante, et donc de n'émettre théoriquement aucune hypothèse à propos de la relation entre ces deux entités. Nous enrichissons ce modèle au cours des prochains chapitres.

2.5.1 Représentation de l'apprenant

Soit Ω_S un espace de représentation “complet” des états de l'apprenant, c'est à dire qui prend exhaustivement en compte toutes les composantes (ou dimensions) S_1, S_2, \dots, S_n sous-jacentes à ces états. Toutes les composantes de ces états ne sont pas toujours modélisables, mais grâce aux différentes méthodes présentées dans ce chapitre, il est possible de modéliser une partie S de Ω_S telle que S soit l'espace produit $S_1 \times S_2 \times \dots \times S_m, m \leq n$ d'une partie des dimensions de Ω_S . L'état d'un apprenant est un élément $s \in S$ représentant sa situation sur chacune des dimensions de l'espace de représentation S , soit $s = \{s_1, s_2, \dots, s_m\}, m \leq n$. Cette représentation reprend notamment les conclusions de (Shute, 2011) portant sur la décomposition de l'apprentissage en un nombre fini de composantes.

Définition 1. *Un apprenant u est représenté dans l'espace S à chaque itération t de son apprentissage par le vecteur associé $\mathbf{s}^{(u)}(t) = \{s_1^{(u)}(t), s_2^{(u)}(t), \dots, s_m^{(u)}(t)\}$.*

2.5.2 Représentation de l'environnement d'apprentissage

Un environnement d'apprentissage peut être représenté par l'ensemble des actions pédagogiques qui pourront être proposées à l'apprenant. On note $A \in \{1, \dots, K\}$.

Définition 2. *On caractérise un environnement d'apprentissage par l'ensemble $A \in \{1, \dots, K\}$ des actions qui peuvent être fournies. À chaque itération de décision t , l'environnement sélectionne une action $I_t \in A$.*

2.5.3 Relation entre apprenant et environnement

La relation entre les représentations de l'apprenant et de l'environnement permet de définir les heuristiques de sélection d'une action pédagogique en fonction de la situation de l'apprenant. Notre approche s'appuie sur les postulats suivants :

Définition 3. *L'impact, noté $\mu_i^{(u)}(t)$, associé à une action $i \in A$, mesure la variation de l'état de l'apprenant suite à la sélection de cette action. Pour tout $i \in A$, l'impact $\mu_i^{(u)}(t)$ est positif si l'apprenant progresse, négatif s'il régresse, et nul si son état reste inchangé ou si l'action i n'a pas été sélectionnée ($I_t \neq i$).*

Hypothèse 1. *Pour tout apprenant u d'état $\mathbf{s}^{(u)}(t) \in S$ à l'instant t , il existe une action optimale $i_t^* \in A$ telle que*

$$\mu_{i_t^*}^{(u)}(t) \geq \mu_i^{(u)}(t), \forall i \in A$$

L'action optimale est alors l'action qui maximisera l'influence sur l'apprentissage de l'apprenant. Tout l'enjeu de la sélection d'actions pédagogiques est donc de sélectionner, à chaque instant t de la séquence de décision, l'action

optimale parmi l'ensemble des actions présentes dans l'ensemble A . L'environnement suit alors une politique optimale de sélection π^* que l'on définit comme suit :

Définition 4. Une politique $\pi : S \rightarrow A$ est une application qui, pour un environnement d'apprentissage défini par l'ensemble A des actions pédagogiques qu'il peut proposer, associe à chaque situation $s^{(u)}(t) \in S$ une action $i \in A$ à sélectionner. Parmi la famille d'application $\pi : S \rightarrow A$, la politique optimale π^* associe à chaque situation $\mathbf{s}^{(u)}(t)$ l'action optimale $i_t^* \in A$ maximisant le gain $\mu_i^{(u)}(t)$ pour tout $i \in A$ à cet instant.

2.5.4 Adaptation basée sur une expertise

Nous nous intéressons dans un premier temps à l'adaptation basée sur une expertise, à l'instar des principaux travaux présentés dans ce chapitre. Il s'agit donc d'intégrer des règles établies par des experts au sein de notre modèle, et nous utilisons pour cela la notion de similarité. Nous avons besoin pour cela de préciser la nature de l'espace S , restée volontairement floue jusqu'ici.

Soient $D_{S_1}, D_{S_2}, \dots, D_{S_m}$ les distances relatives aux dimensions de S . L'espace (S, D_S) est alors l'espace métrique relatif à l'ensemble $S_1 \times S_2 \times \dots \times S_m$ muni de la distance D_S définie par $D_S(\mathbf{s}^{(u1)}, \mathbf{s}^{(u2)}) = M(D_{S_1}, D_{S_2}, \dots, D_{S_m})$, où M est une norme L^p arbitraire sur \mathbb{R}^m . Nous pouvons désormais préciser la définition 3 :

Définition 5. L'impact, noté $\mu_i^{(u)}(t)$, d'une action $i \in A$, mesure la variation de l'état de l'apprenant suite à la sélection de i :

$$\mu_i^{(u)}(t) = \delta_{[I_t=i]} D_S(\mathbf{s}^{(u)}(t), \mathbf{s}^{(u)}(t + \Delta)),$$

avec $\delta_{[I_t=i]}$ valant 1 si l'action i est sélectionnée à l'itération t , 0 dans le cas contraire.

Lors des différents échanges permettant de collecter l'expertise, les experts formulent généralement les règles de décision en indiquant dans quelle(s) situation(s) chaque action devrait être sélectionnée. Si l'on note $T^{(i)} = \{\mathbf{s}^{(i)} | \mathbf{s}^{(i)} \in S, i \in A\}$ l'ensemble des situations qui, selon l'expert, doivent amener l'environnement à sélectionner l'action i , alors la politique π_{expert} suit l'heuristique présentée dans la définition 6.

Définition 6. Dans un environnement d'apprentissage régi par une expertise, et caractérisé par un ensemble $A \in \{1, \dots, K\}$ des actions possibles, l'aide la plus pertinente dans la situation $\mathbf{s}^{(u)}(t) \in S$ est l'action i qui vérifie

$$i = \operatorname{argmin}(\{ \min_{i' \in A} D_S(\mathbf{s}^{(u)}(t), \mathbf{s}^{(i')}) , i' \in A \})$$

Cette propriété pourrait être étendue à une recherche des k actions dont les situations optimales sont les plus similaires à la situation courante si l'on souhaite pouvoir conjuguer les actions pédagogiques. À noter que l'expertise ne

prend pas en compte le caractère individuel de l'apprenant u , et est constante dans le temps. L'algorithme 1 présente le processus de sélection d'actions depuis une expertise selon notre approche.

Algorithme 1 : Formalisation du problème de sélection d'actions depuis une expertise

```

 $t = 1;$ 
 $I_0 =$  tirage aléatoire uniforme parmi les éléments de  $A$ ;
tant que la session n'est pas terminée faire
  proposer l'exercice avec l'action  $I_{t-1}$ ;
  tant que l'exercice n'est pas terminé faire
    enregistrer la performance de l'apprenant;
  fin
  récupérer la situation  $\mathbf{s}^{(u)}(t)$ ;
  sélectionner l'action
   $I_t = \operatorname{argmin}_{i \in A} (\{\min_{\mathbf{s}^{(i)} \in T^{(i)}} D_S(\mathbf{s}^{(u)}(t), \mathbf{s}^{(i)}), i \in A\})$ ;
   $t = t + 1$ ;
fin

```

2.5.5 Caractérisation du geste

La maîtrise de la calligraphie tient principalement à l'acquisition d'une maîtrise de l'outil dans des dimensions de régularité. Ces dimensions de régularité peuvent être relatives à la performance gestuelle (morphocinétique), mais également à la séquence des tracés et à la spatialisation des tracés dans l'espace graphique (topocinétique) qui définit le style de calligraphie (Paillard, 1991). Lors du processus d'apprentissage de la calligraphie, le formateur reconnaît des motifs d'erreurs. Ces erreurs peuvent être liées à la gestuelle de l'apprenant (e.g. utilisation démesurée des doigts), ou à la cognition (e.g. détérioration de l'attention liée à l'organisation spatiale). Depuis cette analyse des erreurs, le formateur fournit à l'apprenant des conseils verbaux, voire effectue des démonstrations de l'activité. En suivant un tel entraînement, les apprenants construisent une connaissance liée à leurs expériences successives et à la mémoire kinesthésique des gestes effectués. Cette connaissance leur permet d'acquérir le contrôle et la régularité essentiels à la pratique de la calligraphie (Doan et al., 2016).

Dans le contexte de la résolution de problèmes, les modèles overlay et les modèles par perturbation représentent des ensembles distincts : le modèle overlay représente le niveau de compétence de l'apprenant par rapport à l'ensemble des compétences formant l'expertise ; le modèle par perturbation représente des erreurs types. La distinction est plus floue dans le cadre du geste si l'on définit le geste expert comme un geste ne présentant pas d'erreur. Le modèle overlay dans le cadre du geste correspond donc à un modèle par perturbation, où l'expert est représenté par une performance vierge de toute erreur.

Notre approche de la représentation du geste revient donc à définir les dimensions de l’espace S comme des types d’erreurs communément commises. Ce sont ces erreurs que l’environnement vise à réduire. Il s’agit alors de définir depuis le geste de l’apprenant des valeurs s_j , $j = 1, \dots, m$ correspondant à l’interprétation de l’erreur commise. Reprenant les principes du modèle overlay proposé par (Brusilovsky and Millán, 2007), on établit que toute valeur de s_j est comprise entre 0 (pas d’erreur) et 1 (erreur grossière). L’importance de l’erreur commise peut alors être mesurée par la probabilité que la performance de l’apprenant présente l’erreur grossière.

2.6 Conclusions

Nous avons présenté dans ce chapitre la littérature relative à la modélisation de l’apprenant dans le cadre d’environnements intelligents d’apprentissage. Si ces environnements s’inscrivent pour la majorité dans des domaines d’apprentissage bien définis (dont font partie les domaines liés à la résolution de problèmes), de récents travaux s’intéressent à des domaines d’apprentissage mal définis, c’est à dire des domaines dont la caractérisation et l’évaluation est plus complexe et où la transposition des méthodes usuelles est impossible.

Suite à la présentation de ces travaux, nous avons proposé une reformulation du problème d’adaptation de comportements. L’approche que nous proposons représente l’apprenant et l’environnement dans un même espace métrique de représentation, et permet la formulation des règles établies par les experts. Au sein de cet espace, nous modélisons le geste en fonction des erreurs commises, et en considérant que l’expertise est représentée par un geste sans erreur (mélange des modèles overlay et par perturbation communément utilisés dans des domaines bien définis). Les règles expertes sont exprimées dans cet espace de représentation par un ensemble de situations pour lesquelles chaque feedback est considéré par l’expert comme pertinent.

Dans le chapitre 3, nous présenterons des méthodes permettant d’interpréter les valeurs d’erreur depuis la pratique gestuelle de l’apprenant. Nous préciserons notamment les méthodes relatives à la classification multilabel qui visent à étiqueter une observation (i.e. un geste) en lui assignant une ou plusieurs classes (i.e. des erreurs).

Nous décrivons également dans ce chapitre une première implémentation de notre approche dans un environnement d’apprentissage du geste de calligraphie. Nous présenterons les dimensions d’erreurs permettant de représenter l’activité de calligraphie, ainsi que l’ensemble des feedbacks implémentés depuis les recommandations d’experts. Finalement, nous présenterons une expérimentation visant à mesurer l’impact d’un tel environnement sur l’apprentissage de la calligraphie.

Chapitre 3

Adaptation basée sur l’expertise : Expérimentation

3.1 Introduction

Suite à la proposition théorique énoncée dans le chapitre 2, nous décrivons ici les caractéristiques de la plateforme CalliSmart, développée pour l’apprentissage de la calligraphie dans le cadre de nos travaux. Nous présentons une première expérimentation visant à évaluer un environnement d’apprentissage sélectionnant de manière automatique des feedbacks parmi un ensemble de feedbacks déterminés par des experts. Nous mesurons notamment l’apport d’un tel environnement face au syndrome de “dépendance à l’environnement d’apprentissage” décrit au chapitre 1.

Au sein du projet Descript, l’objectif est de pouvoir utiliser la plateforme CalliSmart dans le cadre de cours de calligraphie dispensés à l’École Supérieure d’Art et de Design (ESAD) d’Amiens, partenaire du projet. La plateforme CalliSmart est développée sur une table Wacom¹, et doit proposer à l’apprenant des aides adaptées à son niveau. Elle repose pour cela sur les travaux présentés dans le chapitre 2.

La création d’une telle plateforme nécessite une première étape de caractérisation de l’activité. Cette étape, située en amont du processus d’apprentissage, permet de définir les dimensions caractérisant l’activité, et d’implémenter l’ensemble d’actions pédagogiques dont les éléments pourront être sélectionnés de manière automatique par l’environnement. Nous nous positionnons principalement ici par rapport aux travaux relatifs aux environnements d’apprentissage du geste d’écriture et de calligraphie, ainsi qu’à l’expertise de formateurs en calligraphie de l’ESAD.

Une fois l’activité caractérisée, et l’ensemble des actions possibles défini, une seconde étape consiste à permettre à l’environnement de sélectionner automatiquement les aides qu’il fournit lors du processus d’apprentissage. Lors

1. <https://eu-store.wacom.com/Catalog/Pen-Displays/Cintiq/Cintiq-27QHD-and-ArtPen-S02>

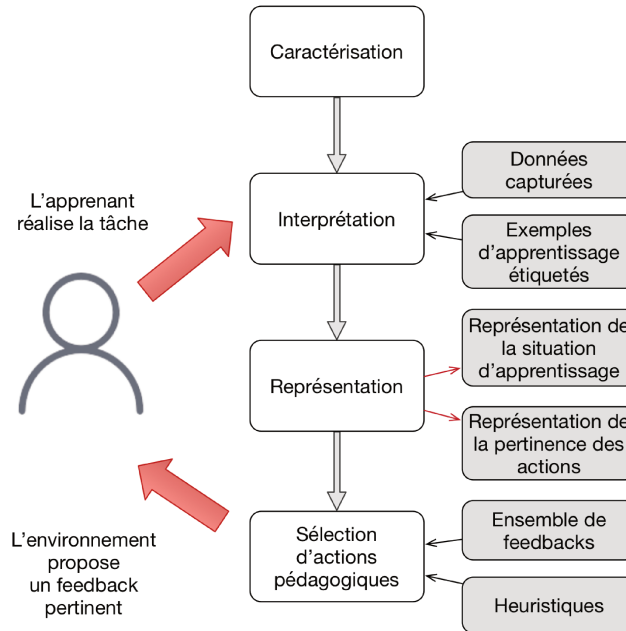


FIGURE 3.1 – Ensemble des opérations nécessaires à l'adaptation automatique de l'environnement

de ce processus, l'environnement doit effectuer une suite d'opérations qui, depuis l'activité d'écriture de l'apprenant jusqu'au choix de l'action pédagogique à proposer, lui permet d'identifier la situation d'apprentissage et la pertinence de chaque action pédagogique (figure 3.1).

Ce chapitre suit l'organisation suivante : la caractérisation de l'activité est décrite dans la section 3.2. Cette section présente également l'ensemble des actions pédagogiques implémentées. La reconnaissance de l'activité, ainsi que l'adaptation de l'environnement par rapport à la représentation qui en découle, sont présentées dans la section 3.3. Dans cette section sont discutés les principaux travaux relatifs à la reconnaissance de gestes, ainsi que l'implémentation et la validation du module de reconnaissance de la plateforme CalliSmart. Enfin, la section 3.4 présente les résultats d'une première expérimentation mettant en avant les avantages liés à l'apport d'un environnement sélectionnant automatiquement son comportement depuis une diversité d'actions pédagogiques en suivant un ensemble de règles définies par des experts.

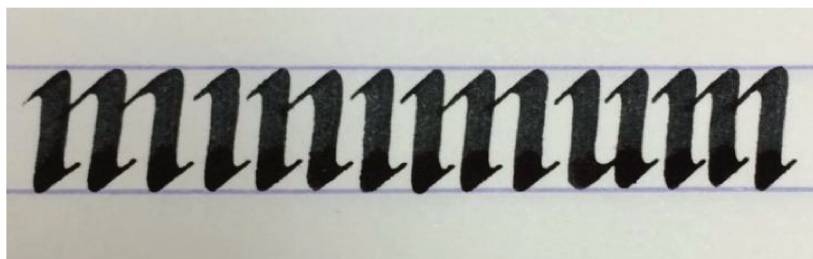


FIGURE 3.2 – L’exercice de l’écriture du mot “minimum” allie contraintes morphocinétiques et topocinétiques

3.2 Environnement CalliSmart (1)

3.2.1 Caractérisation du geste

Nous nous plaçons dans le cadre de l’exercice de l’écriture du mot “minimum” (figure 3.2). Cet exercice, souvent utilisé lors de formation à la calligraphie, présente l’avantage de proposer un motif simple, qui se répète. L’objectif de l’écriture du mot minimum est alors dual, puisqu’il s’agit de répondre aux contraintes morphocinétiques (taille et verticalité du tracé, fluidité du geste) et topocinétique (régularité de la structure du tracé, des espaces entre les différents éléments) liées à l’activité.

En accord avec les experts de l’ESAD, nous avons dégagé trois motifs d’erreur permettant de mesurer la capacité de l’apprenant à répondre à cet objectif :

- une **erreur d’inclinaison** se traduit par la non-verticalité du trait ;
- une **erreur de régularité** se traduit par des espacements inégaux entre les différents éléments verticaux ;
- une **erreur de taille** se traduit par une différence entre la taille du trait et l’espace vertical entre les deux lignes de portée.

Nous caractérisons donc l’activité de l’apprenant en fonction de motifs d’erreur identifiés. Cette modélisation peut être assimilée à un modèle *overlay* de l’apprentissage (voir section 2.3.1.2) dans le sens où l’activité de l’apprenant est évaluée en fonction de l’activité jugée experte. En effet, on peut considérer que l’activité experte se caractérise par des scores d’erreur très faibles sur chacune des dimensions, et la maîtrise de l’apprenant peut donc être mesurée comme la distance à cette expertise.

Puisque l’erreur, dans le cadre du geste, ne peut être évaluée de manière aussi systématique qu’une erreur dans le cadre de la résolution de problème (e.g. résultat d’une somme, d’un produit), il est nécessaire d’utiliser un algorithme de reconnaissance automatique capable d’évaluer l’erreur. Le modèle *overlay* est alors similaire à celui utilisé dans des travaux tels que (Brusilovsky and Millán, 2007), où le score sur chacune des dimensions de l’activité est défini par la probabilité que l’erreur ait été effectuée par l’apprenant. Nous reviendrons en détails dans la section 3.3 sur les travaux relatifs à la reconnaissance

automatique d'erreurs, et leur application dans CalliSmart.

3.2.2 Ensemble d'actions pédagogiques

L'ensemble des actions pédagogiques est déterminé en fonction des dimensions étudiées, c'est à dire ici les composantes liées à l'inclinaison du tracé, à sa régularité, ainsi qu'à sa taille par rapport aux portées définissant l'espace graphique. Sans ignorer que plusieurs travaux ont montré l'impact positif de l'ajout de feedbacks extrinsèques haptiques (Teo et al., 2002; Bara et al., 2004) et sonores (Danna and Velay, 2015) lors de l'apprentissage de l'écriture, nous avons privilégié des retours visuels dans la première phase de conception du projet Callismart². Et si l'utilisation de retours visuels dans ce type d'apprentissage a été critiqué considérant que la modalité visuelle, largement utilisée en tant que feedback intrinsèque, nuit à la lisibilité de tels feedback (Weil and Amundson, 1994), ces critiques sont cependant à mesurer. Effectivement, si l'ajout d'informations visuelles délocalisées par rapport au mouvement est une idée à proscrire, les feedbacks extrinsèques permis par le développement des tablettes tactiles, avec notamment la possibilité de jouer sur les couleurs de la trace, peuvent permettre d'enrichir la boucle sensorimotrice sans détériorer l'attention visuelle de l'apprenant (Danna and Velay, 2015). En outre, les feedbacks visuels peuvent également s'avérer utiles lorsqu'ils sont proposés a posteriori, suite à la performance de l'apprenant (Danna and Velay, 2015). Les feedbacks présentés ici ont été développés en accord avec les experts, et en tenant compte de ces conclusions.

3.2.2.1 Trace et lignes de portées

Sont décrits ici les informations émanant de l'implémentation même de l'environnement numérique de calligraphie : les lignes de portée, et la trace induite par le contact entre l'outil et le plan. Nous prenons soin ici de décrire ces éléments, puisqu'ils entrent dans l'ensemble des informations perceptibles dans la boucle de rétroaction de l'apprenant. Cependant, nous ne considérons pas ces informations comme faisant partie de l'ensemble d'actions pédagogiques, puisqu'elles sont activées en permanence, et ne sont donc pas régies par un quelconque processus de sélection.

L'espace graphique est délimité par deux lignes horizontales centrées verticalement par rapport à l'écran, et séparées par un espace vertical de 14 centimètres. Ces lignes horizontales sont la transposition des lignes de portée utilisées dans le cadre de l'apprentissage de la calligraphie.

La trace numérique est dessinée lors du contact entre l'outil et le plan graphique. La largeur de cette trace dépend de la pression normale au plan appliquée par l'outil. Cette astuce permet de transcrire le feedback naturellement induit par l'écrasement d'un pinceau ou la bavure d'une plume lorsque la pression appliquée est trop forte. La relation entre la largeur de la trace et la

². Une première étape de la sonification du tracé par l'agence Digs a été réalisée en juin 2016 et sera probablement introduite dans l'environnement Callismart.

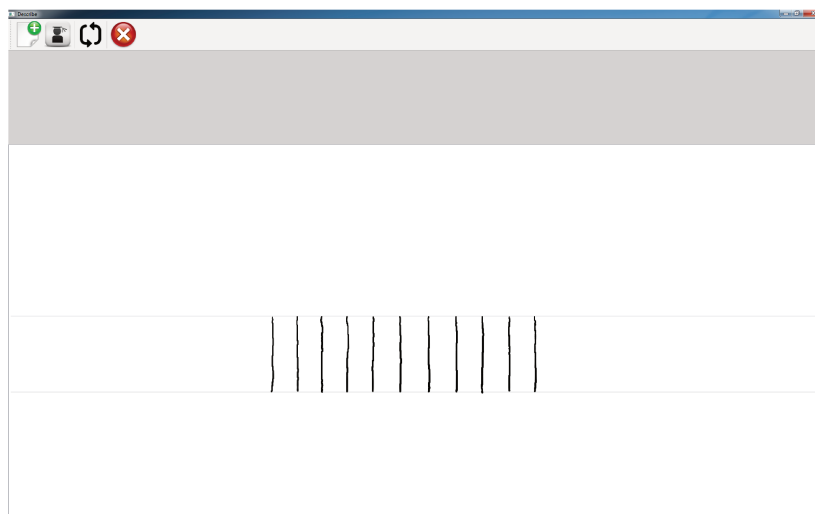


FIGURE 3.3 – Environnement graphique sans feedback supplémentaire

pression normale appliquée est linéaire. La figure 3.3 illustre le rendu visuel lors de la réalisation d'un exercice dans l'environnement graphique (sans feedback supplémentaire).

3.2.2.2 Connaissance du résultat

Le feedback lié à la *connaissance du résultat* (CR) est proposé a posteriori de la performance de l'apprenant. Ce type de feedback permet de donner une évaluation de la performance par rapport aux objectifs initiaux. La connaissance du résultat permet aux apprenants de situer leur performance, ce qui impacte directement leur engagement vis-à-vis de la tâche et améliore leur confiance (Salmoni et al., 1984). Ce type de feedback peut également constituer une première forme de guidage, dans le sens où, suite à la prise en compte du feedback, l'apprenant modifie dans les exercices suivant son activité gestuelle (Salmoni et al., 1984). Le feedback CR fourni par la plateforme CalliSmart propose un score sur 100 sur chacune des trois dimensions caractérisant l'activité (figure 3.4). Ce score est fourni par l'algorithme de reconnaissance d'erreurs décrit dans la section 3.3. Ce feedback est affiché à la fin de chaque exercice.

3.2.2.3 Feedback supplémentaire

Les feedbacks supplémentaires correspondent à l'ensemble des actions pédagogiques parmi lesquelles l'environnement sélectionne, à chaque tour, l'élément le plus pertinent par rapport à l'état de l'apprenant. Ces feedbacks ont été choisis par des experts et correspondent à des retours qu'ils donnent réguliè-

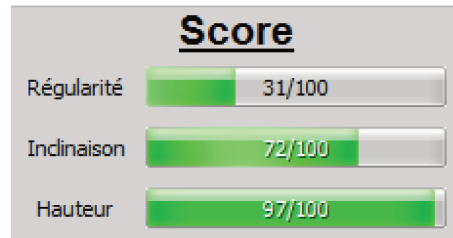


FIGURE 3.4 – Le feedback CR (connaissance du résultat) propose un score sur 100 sur chacune des dimensions d'erreur considérées



FIGURE 3.5 – Feedback illustrant une erreur d'inclinaison d'un élément

rement, dans le cadre de cours de calligraphie, pour travailler les dimensions que nous avons sélectionnées ici. En outre, ces feedbacks font soit écho à la dimension morphocinétique de l'activité, soit à sa dimension topocinétique. Ils informent l'apprenant en temps réel en fonction de son activité, et agissent donc comme des éléments constitutifs de la boucle proprioceptive de l'apprenant. Une fois l'exercice terminé, ces feedbacks peuvent être vus comme des informations relatives à la *connaissance de la performance* (CP), puisqu'ils décrivent de manière qualitative la performance de l'apprenant par rapport à la dimension considérée.

Dimension relative à l'inclinaison des éléments Le feedback illustrant l'inclinaison colore la trace issue du geste en fonction de son inclinaison par rapport à l'axe vertical. Un tracé parfaitement vertical est coloré en vert, tandis que l'inclinaison se traduit par une couleur de tracé virant vers le rouge en fonction de la valeur de l'angle (figure 3.5).

Dimension relative à la régularité des éléments Le feedback visant à travailler le geste sur la dimension de la régularité affiche, lors du tracé d'un nouvel élément, une flèche indiquant la position horizontale qui correspond à une régularité parfaite par rapport aux précédents éléments tracés (figure 3.6).

Dimension relative à la taille des éléments Le feedback relatif à la taille



FIGURE 3.6 – Feedback illustrant la spatialisation des éléments

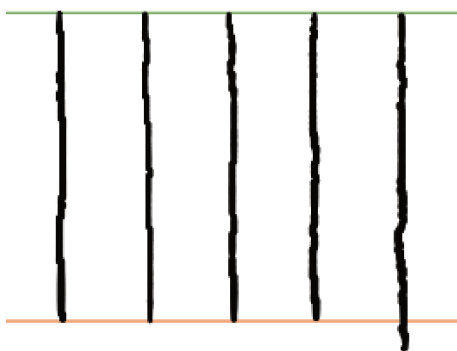


FIGURE 3.7 – Feedback illustrant une erreur sur la taille d'un élément

tient principalement un rôle attentionnel, les apprenants focalisés sur d'autres aspects inhérents à leur gestuelle oubliant parfois l'importance de limiter leur production dans l'espace graphique. Ce feedback colore alors la ligne de portée dans un intervalle allant du vert au rouge en fonction de l'ampleur de l'erreur (figure 3.7).

3.3 Adaptation dans CalliSmart

L'objectif du processus d'adaptation automatique est de permettre à l'environnement de retranscrire la méthode pédagogique utilisée par le formateur en calligraphie : celui-ci identifie les erreurs, et sélectionne l'action qui permettra à l'apprenant de prendre conscience et d'améliorer sa performance par rapport à cette erreur. L'identification de/des erreur(s), puis la sélection du feedback constituent donc les deux étapes distinctes du processus d'adaptation que nous décrivons ici.

3.3.1 Travaux relatifs à la reconnaissance de gestes

Ces dernières années, les améliorations et la réduction du coût des capteurs (caméra infrarouge, caméra de profondeur, gyroscope) ainsi que l’intérêt croissant pour les interfaces interactives intelligentes, ont amené de nombreux chercheurs à proposer des méthodes permettant de reconnaître des motifs de gestes (tâches de *segmentation* et de *classification*, (Mitra and Acharya, 2007)). La segmentation consiste à définir un cadre temporel et spatial au geste : quand il démarre, quand il s’arrête. La classification consiste à étiqueter le geste en lui affectant un ou plusieurs label(s), chaque étiquette donnant une information relative à ce geste (e.g. nature du geste : “marche”, “course”, “saut” ; évaluation du geste : “novice”, “expert” ; détection d’erreur : “mauvaise posture”, “flexion trop importante du genou”). Dans CalliSmart, le geste est segmenté en fonction de la position de l’outil par rapport au plan. Les capteurs intégrés dans une table Wacom permettent de capturer l’outil, non seulement lorsqu’il est en contact avec l’écran, mais également lorsqu’il se situe dans une zone suffisamment proche de l’écran (de l’ordre de quelques centimètres). Nous considérons donc qu’un geste démarre (resp. s’arrête) lorsque l’outil entre (resp. sort) de cette zone. L’environnement analyse alors la trace produite par l’utilisateur. Si aucune trace n’a été produite (l’utilisateur s’est simplement approché puis éloigné sans tracer), le geste n’est pas analysé. Dans le cas contraire, l’environnement cherche à classifier le geste enregistré. Nous nous intéresserons donc principalement ici aux travaux relatifs à la classification de gestes.

Un geste peut être vu comme un jeu d’observations spatio-temporelles, c’est à dire que les données consistent en une suite de positions et d’orientations évoluant au cours du temps. Les méthodes de classification peuvent être divisées en deux groupes suivant qu’elles représentent ces données comme un vecteur, ou comme une séquence temporelle :

- Une représentation séquentielle du geste considère une séquence d’observations $x = (x_1, \dots, x_T)$, où chaque observation x_t pour $t \in \{1, \dots, T\}$ représente la situation au temps t (e.g. positions, orientations, descripteurs). L’objectif est d’assigner à x une étiquette y (classification dite *monolabel*), ou un ensemble d’étiquettes $y = (y_1, \dots, y_D)$ (classification dite *multilabel*). L’avantage d’une représentation séquentielle est de conserver la progression temporelle du geste.
- Une représentation vectorielle du geste réduit le geste à un unique vecteur x qui décrit à lui seul l’observation toute entière. Comme pour la représentation séquentielle, la classification consiste alors à attribuer à x une étiquette y ou un ensemble d’étiquettes (y_1, \dots, y_D) . Les représentations vectorielles s’avèrent intéressantes lorsque des descripteurs globaux discriminants vis-à-vis des étiquettes possibles peuvent être calculés depuis l’ensemble d’observations.

Puisque nous nous positionnons ici sur une reconnaissance de la qualité du geste à l’échelle de l’ensemble de l’exercice (afin de pouvoir mesurer les dimensions de régularité), et puisqu’il existe des caractéristiques pertinentes à l’échelle de l’exercice (inclinaison des tracés, taille du tracé), nous privilégierons

une représentation vectorielle.

Outre cette division des méthodes de classification selon deux modes de représentation du geste, il est également possible de diviser ces méthodes par la façon dont elles prédisent l'étiquette y en fonction du vecteur (dans une représentation vectorielle), ou de la séquence d'observations (dans une représentation séquentielle).

3.3.1.1 Approches basées sur des exemples de référence

Les approches basées sur des exemples de référence comparent l'observation x à l'ensemble des observations de référence dont la classe est connue, c'est à dire dont on connaît l'étiquette y . Ces méthodes assignent ensuite à l'observation x une étiquette y en fonction des observations les plus similaires. Parmi ces méthodes, les plus utilisées sont la méthode des *k plus proches voisins* (Efros et al., 2003) (représentation vectorielle), qui assigne l'étiquette la plus représentée parmi les k observations connues les plus similaires à x (le paramètre k pouvant être obtenu par validation croisée), et l'algorithme de *déformation temporelle dynamique* (ten Holt et al., 2007) (représentation séquentielle), qui mesure la similarité entre deux séries temporelles, et détermine un alignement optimal entre les deux séquences.

3.3.1.2 Approches paramétriques

Les approches paramétriques entraînent un modèle permettant de généraliser la relation entre les observations x et les étiquettes y . Adoptant une représentation séquentielle du geste, les approches basées sur les processus de Markov, et notamment les modèles de Markov à états cachés (HMM) (Rabiner, 1989), font partie des méthodes les plus utilisées (Bevilacqua et al., 2009; Manitsaris et al., 2014). Les HMM suivent un processus de Markov comprenant un nombre fini d'états reliés par des probabilités de transition. Les états et les observations sont découplés, et la distribution des observations est modélisée par une loi de probabilité connue associée à l'état courant (les HMM standards utilisant des mélanges de distributions gaussiennes). Utilisé en classification de gestes, le HMM est une méthode générative³. Une méthode courante est alors d'entraîner un modèle par type de geste. Pour une séquence d'observations x donnée, il s'agit alors de déterminer le modèle ayant la plus grande probabilité de reproduire la séquence d'observations. Les algorithmes de Baum-Welch (pour l'entraînement des modèles) et de Viterbi (pour déterminer une séquence d'états optimale) sont alors utilisés. À noter qu'il existe également des modèles discriminants, utilisés dans le cadre de la reconnaissance de gestes, basés eux aussi sur des processus de Markov tels que les champs aléatoires conditionnels (CRF, (Lafferty et al., 2001)), et les champs aléatoires conditionnels à états cachés (HCRF, (Quattoni et al., 2007)).

3. Autrement dit, il est entraîné à maximiser la vraisemblance jointe des observations et des classes. Nous précisons dans la section 3.3.1.3 les différences entre modèles génératifs et discriminants.

Dans le cadre de représentations vectorielles, les séparateurs à vaste marge (SVM, (Burges, 1998)), les classifieurs logistiques et les classifieurs bayésiens naïfs sont des méthodes communément utilisées (Wu et al., 2009; Biswas and Basu, 2011; Wang et al., 2012). Les SVMs reposent sur les concepts de *noyau* et de *marge maximale*. L'astuce du noyau consiste à modifier l'espace de représentation en un espace de plus grande dimension où les échantillons de classes différentes sont linéairement séparables. La maximisation de la marge permet de trouver un hyperplan optimal permettant de séparer les classes. Il a été démontré par Vapnik que l'hyperplan optimal existe et qu'il est celui qui maximise la marge entre les échantillons et lui-même. Les SVMs ont démontré de très bonnes performances dans le cadre de la reconnaissance de gestes, mais leurs résultats sont difficilement interprétables. En effet, un résultat donné par un SVM n'est pas probabiliste, mais est déduit de la position de l'observation à étiqueter par rapport à l'hyperplan séparateur. Puisque nous adoptons un modèle overlay semblable à celui proposé par (Brusilovsky and Millán, 2007), où la coordonnée sur chacune des dimensions du modèle de l'apprenant est déterminée par la probabilité d'effectuer l'erreur, nous préférons nous intéresser aux classifieurs logistiques et aux classifieurs bayésiens naïfs, qui sont eux des modèles probabilistes, respectivement discriminants et génératifs.

Les régressions logistiques reposent sur l'hypothèse que l'expression $\log(\frac{p}{1-p})$ est une combinaison linéaire des éléments $x_i, i = 1, \dots, N$ du vecteur x . On note Θ le vecteur de paramètres du modèle. Contrairement à l'hypothèse linéaire $p = \Theta_0 + \Theta \cdot x$, l'expression $\log(\frac{p}{1-p}) = \Theta_0 + \Theta \cdot x$ permet d'obtenir des valeurs probabilistes, et permet de modéliser les phénomènes où la relation entre p et x n'est pas homogène suivant la valeur de p (e.g. lorsqu'augmenter la valeur de p de 0.1 requiert une différence plus importante sur x lorsque p est proche de 0 que lorsqu'il avoisine 0.5). Le classifieur logistique binaire consiste alors à modéliser les observations par la fonction logistique $p(Y = 1|X = x) = \frac{1}{1+e^{-(\Theta_0+\Theta \cdot x)}}$, où Θ est un vecteur de paramètres. Le classifieur associe à x la classe associée à $Y = 1$ si $p(Y = 1|X = x) > 0.5$, ou la classe associée à $Y = 0$ dans le cas contraire.

Le modèle bayésien naïf est un modèle conditionnel exprimant, pour le vecteur $x = (x_1, \dots, x_N)$, les probabilités $p(C|x_1, \dots, x_N)$ d'avoir une classe sachant le vecteur d'observations. Partant du théorème de Bayes : $p(C|x_1, \dots, x_N) = \frac{p(x_1, \dots, x_N|C)p(C)}{p(x_1, \dots, x_N)}$, et en émettant l'hypothèse (forte) que les descripteurs x_i sont indépendants conditionnellement à la classe⁴, on obtient $p(C|x) = \frac{p(C) \prod_{i=1}^N p(x_i|C)}{p(x_1, \dots, x_N)} = \frac{p(C) \prod_{i=1}^N p(x_i|C)}{\sum_{C'} p(C') \prod_{i=1}^N p(x_i|C')}$. Les paramètres du classifieur bayésien naïf sont usuellement estimés par la méthode du *maximum a posteriori*. En inférence, on assigne alors à l'observation x l'étiquette $\hat{y} = C_k$ vérifiant $\hat{y} = \operatorname{argmax}_{k \in 1, \dots, K} p(C_k|x)$.

4. Soit $p(x|C) = \prod_{i=1}^N p(x_i|C)$.

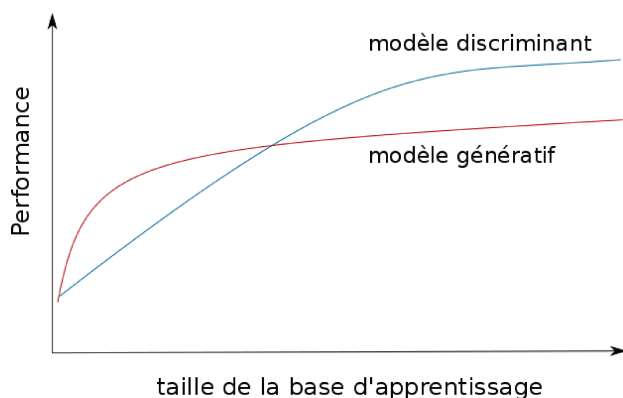


FIGURE 3.8 – Comportements schématiques possibles de modèles génératifs et discriminants (Soullard, 2013)

3.3.1.3 Modèles génératifs vs. modèles discriminants

Certaines approches, telles que les classifieurs bayésiens naïfs ou les HMMs sont dits génératifs, alors que d'autres comme les classifieurs logistiques ou les CRF sont dit discriminants. Un modèle est dit génératif lorsque le critère d'apprentissage maximise la vraisemblance jointe des observations et de leurs classes, et discriminant lorsque le critère d'apprentissage maximise la vraisemblance conditionnelle des classes sachant les observations. Suite à l'étude d'une paire de modèles génératif-discriminant basée sur les modèles bayésiens naïfs et les régressions logistiques, (Ng and Jordan, 2002) établissent qu'un modèle discriminant a une meilleure performance asymptotique. Les auteurs démontrent néanmoins que le modèle génératif peut converger plus rapidement vers sa performance optimale que le modèle discriminant en fonction de la taille de la base d'apprentissage (figure 3.8), ce qui est cohérent avec les conclusions de (Bouchard and Triggs, 2004) estimant les modèles discriminants comme souffrant de davantage de variance que les modèles génératifs. Le modèle génératif apparaît donc plus fiable sur de petites bases de données d'apprentissage (Soullard, 2013). Le modèle discriminant est par contre plus performant lorsque la base de données d'apprentissage s'avère plus conséquente, ce qui est également cohérent avec les commentaires de (Bouchard and Triggs, 2004), les modèles discriminants ayant un biais moins important que les modèles génératifs.

3.3.1.4 Classification multilabel, multiclasse et rapport au geste

Nous avons présenté ici différentes méthodes de classification, indépendamment du type de classification. Le type de classification le plus simple est la classification binaire, c'est à dire lorsque l'étiquette y ne peut prendre que deux

valeurs, soit 0, soit 1⁵. Il arrive pourtant que l’ensemble des classes possibles ait plus de deux éléments, c’est notamment le cas de notre représentation du geste puisque nous souhaitons pouvoir assigner les trois classes d’erreur présentées dans la section 3.2. Existente alors deux types de classification : la classification multiclasse, où l’exercice consiste à prédire une classe parmi l’ensemble des classes pour une observation donnée ; et la classification multilabel qui consiste à prédire une ou des classes parmi cet ensemble. La différence entre classification multiclasse et multilabel peut être illustrée par un exemple simple : si la classification consiste, à partir d’un article, à savoir en quelle langue il est rédigé, la classification est multiclasse (une seule classe possible). Si par contre, par rapport à ce même article, il s’agit de définir si l’article traite de politique, de sport, ou d’information, la classification est multilabel (l’article peut relever de l’information, et évoquer des tensions politiques en rapport avec un évènement sportif).

En rapport au geste, et à une caractérisation du geste par types d’erreur, le choix d’un cadre multiclasse ou multilabel n’est pas trivial. Il s’agit en effet de faire l’hypothèse ou non de l’exclusion mutuelle de chacune des erreurs considérées. Eu égard à la définition donnée des dimensions d’erreur considérées dans la section 3.2, nous considérons que ces dimensions ne peuvent pas être considérées comme mutuellement exclusives. Il apparaît en effet possible qu’un calligraphe novice puisse faire à la fois une erreur d’inclinaison, de taille, et de régularité. Nous nous positionnons donc dans un cadre multilabel.

L’adaptation au cadre multilabel des modèles classiques (permettant une classification binaire) fut ces dernières années le sujet de multiples travaux (Zhang and Zhou, 2014). Deux types de méthodes peuvent être dégagées : la transformation du problème, qui consiste à simplifier le problème de classification multilabel en un problème connu (classification multiclasse, plusieurs sous-problèmes de classification binaire, ordonnancement de classes) ; et l’adaptation du modèle, qui consiste à étendre un modèle de classification binaire afin qu’il puisse gérer plusieurs classes, et notamment sélectionner plusieurs classes pour une même observation. Nous nous positionnons ici dans le cadre de méthodes de transformation du problème, qui permettent d’utiliser les principales méthodes de classification décrites plus tôt dans ce chapitre.

Les méthodes du *label powerset* et de *binary relevance* font partie des méthodes les plus utilisées dans le cadre de la transformation de problème. La méthode du label powerset consiste à transformer le cadre multilabel en un cadre multiclasse, en considérant toute combinaison possible d’étiquettes dans le contexte multilabel en une classe particulière dans le contexte multiclasse. Outre le résultat multilabel (une classe pouvant regrouper une ou plusieurs des étiquettes initiales), il est possible de retrouver la probabilité d’obtenir chacune des étiquettes initiales en calculant la loi marginale sur cette étiquette. Deux des principales critiques émises à l’encontre de cette méthode sont son

5. Suivant les modèles, la notation usuelle veut parfois que les deux valeurs possibles soient -1 et 1 , mais cela n’a que peu d’importance et nous resterons ici sur les valeurs 0 et 1 .

incomplétude (la méthode ne peut prédire que les combinaisons d'étiquettes présentes dans le jeu de données d'entraînement, et ne peut pas généraliser aux combinaisons n'y apparaissant pas), et son inefficacité lorsque le nombre d'étiquettes est grand (le nombre de combinaisons possibles étant de 2^n avec n le nombre d'étiquettes). La binary relevance consiste à créer un jeu de données par classe, et à étiqueter positivement ($y = 1$) les données relatives à cette classe, et négativement ($y = 0$) toutes les autres données. Chaque jeu de données entraîne ensuite un modèle de classification binaire. Pour classer une nouvelle observation, chaque modèle détermine si l'observation appartient ou non à la classe pour laquelle il a été entraîné. Bien que largement utilisé, ce modèle est critiquable du fait qu'il ne prend pas en considération une potentielle corrélation entre données de différentes classes, contrairement au label powerset. Le choix de l'une ou l'autre de ces méthodes dépend donc des caractéristiques du jeu de données. Nous évaluerons ces deux approches dans le cadre de notre expérimentation (section 3.3.2.2).

3.3.2 Intégration dans CalliSmart

Dans ce chapitre, nous avons jusqu'à présent présenté la manière dont nous caractérisons l'activité de l'utilisateur, l'ensemble des feedbacks parmi lesquels l'environnement devra automatiquement choisir l'élément le plus pertinent pour l'apprentissage de l'apprenant, et les travaux relatifs à la reconnaissance de gestes. Dans cette partie, nous présentons plus précisément la façon dont le geste est capturé, puis décrit. Depuis cette description, nous présentons comment les motifs d'erreurs sont reconnus, permettant ensuite, grâce à la politique de sélection illustrant les règles expertes, de sélectionner le feedback le plus pertinent.

3.3.2.1 Capture du geste

Le geste de calligraphie est capturé à l'aide des différents capteurs intégrés dans le stylet ainsi que sur l'écran d'une table Wacom (figure 3.9). Ces capteurs permettent de mesurer la position de l'outil par rapport à l'écran, la pression normale appliquée par l'outil sur l'écran, et ce sur 2048 niveaux de pression. L'inclinaison du stylet est mesurable dans un cône défini par un angle de 40 degrés autour de l'axe normal à l'écran. La fréquence de capture est de 14 millisecondes. L'ensemble des spécifications est disponible sur le site du constructeur⁶.

L'interface de l'application permet à l'utilisateur de démarrer un nouvel exercice. Tout au long de cet exercice, les informations relatives au temps, à la position, à la pression et aux différentes inclinaisons du stylet sont sauvegardées. L'utilisateur informe lui-même de la fin de l'exercice. À cet instant, le vecteur de caractéristiques de l'exercice est calculé. Ces descripteurs sont présentés dans le tableau 3.1.

6. <http://www.wacom.com/en-us/products/pen-displays/cintiq-22-hd#Specifications>

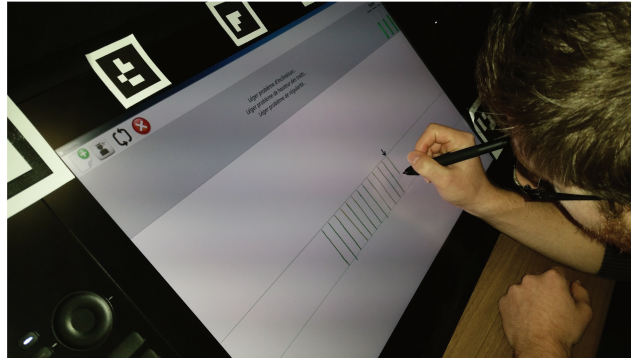


FIGURE 3.9 – La plateforme CalliSmart, intégrée sur une table Wacom

TABLE 3.1 – Caractéristiques calculées depuis l'activité de l'apprenant

Erreur d'inclinaison	Maximum des angles entre l'axe vertical et les tracés. Maximum des angles entre l'axe vertical et les tracés, en interpolant les tracés par des régressions linéaires.
Erreur de régularité	$\frac{\max(\text{distance})}{\min(\text{distance})}$, en utilisant la distance entre le centre de chaque motif. $\frac{\max(\text{area})}{\min(\text{area})}$, en utilisant l'aire entre chaque interpolation de motif par des régressions linéaires. $\frac{(\max(\text{area}) - \min(\text{area}))}{\mu(\text{area})}$, en utilisant l'aire entre chaque interpolation de motif par des régressions linéaires.
Erreur de taille	Maximum des différences entre l'espace vertical entre les deux lignes de portée, et la longueur de chaque motif.

3.3.2.2 Reconnaissance du geste

Une phase de collecte de données, menée à l'ESAD par Patrick Doan et Morgane Rébulard, membres du projet Descript, a permis la création d'une base de données d'entraînement constituée de 138 exercices réalisés par 46 participants n'ayant pas ou peu d'expérience quant à la pratique de la calligraphie et l'utilisation de tablettes graphiques. Pendant cette phase de collecte, il était demandé aux participants d'effectuer trois exercices, chaque exercice étant composé de dix à quinze éléments de tracé. L'ensemble de ces données a été étiqueté par les deux experts via l'examen du geste (chaque exercice étant filmé) et du résultat (une capture d'écran étant enregistrée à la fin de chaque exercice), et suivant les trois classes d'erreurs présentées dans la section 3.2.

Comme relevé précédemment, l'étiquetage est multilabel, puisque plusieurs erreurs peuvent apparaître au sein d'un même exercice. Le tableau 3.2 illustre des caractéristiques du jeu de données. L'ensemble des combinaisons possibles

TABLE 3.2 – Propriétés du jeu de données multilabel

Cardinalité des étiquettes	2.0412
Densité des étiquettes	0.6804
Diversité des étiquettes	7
Diversité normalisées des étiquettes	0.0722

sont présentes dans le jeu de données, puisque les exercices dit “experts”, c’est à dire au cours desquels aucune erreur n’apparaît, ne participent pas à la phase de reconnaissance d’erreurs (il existe alors $2^3 - 1 = 7$ combinaisons). De nombreuses erreurs (2 en moyenne) apparaissent lors des exercices, ce qui engendre une forte densité d’étiquettes assignées à chaque exemple (0.6804), et peut compliquer la tâche de classification.

Nous comparons les deux méthodes de transformation de problème discutées précédemment (binary relevance BR, et label powerset LP), couplées à des modèles bayésiens naïfs. Les tableaux 3.3 à 3.6 présentent les résultats de la reconnaissance en validation croisée, avec 5 sous-ensembles. Ces résultats sont présentés suivant des métriques de validation classiques d’un classifieur multilabel (Zhang and Zhou, 2014). Nous renvoyons donc à (Zhang and Zhou, 2014) pour une description détaillée des métriques permettant l’évaluation dans un cadre multilabel. Nous nous contenterons ici de définir brièvement les différentes catégories de métriques :

- les métriques **orientées exemples** évaluent la performance du modèle sur chaque exemple du jeu de test, puis retournent la moyenne ;
- les métriques **orientées classes** évaluent la performance du modèle sur chaque classe présente dans le jeu de test, puis retournent la moyenne.

Les métriques orientées exemples et classes peuvent une fois encore être divisées selon qu’elles évaluent le modèle suivant une **perspective de classification** (est-ce que la capacité prédictive du modèle se généralise en dehors du jeu d’entraînement) ou suivant une **perspective de ranking** (est-ce que le modèle permet d’établir une relation d’ordre entre les étiquettes). Ces deux perspectives sont intéressantes dans le cadre de CalliSmart, puisque nous souhaitons non seulement reconnaître les erreurs faites par l’apprenant, mais également savoir en quelles proportions sont faites ces erreurs. En effet, nous évaluons l’apprenant sur une dimension en mesurant la probabilité qu’il ait commis l’erreur relative à cette dimension⁷. Il est donc primordial que les probabilités fournies par le modèle sur chacune des classes illustrent l’importance des erreurs commises.

Une première lecture des résultats met en évidence que les meilleurs résultats sont obtenus par la méthode LP, ce qui peut s’expliquer par le fait qu’elle prend en compte les potentielles corrélations entre les classes, contrairement à la méthode BR. Cette dernière se distingue néanmoins par de meilleures performances pour la reconnaissance des erreurs de longueur (tableaux 3.3 et

7. En reprenant le modèle de (Brusilovsky and Millán, 2007), si $p(S_1) = 0$, l’apprenant est considéré comme expert, si au contraire $p(S_1) = 1$, l’apprenant est considéré comme novice.

TABLE 3.3 – Résultats de la classification multilabel sur le jeu de données (classe, classification) en validation croisée sur cinq sous-ensembles

		BR	LP
Inclinaison	Precision	0.8889 ± 0.0832	0.8651 ± 0.0337
	Recall	0.7223 ± 0.1018	0.8845 ± 0.0529
	F1 score	0.7866 ± 0.0692	0.8737 ± 0.0350
Régularité	Precision	0.9183 ± 0.0883	0.8924 ± 0.0619
	Recall	0.7168 ± 0.1331	0.7725 ± 0.0826
	F1 score	0.7966 ± 0.0994	0.8246 ± 0.0634
Longueur	Precision	0.8705 ± 0.0958	0.8056 ± 0.1257
	Recall	0.7739 ± 0.1108	0.7957 ± 0.1227
	F1 score	0.8085 ± 0.0696	0.7885 ± 0.0892

TABLE 3.4 – Résultats de la classification multilabel sur le jeu de données (exemple, classification) en validation croisée sur cinq sous-ensembles

	BR	LP
Hamming loss	0.2446 ± 0.0474	0.2169 ± 0.0468
Subset accuracy	0.4116 ± 0.0983	0.4928 ± 0.0908
F1 score	0.7928 ± 0.0591	0.8368 ± 0.0383
Recall	0.7281 ± 0.0964	0.8213 ± 0.0577
Precision	0.8960 ± 0.0598	0.8645 ± 0.0374

TABLE 3.5 – Aires sous la courbe ROC par motif d'erreur (classe, ranking) en validation croisée sur cinq sous-ensembles

	BR	LP
Inclinaison	0.7760 ± 0.2236	0.8347 ± 0.1069
Régularité	0.8386 ± 0.0550	0.8652 ± 0.0645
Longueur	0.9502 ± 0.0280	0.9464 ± 0.0405

TABLE 3.6 – Résultats de la classification multilabel sur le jeu de données (exemple, ranking) en validation croisée sur cinq sous-ensembles

	BR	LP
Coverage	2.1756 ± 0.0372	2.1474 ± 0.0741
Ranking Loss	0.0660 ± 0.0206	0.0577 ± 0.0297
Average Precision	0.9593 ± 0.0276	0.9734 ± 0.0130

TABLE 3.7 – Comparaison des aires sous la courbe ROC entre les modèles bayésiens naïfs (génératifs) et régressions logistiques (discriminants)

	Naïve Bayes	Régressions Logistiques
Inclinaison	0.8347 ± 0.1069	0.5423 ± 0.1453
Régularité	0.8652 ± 0.0645	0.8371 ± 0.0456
Longueur	0.9464 ± 0.0405	0.9560 ± 0.0333

3.5). De manière plus générale, et comme les tableaux 3.3 et 3.4 l’illustrent, la méthode BR obtient une précision intéressante (sensiblement meilleure que la méthode LP), mais un rappel nettement inférieur à la méthode LP. Cette nette différence se retrouve dans le tableau 3.5 qui illustre les aires sous la courbe ROC, c’est à dire le ratio entre le taux de vrais et de faux positifs. L’aire sous la courbe precision-recall (Average Precision, tableau 3.6) montre également la supériorité de la méthode LP. Ainsi la méthode LP semble à privilégier dans notre contexte.

La méthode LP permet d’identifier exactement près de la moitié des échantillons (*subset accuracy*, tableau 3.4). Si ce résultat peu paraître décevant, le *subset accuracy* est connue pour être une métrique très stricte, d’autant plus lorsque, comme ici, la quantité de données par classe est peu élevée, et la densité d’étiquette par échantillon est importante (Tsoumakas and Katakis, 2006).

Remarque. *Afin de renforcer le choix des modèles bayésiens naïfs par rapport aux régressions logistiques, choix basé sur les conclusions de (Ng and Jordan, 2002) préférant l’approche générative à l’approche discriminante lorsque la quantité de données d’entraînement est faible, nous avons souhaité comparer ces deux approches dans le cadre de la méthode LP. Les modèles bayésiens naïfs ont confirmé leur supériorité dans ce cadre précis, comme l’illustre le tableau 3.7 (dans un but de simplicité, seules les métriques de classification obtenues par une moyenne sur les exemples ont été présentées). En particulier, les régression logistiques montrent un grand nombre de faux positifs sur la classe inclinaison, qui est la classe la plus représentée.*

3.3.2.3 Décision

Depuis la capture de l’activité de l’apprenant sur la table Wacom, le module de reconnaissance permet la constitution d’une représentation de l’apprenant à la fin de chaque exercice. Cette représentation positionne l’apprenant dans un espace dont les trois dimensions illustrent les types d’erreurs sélectionnés par les experts : inclinaison, régularité et taille. On obtient, à la fin de chaque exercice, un vecteur de représentation $\mathbf{s} = \{p(\text{inclinaison}), p(\text{regularite}), p(\text{taille})\}$ contenant les probabilités d’erreur. Cette représentation reprend le modèle overlay de (Brusilovsky and Millán, 2007), les vecteurs de représentation $s_{\text{expert}} = \{0, 0, 0\}$ et $s_{\text{novice}} = \{1, 1, 1\}$ représentant respectivement l’expert et le novice. La similarité entre deux apprenants caractérisée par les vecteurs \mathbf{s}_1 et \mathbf{s}_2 peut alors se mesurer par une similarité globale $\text{sim}(\mathbf{s}_1, \mathbf{s}_2)$.

Les experts peuvent alors définir une représentation des actions pédagogiques en déterminant, pour chaque action, la situation dans laquelle elle est, d'après leur expertise, la plus appropriée. Pour chaque action $i \in A$, cela revient, en reprenant les notations du chapitre 2, à définir un vecteur de représentation $\mathbf{s}^{(i)} \in S$. En indexant les dimensions d'erreur de la manière suivante :

- S_1 = erreur d'inclinaison ;
- S_2 = erreur de régularité ;
- S_3 = erreur de taille,

et les feedbacks de la manière suivante :

- $s^{(1)}$ = feedback inclinaison ;
- $s^{(2)}$ = feedback régularité ;
- $s^{(3)}$ = feedback taille,

l'expertise fournit les représentations d'action suivantes :

- $\mathbf{s}^{(1)} = \{1, 0, 0\}$
- $\mathbf{s}^{(2)} = \{0, 1, 0\}$
- $\mathbf{s}^{(3)} = \{0, 0, 1\}$

L'expertise exprime un fort couplage entre action pédagogique et erreur, c'est à dire que l'expert considère chaque feedback comme répondant directement à une erreur particulière. Dans les prochains chapitres, nous étudierons un contexte où cette relation n'est pas aussi bien établie, un même feedback pouvant permettre à l'apprenant de travailler sur plusieurs dimensions de son activité.

3.4 Expérimentation

Nous souhaitons, au cours d'une première expérience, mesurer l'impact sur l'apprentissage d'un environnement proposant automatiquement des feedbacks en se basant sur des règles expertes. En particulier, nous souhaitons mesurer la performance des apprenants en conditions réelles après la période d'entraînement, afin d'identifier un éventuel syndrome de dépendance à l'environnement d'apprentissage. Enfin, nous souhaitons évaluer l'adoption par les utilisateurs de cet environnement. Nous utilisons pour cela la plateforme CalliSmart présentée dans ce chapitre.

3.4.1 Hypothèses

L'expérimentation doit permettre de tester les hypothèses suivantes :

- hypothèse 1 (**H1**) : les retours diversifiés et sélectionnés depuis l'expertise permettent aux apprenants d'obtenir de meilleures performances lors de la phase d'apprentissage. Les retours permettent en outre aux apprenants de rester engagés dans la tâche, ce qui induit une réduction de la variance sur les performances ;
- hypothèse 2 (**H2**) : les retours diversifiés et sélectionnés depuis l'expertise permettent aux apprenants de reproduire leurs progrès en conditions réelles (sans feedback) ;

- hypothèse 3 (**H3**) : les apprenants reconnaissent la capacité de l’environnement à sélectionner des actions adaptées à leurs besoins.

3.4.2 Protocole

L’expérimentation se déroule suivant le protocole suivant. Chaque participant peut interagir avec l’environnement pendant le temps qu’il juge nécessaire, afin de se familiariser avec l’interface. Une fois familiarisé avec l’environnement, il est demandé à chaque participant de calligraphier cinq séries (une série est considérée comme un exercice), chaque série étant composée de dix à quinze éléments (les éléments verticaux du mot minimum). Les dimensions d’évaluation (inclinaison, taille des traits, régularité spatiale) sont expliquées aux participants avant le premier exercice. Les participants sont divisés en deux groupes de manière aléatoire :

- le groupe **contrôle** travaille sans feedback. Ce groupe permet d’avoir une valeur étalon de la progression naturelle des apprenants lors de pratiques successives en conditions réelles. La progression naturelle, dans le cadre d’exercices courts impliquant une gestuelle simple s’avère parfois aussi, voire plus importante que la progression lors d’entraînement sur un système d’aide (Smith et al., 2006) ;
- le groupe **expert** se voit proposer un feedback par exercice. Ce feedback est sélectionné en fonction de la performance accomplie lors de l’exercice précédent, en utilisant les modules de reconnaissance et de décision décrits dans ce chapitre (sections 3.3.2.2 et 3.3.2.3).

Suite aux cinq séries constituant la phase d’entraînement, il est demandé à chaque participant d’accomplir un dernier exercice, sans feedback (aussi bien pour le groupe contrôle que pour le groupe expert). Ce dernier exercice nous permet de mesurer la faculté des participants du groupe expert à reproduire leurs progrès en conditions réelles. À noter que nous entendons par “conditions réelles” la pratique de la calligraphie sur la table Wacom, sans feedback. Nous ne mesurons pas ici la faculté des apprenants à reproduire leurs progrès sur papier. Cette étude reste cependant une question intéressante qui mériterait d’être étudiée dans le futur. Finalement, un bref questionnaire est proposé aux participants. Les participants des deux groupes doivent y noter, entre 1 (aucune progression) et 10 (forte progression), leur perception de leurs performances successives lors de l’expérimentation. Les participants du groupe expert doivent répondre à deux questions supplémentaires : ont-ils le sentiment que le système reconnaît leurs erreurs (note de 1 ; reconnaissance très imprécise, à 5 ; reconnaissance très précise) ; ont-ils l’impression que les aides proposées sont adaptées aux erreurs commises (note de 1 ; pas adaptées du tout, à 5, parfaitement adaptées).

3.4.3 Participants

28 participants ont participé à cette étude. Ces participants sont principalement des étudiants en génie informatique, design et génie mécanique, n’ayant

pas ou très peu d'expérience quant à la pratique de la calligraphie et l'utilisation de table graphique. La distribution des participants au sein des deux groupes était faite de façon aléatoire, en s'assurant tout de même de garder un nombre égal de participants dans chacun des groupes.

3.4.4 Évaluation

L'évaluation de la performance des participants est mesurée par rapport à des exercices experts réalisés dans le cadre de la campagne de mesure ayant permis la constitution d'un jeu de données d'entraînement. D'une façon similaire à l'algorithme des k plus proches voisins, nous évaluons chaque vecteur caractérisant la performance du participant par la distance moyenne aux vecteurs caractérisant les trois performances expertes les plus proches. Cette méthode permet de réduire le potentiel biais que pourrait provoquer l'utilisation d'un modèle paramétrique.

3.4.5 Résultats et discussion

L'expérimentation ne réunissant qu'un nombre réduit d'utilisateurs, le test de Shapiro-Wilk appliqué aux données a montré que ces dernières ne suivaient pas une distribution normale, ce qui ne permet pas d'utiliser un test paramétrique. Prenant acte de ces résultats, nous utilisons le test non-paramétrique de Mann-Whitney pour valider nos hypothèses.

La figure 3.10 illustre une amélioration significative des performances des participants du groupe expert, depuis une distance médiane de 2.88 à la fin du premier exercice, vers une distance de 1.75 à la fin du cinquième. Le groupe contrôle ne progresse quant à lui que très faiblement, depuis une distance de 3.20 vers une distance de 3.11. La différence entre les deux groupes lors du cinquième exercice est significative, avec une p -valeur égale à 0.029 résultant du test de Mann-Whitney. La figure 3.10 fait également apparaître une diminution de la variance au long des itérations pour les participants du groupe expert. Ceci conforte notre première hypothèse **H1**. Les résultats obtenus par le groupe contrôle (similaires aux résultats du groupe expert sur les premières itérations, puis plus distants des exemples experts jusqu'à la fin du processus) peuvent être expliqués par deux facteurs que sont la difficulté à évaluer la performance, et la lassitude. La tâche proposée est en effet répétitive, et les participants du groupe contrôle n'ayant que peu de connaissances en calligraphie, et ne recevant pas d'information de l'environnement, ne pouvaient percevoir s'ils évoluaient ou non dans la bonne direction. À partir du quatrième exercice, ces participants semblaient subir une perte d'intérêt et de concentration vis-à-vis de la tâche. Ces conclusions sont tout de même à mettre en perspective, puisque certains participants du groupe contrôle ont continué à progresser. Cette différence s'illustre dans les résultats par la grande variance intra-groupe pour le groupe contrôle. Cette différence s'illustre également en terme de perception de leur progression (figure 3.12), où l'on peut noter que les participants du groupe

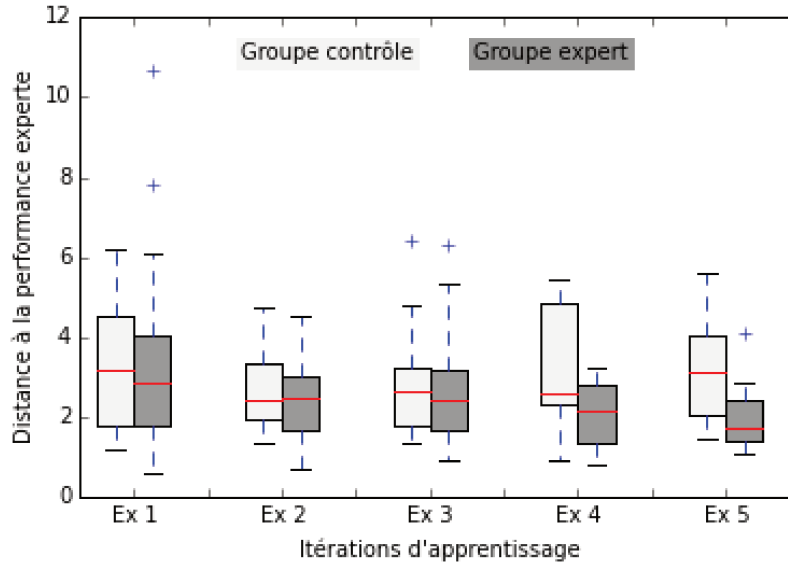


FIGURE 3.10 – Diagramme en boîtes illustrant la distance à la performance experte des participants des groupes contrôle (sans feedback) et expert (avec feedback et adaptation basée sur des règles expertes)

contrôle peuvent avoir des impressions très différentes quant à leur progression au cours de l'expérimentation.

Les résultats des deux groupes lors du dernier exercice (post-test, les deux groupes étant dans les mêmes conditions, c'est à dire sans aide de l'environnement, figure 3.11) montrent un regain d'intérêt des participants du groupe contrôle, dont les performances restent malgré tout en deçà de celles accomplies par les participants s'étant entraînés dans la configuration avec feedback. Nous notons surtout que les performances du groupe expert lors de ce post-test restent proches de celles accomplies avec des aides (la distance aux experts n'augmentant que de 1.75 à 1.88 entre le dernier exercice avec aides et le post-test). Ce résultat est encourageant en ce qui concerne le syndrome de dépendance à l'environnement d'apprentissage. Il est cependant nécessaire de nuancer ces commentaires au regard de l'augmentation de la variance entre les performances lors de ce dernier exercice, qui se traduit par une p -valeur de 0.0549 entre les performances des deux groupes, ce qui est significatif pour un seul de 0.1. Notre seconde hypothèse **H2** semble donc se vérifier, mais une expérimentation complémentaire, rassemblant un plus grand nombre de participants, serait nécessaire pour émettre des conclusions plus affirmées.

Les diagrammes 3.12 à 3.14 résument les résultats du questionnaire proposé. La période d'entraînement étant très courte, il n'est pas étonnant que le sentiment de progression soit mitigé. Il est toutefois intéressant, comme

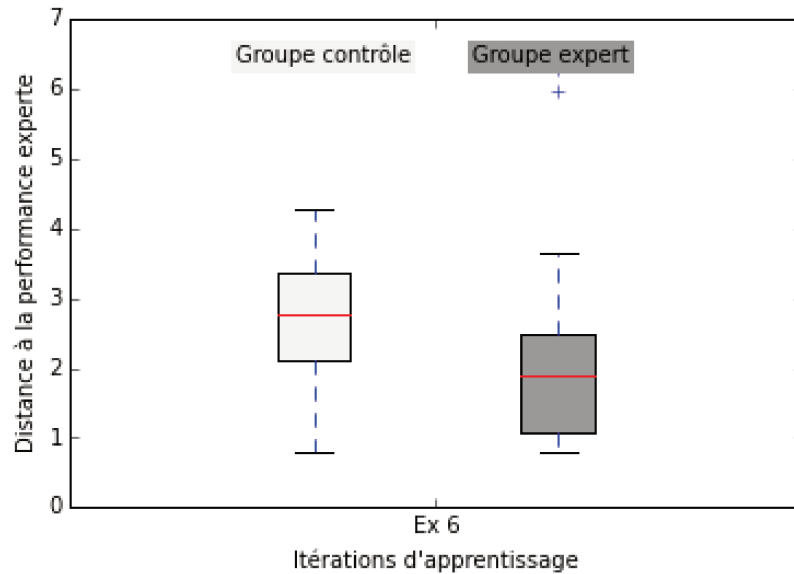


FIGURE 3.11 – Diagramme en boîtes illustrant la distance à la performance experte des participants des groupes contrôle (sans feedback) et expert (avec feedback et adaptation basée sur des règles expertes) lors du post-test

décrit précédemment, de noter les fortes disparités de réponses au sein du groupe contrôle. Nous notons également, bien que cette variance très importante nuance ce résultat, qu'en moyenne les participants du groupe expert ont davantage l'impression d'avoir progressé. Malgré la brièveté de la période d'entraînement, leur impression est en moyenne positive (5.72/10). Le diagramme 3.13 montre la bonne réception des participants par rapport à la reconnaissance automatique des erreurs fournies par le système, 12 participants sur 14 ayant donné une note très positive (supérieure ou égale à 4), et aucun n'ayant donné une note négative. L'avis est plus partagé en ce qui concerne l'adaptation de la sélection de feedback par rapport aux erreurs (figure 3.14). Cette réticence vis-à-vis de la sélection automatique peut être expliquée par deux raisons : le faible nombre d'actions pédagogiques possibles, et la rigidité des règles expertes. Il est tout à fait envisageable, en ne proposant que trois types d'actions pédagogiques, qu'aucune d'entre elles ne conviennent à certains participants. Il serait donc intéressant de proposer une expérimentation de plus grande envergure, avec un environnement pouvant choisir parmi un plus grand nombre de possibilités. Nous nous intéressons davantage à l'aspect monolithique des règles expertes : si certains participants se sont montrés convaincus par la sélection faite par l'environnement, d'autres se sont agacés de se voir proposer de manière répétée le même feedback, alors même qu'ils ne progressaient pas. En effet, si le feedback proposé reste le plus proche de leur représentation, il conti-

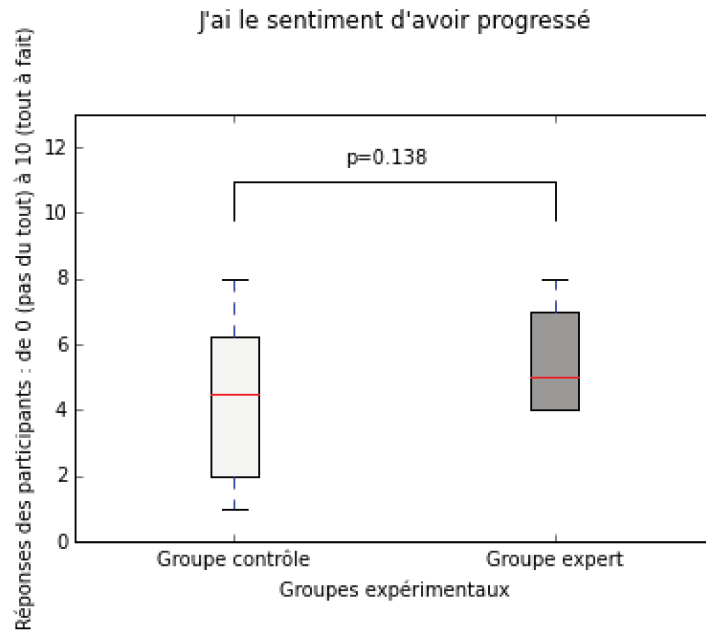


FIGURE 3.12 – Réponses des participants face à la proposition : “J’ai le sentiment d’avoir progressé”

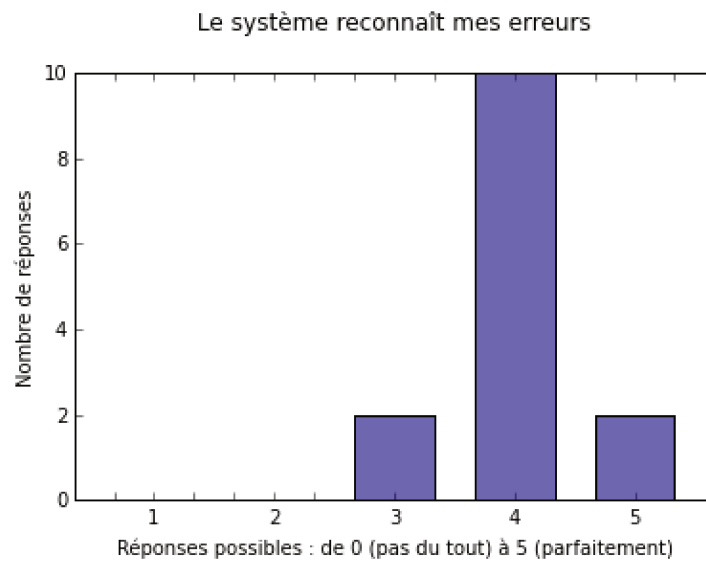


FIGURE 3.13 – Réponses des participants du groupe expert face à la proposition : “Le système reconnaît mes erreurs”

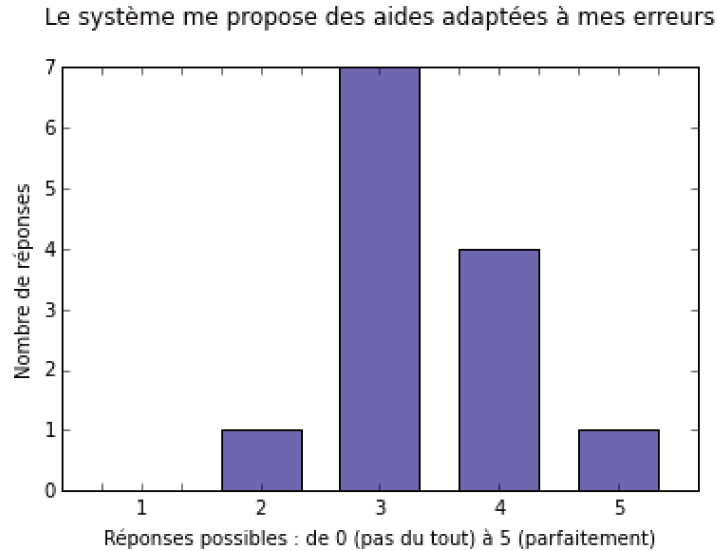


FIGURE 3.14 – Réponses des participants du groupe expert face à la proposition : “Le système me propose des aides adaptées à mes erreurs”

nuera à être sélectionné malgré le fait que l’apprenant ne progresse pas. Pour pallier ce problème, il est nécessaire que l’environnement puisse, non seulement suivre des règles de décision lui permettant de sélectionner le feedback le plus approprié à la situation, mais également remettre en cause, le cas échéant, ces règles qui régissent la sélection.

3.5 Conclusions

Dans ce chapitre, nous avons présenté la première version de la plateforme CalliSmart développée dans le cadre de nos travaux. Nous avons présenté la manière dont nous interprétons l’activité de l’apprenant, en caractérisant des dimensions d’erreur et en mesurant son niveau sur chacune de ces dimensions par la probabilité que l’erreur ait été commise via des classifieurs bayésiens naïfs. Nous avons pour cela présenté les principaux modèles présents dans la littérature, et présenté une comparaison, en utilisant des données issues d’exercices de calligraphie, des méthodes de binary relevance et de label powerset, permettant une classification multilabel.

Nous avons présenté l’ensemble des actions pédagogiques implémentées dans le cadre d’une première expérimentation visant à mesurer l’impact d’un environnement adaptatif pour l’apprentissage de la calligraphie. Nous concluons de cette expérimentation que proposer une diversité de feedback, l’environnement sélectionnant le feedback à proposer en se basant sur des règles de décision expertes, est bénéfique pour l’apprentissage. En outre, l’amélioration des per-

performances enregistrées lors des phases d'apprentissage se retranscrit lors des performances effectuées en conditions réelles suite à l'apprentissage.

Toutefois, la variance importante accompagnant les performances du groupe travaillant sur l'environnement adaptatif, ainsi que les avis divergeants des participants quant à la pertinence de la sélection des feedbacks, sont autant d'indices nous amenant à penser qu'il serait intéressant de pouvoir fournir un degré d'adaptation supplémentaire au système. Par degré d'adaptation supplémentaire, nous entendons la possibilité, non seulement de pouvoir sélectionner différents feedbacks pour faire face à différentes situations, mais également de pouvoir remettre en cause ces règles suite aux interactions avec l'apprenant, afin de répondre au mieux à ses besoins. Le prochain chapitre porte sur cette problématique d'adaptation. Nous y décrirons les principaux travaux relatifs à cette question, et nous proposerons un modèle théorique, reprenant le cadre présenté dans le chapitre 2, permettant d'adapter dynamiquement les règles de sélection d'actions pédagogiques.

Chapitre 4

Adaptation automatique : Modélisation de l'interaction

4.1 Motivations

Dans le chapitre 3, nous avons présenté une première expérimentation où nous considérons l'expertise comme suffisamment précise pour approcher la politique optimale π^* , cette politique étant la même pour tous les apprenants. Les actions pédagogiques proposées ayant été directement pensées par les experts par rapport aux dimensions considérées de l'apprentissage, l'environnement a permis aux apprenants de progresser lors de la phase d'entraînement, tout en limitant les effets du syndrome de dépendance à l'environnement d'apprentissage. Le contexte dans lequel nous nous sommes positionnés au sein du chapitre 3 est néanmoins particulier : il n'est pas toujours possible de déterminer quelle action pourra s'avérer avantageuse par rapport à une situation (e.g. erreur) donnée ; il est souvent difficile, et coûteux, de réunir les connaissances expertes permettant de faire le lien entre situation d'apprentissage et sélection d'actions pédagogiques ; les connaissances récupérées peuvent s'avérer imprécises et/ou subjectives, et ne peuvent pas considérer le caractère individuel de l'apprenant (il n'est d'ailleurs pas directement possible d'analyser si le facteur déterminant de l'apport sur l'apprentissage, lors de l'expérience du chapitre 3, fut lié à l'adaptation ou simplement à la diversité des actions pédagogiques proposées).

Puisque la problématique liée à une expertise incomplète (ne considérant pas tous les paramètres inhérents à l'apprentissage), ou imprécise (différence non négligeable entre l'expertise et la réalité) se démocratise au même titre que la recherche sur les apprentissages dans les domaines mal définis (Lynch et al., 2006; Conati, 2009; Fournier-Viger et al., 2010b), nous nous intéressons dans ce chapitre à la manière dont l'environnement pourrait adapter automatiquement, et dynamiquement, les règles permettant la sélection d'actions en fonction de la situation d'apprentissage. Cette adaptation doit permettre à l'environnement de réduire l'écart entre réalité et expertise dû aux problèmes d'incomplétude et d'imprécision de cette dernière.

La partie 4.2 dresse un tour d'horizon des principaux travaux liés à l'adaptation de la sélection d'actions pédagogiques au sein des environnements d'apprentissage. Suite à cet état de l'art, nous motivons et présentons notre approche, qui étend le modèle proposé dans le chapitre 2 en modélisant les interactions entre l'environnement et l'apprenant comme un problème de bandit.

4.2 Travaux similaires

Les environnements d'apprentissage "réellement" adaptatifs, c'est à dire qui ne se contentent pas de s'adapter à l'activité de l'apprenant en suivant des règles, mais modifient dynamiquement ces règles en fonction de l'apprenant, ont vu leur popularité croître ces dernières années. Cette problématique peut être vue comme un problème d'optimisation : quelle est l'activité à proposer à l'apprenant (ou dans notre cas, quels sont les feedbacks à proposer lors de l'exercice) qui, étant donnée sa représentation dans l'environnement, maximisera certains aspects de son apprentissage. Cette analogie entre sélection d'actions pédagogiques et optimisation a amené la communauté à s'intéresser aux techniques relatives aux séquences de prises de décision dans des contextes incertains (Chi et al., 2011a,b).

La démocratisation des cours en ligne (*Massive Open Online Courses*, ou *MOOCs*), qui rassemblent parfois des milliers voire des millions d'utilisateurs, a provoqué non seulement une prise de conscience des apports de tels environnements (accessibilité et individualisation de la formation), mais a également permis la collecte de quantités de données relatant l'apprentissage jamais vues jusqu'alors. Ces nouveaux jeux de données ont permis d'ouvrir le domaine des environnements d'apprentissage à des méthodes telles que les systèmes de recommandation. Preuve de cette ouverture, depuis 2010 se tient tous les deux ans le workshop RecSysTEL (Recommender Systems for Technology-Enhanced Learning), désormais renommé EdRecSys (Educational Recommender Systems). D'autres approches, telles que les approches par renforcement, ont également été proposées. Nous proposons ici un tour d'horizon des principales méthodes proposées ces dernières années.

4.2.1 Systèmes de recommandation

Très utilisés dans le cadre du web (e.g. publicité ciblée, recommandation de contenu comme des séries TV ou des films), les systèmes de recommandation se positionnent dans la catégorie du filtrage de l'information. Les systèmes de recommandation cherchent ainsi à prédire l'avis que l'utilisateur pourrait avoir vis-à-vis d'une ressource. Dans le cadre d'environnement d'apprentissage, les systèmes de recommandation peuvent chercher à prédire l'impact (plutôt que l'avis) que pourra avoir une ressource (e.g. un feedback) sur l'apprentissage de l'apprenant¹. Plusieurs approches ont été proposées ces dernières années pour

1. D'autres problématiques, telles que la recommandation d'activités liées à la géolocalisation ou à l'environnement (e.g. proposer une activité simple lorsque l'utilisateur est dans un

établir des profils d'utilisateur et de ressource permettant de prédire l'avis d'un utilisateur par rapport à une nouvelle ressource.

4.2.1.1 Approches basées sur le contenu

Les approches basées sur le contenu (Pazzani and Billsus, 2007) cherchent à filtrer les données par rapport aux propriétés de l'utilisateur d'une part, et de la ressource d'autre part. L'intuition derrière cette approche consiste à penser que des utilisateurs appartenant au même type de profil auront tendance à partager les mêmes avis concernant un même type de ressource. (Drachsler et al., 2008) identifient les méthodes de raisonnement à partir de cas et les méthodes basées sur les attributs comme deux des principales approches basées sur le contenu.

Le *raisonnement à partir de cas* (Bridge et al., 2005) recommande des ressources similaires (en utilisant une mesure de similarité tenant compte des différentes propriétés de la ressource) à des ressources ayant précédemment reçu un avis favorable de l'utilisateur. Cette méthode montre de bons résultats mais nécessite que l'utilisateur ait déjà donné son avis plusieurs fois, sur différents types de ressources. Cette technique ne permet donc pas de recommander un contenu à un nouvel utilisateur (on parle de démarrage à froid, ou *cold start*). Une autre critique adressée au raisonnement à partir de cas est son incapacité à donner un quelconque avis concernant une ressource très différente de celles déjà notées par l'utilisateur (e.g. si un utilisateur ne note que des films romantiques et d'action, il est impossible de lui recommander une comédie).

Les méthodes *basées sur les attributs* (Drachsler et al., 2008) ne sont pas sujettes à ces deux problèmes, et se basent sur une expertise reliant une ressource à un type d'utilisateur, d'une manière assez proche de ce que nous avons présenté dans les chapitres 2 et 3. Cette approche souffre néanmoins des mêmes inconvénients que les méthodes présentées dans ces chapitres : le coût de création des règles et leur précision, la nature monolithique de ces règles, qui ne permettent pas la personnalisation de la recommandation, la modularité du système qui nécessite que tout ou partie des règles soient revues dès lors que l'on ajoute une nouvelle ressource.

4.2.1.2 Approches sociales

Les approches sociales, dont le principal représentant est la méthode de filtrage collaboratif (Herlocker et al., 2004) se basent sur les avis proposés par toute la communauté lors de la prédiction d'une ressource pour un utilisateur. Les approches collaboratives peuvent se focaliser sur les utilisateurs, en établissant des corrélations entre avis pour former des groupes d'utilisateurs. Un utilisateur appartenant à un groupe se verra alors recommandées des ressources

environnement bruyant, inciter plusieurs utilisateurs proches les uns des autres à collaborer sur un exercice) sont également de plus en plus étudiées. Nous nous focalisons cependant ici sur la problématique qui est la nôtre, la sélection d'action par rapport à une représentation de l'activité de l'apprenant. Des informations complémentaires à propos des différentes problématiques de recommandation et d'apprentissage peuvent être trouvées dans (Verbert et al., 2012).

ayant reçu de bons avis par les autres membres de ce groupe. Le filtrage collaboratif peut également se focaliser sur les ressources en établissant de la même manière des groupes de ressources similaires, et en proposant aux utilisateurs des ressources similaires à d'autres ressources ayant reçu des avis positifs. Le principal avantage du filtrage collaboratif est de profiter de la masse d'information collectée par l'ensemble des utilisateurs. Ces méthodes souffrent par contre, comme le raisonnement à partir de cas, du problème de démarrage à froid lorsqu'un nouvel utilisateur/une nouvelle ressource est ajouté(e) au système. Les approches sociales sont également très dépendantes des modélisations parfois peu précises des utilisateurs ou des ressources, amenant alors le système à proposer des ressources incohérentes (Burke, 2002). Enfin, les recommandations proposées tendent à privilégier les ressources populaires aux ressources moins populaires (bien que de qualités égales), et les utilisateurs ayant des goûts différents de la masse ne recevront que très peu de ressources qu'ils jugeront pertinentes (syndrome du mouton gris, (Burke, 2002)).

4.2.1.3 Recommandation pour les environnements d'apprentissage

Initialement développés dans des contextes tels que la publicité ciblée ou la recommandation de films, les systèmes de recommandation ne sont pas directement transposables aux environnements d'apprentissage. Ces derniers proposent en effet des problématiques nouvelles identifiées par (Drachler et al., 2008) :

- la prise en compte du contexte : les besoins de l'apprenant ne dépendent pas uniquement de son profil d'apprentissage (en imaginant qu'il soit identifiable), mais également du contexte d'apprentissage (i.e. qu'est-ce que l'apprenant est en train de travailler, dans quel état est-il ?) ;
- la spécificité des besoins de l'apprenant : bien qu'il soit potentiellement possible de dégager des profils d'apprentissage, il existe toujours une part de spécificité quant aux besoins des apprenants, ces derniers pouvant avoir une sensibilité, ou une compréhension, différente de l'activité ou du feedback proposé ;
- la difficulté des apprenants à appréhender l'utilité du comportement adopté par l'environnement (choix d'une activité, d'un feedback) : les systèmes de recommandation se reposent sur une phase d'acquisition de données où les utilisateurs sont amenés à noter les ressources qui leur sont proposées (e.g. note sur cinq pour un film, note binaire déterminant si une information est utile ou non). Si chacun peut déterminer à quel point il apprécie un film, il est rare qu'un apprenant (d'autant plus s'il est novice) puisse déterminer l'utilité d'une activité, ou d'un feedback, sur son apprentissage. Ainsi, un environnement sélectionnant le comportement à adopter depuis l'avis d'utilisateurs apparaît être, dans le cadre de l'apprentissage, un exercice non seulement coûteux mais incertain.

4.2.1.4 Commentaires

En somme, les systèmes de recommandation constituent une piste de recherche importante pour les environnements d'apprentissage, grâce notamment au développement d'outils de formation en ligne qui favorisent la collecte de quantités importantes de données. La pertinence de tels outils dans des cadres où cette collecte est plus complexe (e.g. peu d'apprenants, impossibilité pour les apprenants d'évaluer la pertinence des feedbacks/activités proposées) est néanmoins discutable, ou nécessite tout du moins une révision en profondeur des théories constituant la fondation même de ces approches.

4.2.2 Approches par renforcement

L'apprentissage par renforcement (Sutton and Barto, 1998) constitue, avec l'apprentissage supervisé, l'une des principales thématiques de l'apprentissage automatique étudiées ces trois dernières décennies. Contrairement à l'apprentissage supervisé, qui consiste à apprendre depuis un ensemble d'informations qui peut être vu comme une expérience, l'apprentissage par renforcement consiste à apprendre en interagissant, l'information étant acquise au fil de ces interactions. Les actions prises par l'environnement se voient attribuer un gain, et l'objectif de l'apprentissage par renforcement est d'apprendre quand et quelle action choisir afin de maximiser ce gain. Deux principaux enjeux se dressent alors : la modélisation de l'influence des actions sélectionnées, c'est à dire la façon dont l'environnement évalue le gain d'une action ; et le raisonnement qui permettra de converger vers des règles de sélection qui maximisent la somme des gains. On reconnaît ici les deux principales problématiques qui sont les nôtres, la question de la modélisation ayant été adressée lors des précédents chapitres. Concernant cette seconde problématique, l'un des principaux dilemmes est celui du compromis entre exploration et exploitation. L'exploration signifie la sélection de nouvelles actions dont la pertinence est jusqu'alors inconnue, ce qui peut permettre de trouver une meilleure solution, mais en contrepartie peut également résulter en une dégradation des performances si cette action n'est finalement pas pertinente. L'exploitation réside au contraire en la sélection des seules actions dont le système possède déjà des connaissances. Intuitivement, ne faire que de l'exploitation a l'avantage de réduire à zéro le risque de choisir une action moins pertinente que celles connues. Néanmoins, cela réduit également à zéro la chance de découvrir une action qui s'avèrera plus, voire beaucoup plus pertinente que celles utilisées jusque là.

De manière générale, un modèle d'apprentissage par renforcement se caractérise par trois éléments :

- l'espace des états possibles $S = \{S_1, \dots, S_m\}$;
- l'espace des actions possibles $A = \{A_1, \dots, A_k\}$;
- le modèle de gain $\nu = \mu(s, s', a_i)$, $s, s' \in S$, $a_i \in A$ qui assigne un gain à l'action i ayant permis la transition de l'état s à l'état s' .

L'objectif du modèle est de trouver la politique π^* qui relie chaque état à l'action qui maximisera le gain. On retrouve bien évidemment ici une par-

tie du cadre théorique présenté dans le chapitre 2. Les méthodes par renforcement peuvent être divisées en deux catégories : les méthodes avec et sans modèles (Kaelbling et al., 1996). Nous présentons ici le principe de ces deux catégories, et leur utilisation dans le cadre d'environnements d'apprentissage adaptatifs.

4.2.2.1 Approches sans modèle

Les approches sans modèles apprennent directement une fonction objectif depuis les interactions entre l'agent apprenant (dans notre contexte, l'environnement d'apprentissage), et les agents extérieurs (l'apprenant). Au fil de ces interactions, l'agent apprenant précise sa politique de décision mais il n'acquiert pas de connaissances explicites sur l'environnement (i.e. structure des états, transitions entre les états). Ces approches, qui rassemblent entre autres les méthodes de Monte Carlo et les méthodes de différences temporelles, sont privilégiées lorsque la collecte d'information ne constitue pas un problème (Chi et al., 2011a).

Dans le cadre d'environnements d'apprentissage, ces méthodes ont été employées en utilisant, pour la phase d'apprentissage, des apprenants simulés (Beck et al., 2000; Iglesias et al., 2009). La simulation d'apprenants permet en effet de collecter une masse importante de données à moindre coût. Cette technique (simulation pour l'entraînement d'une méthode par différences temporelles) a permis la réduction significative du temps de réponse de l'apprenant dans le cadre du tuteur AnimalWatch (Beck et al., 2000) (ce qui, comme noté par (Chi et al., 2011a), ne présage en rien de la qualité de l'apprentissage). Dans le cadre de l'apprentissage de la conception de base de données (Iglesias et al., 2009), la simulation (permettant ici d'entraîner un modèle de Q-learning) a permis à l'environnement de proposer aux apprenants des conseils visant à augmenter leur efficacité en réduisant les ressources consultées aux plus pertinentes, ce qui n'a par contre pas induit de différences significatives en terme d'apprentissage.

L'utilisation de données simulées reste une méthode critiquable, puisque la précision des politiques collectées devient entièrement dépendante de la faculté de la simulation à retranscrire le comportement (souvent très complexe, et toujours mal, ou uniquement partiellement, compris). Utiliser de "vraies" données peut également être un exercice compliqué dans le cadre d'environnement d'apprentissage, ces derniers pouvant avoir, au fil des travaux de recherche, des buts différents. Ainsi (Tetreault and Litman, 2006) utilisent des données d'un précédent système. L'objectif du précédent système étant légèrement différent du système final, la politique générée par le modèle par renforcement s'est avérée incomplète. C'est pourquoi, comme le conclut (Chi et al., 2011a), la seule méthode optimale d'acquisition de données apparaît être l'utilisation du système dans sa configuration finale. Ceci nécessite un investissement important en terme de temps et de coût puisqu'à chaque amélioration, l'ensemble des données précédemment générées devient inutilisable.

4.2.2.2 Approches avec modèle

Les approches avec modèle, qui s’opposent aux approches sans modèle présentées précédemment, apprennent la structure de transition entre les états, et visent à travers cette structure à déterminer l’action à privilégier. Cette structure² peut être modélisée comme un processus de décision Markovien (MDP), ou comme un processus de décision Markovien partiellement observable (POMDP), ces derniers permettant, de manière analogue aux modèles de Markov à états cachés, de différencier états et observations.

Ces approches, qui incluent les méthodes de programmation dynamique comme l’itération sur les valeurs, ou l’itération sur les politiques, restent applicables dans des contextes où la collecte d’information est coûteuse (Chi et al., 2011a)³. Les approches basées sur les modèles sont par contre computationnellement plus coûteuses, puisqu’elles reposent sur de nombreux paramètres définissant la structure du modèle. Ces paramètres sont en outre difficilement identifiables (Beck and Chang, 2007), et l’on rejoint ici les problématiques de domaines bien et mal définis énoncées dans le chapitre 2.

4.2.2.3 Commentaires

Les approches avec et sans modèle proposées dans le cadre de l’apprentissage par renforcement commencent à être utilisées dans le cadre d’environnement d’apprentissage. Toutefois, lorsque l’acquisition d’une quantité suffisamment grande d’exemples d’apprentissage est impossible (dans le cadre d’approches sans modèle), ou lorsqu’il est difficile de définir un modèle (dans le cadre d’approches avec modèles), l’apport de telles approches reste un problème ouvert.

4.2.3 Problème de bandit

Le problème de bandit (Bubeck and Cesa-Bianchi, 2012), ainsi nommé en référence aux machines à sous⁴, étudie le compromis entre exploitation et exploration présent dans les problèmes d’apprentissage par renforcement. L’analogie aux machines à sous est la suivante : un joueur de casino fait face à une machine à sous possédant k leviers. Chaque levier distribuant un gain différent, et les distributions de gains étant inconnues, quelle stratégie doit adopter le joueur pour maximiser ses gains ? Le problème de bandit est donc celui de la maximisation de gains lors d’une séquence de décisions sans connaissances a priori. Les problèmes de bandits (au pluriel car de nombreuses variantes du problème initial existent) sont très étudiés, et ce depuis de nombreuses années.

2. Dans le cadre des environnements d’apprentissage, cette structure correspond au modèle de l’apprenant. Ces approches se reposent notamment fortement sur les tuteurs cognitifs présentés dans le chapitre 2.

3. Une nuance est toutefois à apporter à cette affirmation : ces méthodes restent applicables dans des situations où la collecte d’information est coûteuse, à condition que la structure sous-jacente, le modèle de l’apprenant, soit bien définie.

4. En effet, les machines à sous sont parfois désignées par le terme de “bandits manchots” (en anglais sous le terme de “one-armed bandit”) pour leur capacité à dérober l’argent des joueurs, et cela sans avoir de bras (ou un seul, le levier, dans l’expression anglo-saxonne).

L'intérêt porté à ces problèmes provient de la simplicité du problème initial, qui ne repose que sur très peu d'hypothèses et peut donc s'appliquer dans de très nombreux contextes (apprentissage par renforcement, donc, mais également la médecine, et plus particulièrement les essais cliniques où le problème de la sélection du traitement à donner au patient peut se résumer à un problème de bandit).

À notre connaissance pourtant, les résultats des travaux relatifs au problème de bandit n'ont été que très peu utilisés dans le cadre d'environnement d'apprentissages adaptatifs. À notre connaissance, aucun travail n'y faisait référence en 2013 lorsque nos propres travaux ont débutés. Récemment, quelques travaux se sont dirigés dans ce sens, en particulier les travaux de l'équipe *Flowers* de l'Inria Bordeaux Sud-Ouest (Clement et al., 2013, 2014). Les nombreuses découvertes théoriques récentes quant aux problèmes de bandit, et l'intérêt croissant vers les environnements d'apprentissage adaptatifs, nous laissent cependant penser que l'analogie bandit/environnement d'apprentissage devrait se démocratiser dans le futur, d'autant plus dans le cadre de domaines mal définis où la constitution d'un modèle précis de l'apprenant est problématique. Dans la suite de cette partie, nous décrivons les deux principaux types de bandits⁵, les bandits stochastiques et les bandits contre un adversaire, et nous décrivons les possibilités d'utilisation de ces modèles au sein d'environnements d'apprentissage. Finalement dans la section 4.3 qui clôt ce chapitre, nous présentons notre approche.

4.2.3.1 Bandit stochastique

La première hypothèse à effectuer, lors de la résolution d'un problème de bandit, consiste à déterminer la nature (stochastique ou déterministe) de la distribution des gains de chaque levier/bras. L'hypothèse stochastique est la plus forte puisqu'elle suppose que la distribution des gains de chaque bras suit une loi de distribution connue. Cette hypothèse, lorsqu'elle est vérifiée, permet en contrepartie d'obtenir de meilleurs résultats que le bandit contre un adversaire, ce dernier ne pouvant écarter aucune hypothèse concernant la distribution des gains. La définition suivante définit le problème de bandit dans un cadre stochastique :

Définition (Problème de bandit stochastique). *À chaque itération $t = 1, 2, \dots$ d'une séquence de décision, l'agent choisit une action $I_t \in \{1, \dots, K\}$ et reçoit un gain $\mu_{I_t, t}$ tiré depuis une distribution ν_{I_t} (indépendante du passé).*

Le problème de bandit stochastique a été très étudié (Auer et al., 2002a; Bubeck and Cesa-Bianchi, 2012), et est aujourd'hui bien compris. Les principales techniques proposées se basent sur le principe d'optimisme dans l'incertitude (Auer et al., 2002a). Les algorithmes faisant face à ce problème

5. Pour un état de l'art très détaillé des différents types de bandits, évoquant notamment des variantes telles que les bandits linéaires que nous n'évoquons pas ici, se référer à (Bubeck and Cesa-Bianchi, 2012).

sont évalués par rapport au pseudo-regret par rapport à la meilleure stratégie constante, autrement dit la différence entre la meilleure distribution de gains $\hat{\mu}^* = \max_{i=1,\dots,K} \hat{\mu}_i$, $\hat{\mu}_i$ correspondant à la moyenne des gains reçus suite à la sélection du levier i , et les distributions des actions I_t sélectionnées à chaque itération par l'algorithme :

$$R_t = t\hat{\mu}^* - \mathbb{E} \sum_{l=1}^t \hat{\mu}_{I_l} \quad (4.1)$$

L'analogie entre bandits stochastiques et environnements d'apprentissage est loin d'être évidente. De part la complexité et le nombre de variables qui font qu'un feedback ou une activité va profiter ou pas à un apprenant, le nombre de facteurs pouvant bruite ce profit (de manière positive : motivation, concentration, ou négative : fatigue, lassitude), et la variabilité entre les besoins de différents apprenants, le gain ne peut être modélisé comme un phénomène purement stochastique. Nous ne nous attarderons donc pas sur les différentes solutions proposées dans ce cadre, et nous concentrerons sur le contexte de bandit contre un adversaire dont l'hypothèse constituante est moins forte et mieux adaptée aux gains d'apprentissage. Nous renvoyons tout lecteur intéressé vers les études de (Auer et al., 2002a; Bubeck and Cesa-Bianchi, 2012).

4.2.3.2 Bandit contre un adversaire

Le problème de bandit contre un adversaire⁶ (Auer et al., 2002b) se pose dans le cadre de distributions déterministes des gains où l'on suppose qu'un adversaire choisit, à chaque tour, le gain reçu suite à la sélection d'une action. D'aucuns pourraient se poser la question de l'intérêt d'un cadre si pessimiste du problème de bandit, et de la possibilité d'y apporter une réelle solution sachant que l'adversaire peut choisir à chaque tour de jeu le plus mauvais gain possible. Cet intérêt est pourtant certain : si l'adversaire choisit à chaque tour un gain nul indépendamment de l'action choisie, alors le problème est trivial puisque toutes les actions amènent au même résultat, et il suffit d'en choisir une aléatoirement pour maximiser le gain ; si le gain n'est pas toujours nul, c'est qu'il existe une séquence optimale, et la performance de l'algorithme peut alors être définie par rapport à cette séquence (ou si elle est inconnue, en comparant plusieurs algorithmes). Ces méthodes peuvent en outre être appliquées dans tous les contextes (dont les contextes stochastiques), l'hypothèse de départ étant moins forte. La définition suivante définit formellement le problème de bandit contre un adversaire :

Définition (Problème de bandit contre un adversaire). *À chaque itération $t = 1, 2, \dots$ d'une séquence de décision, l'agent choisit une action $I_t \in \{1, \dots, K\}$ et, simultanément à ce choix, un adversaire assigne à chaque action $i = 1, \dots, K$ un gain $\mu_i(t)$.*

6. En Anglais, adversarial bandit.

Toute stratégie est comparée par rapport à l'ensemble des stratégies constantes, c'est à dire, pour rappel, les stratégies qui choisissent à chaque itération la même action $i \in 1, \dots, K$ à chaque tour. Ainsi le *regret* par rapport à la stratégie constante k est défini selon :

$$R_t(k) = \max_{i=1, \dots, K} \sum_{l=1}^t \mu_i(l) - \sum_{l=1}^t \mu_{I_l}(l) \quad (4.2)$$

L'objectif est alors de trouver une stratégie dont le regret, pour toute séquence de gains fournie par l'adversaire, est borné par un ϵ petit.

Les stratégies les plus simples sont les stratégies semi-uniformes, c'est à dire qu'elles font une distinction binaire entre exploitation (le choix dit "glouton") et exploration (une probabilité uniforme sur un ensemble d'actions). Parmi ces stratégies, la stratégie *ϵ -first* consiste à commencer par la phase d'exploration. Pour un nombre total de $t_{\text{explore}} \in \mathbb{N}$ de tours, les actions sont choisies aléatoirement pendant $\epsilon \times t_{\text{explore}}$ tours. Pendant les $(1 - \epsilon) \times t_{\text{explore}}$ tours suivants, l'action ayant reçue en moyenne le meilleur gain est choisie. Effectuer une phase de pure exploration pendant une durée fixe rend cette stratégie sous-optimale. Conçue pour palier à ce problème, la stratégie *ϵ -decreasing* consiste quant à elle à sélectionner avec une probabilité $(1 - \epsilon_t)$ la meilleure action en moyenne, et avec une probabilité ϵ_t une action de manière aléatoire, la probabilité ϵ_t diminuant au fur et à mesure des tours : $\epsilon_t = \min(1, \frac{\epsilon_0}{t})$, la valeur de ϵ_0 étant laissée au choix de l'utilisateur. Les stratégies semi-uniformes offrent l'avantage d'être très simples et peu coûteuses. Néanmoins, lors des phases d'exploration elles sélectionnent les actions totalement aléatoirement, ce qui, intuitivement, peut être considéré comme naïf.

Afin d'éviter la sélection purement aléatoire lors des phases d'exploration, tout en cherchant le compromis entre exploitation et exploration, les méthodes probabilistes choisissent une action suivant la probabilité que cette action soit optimale. Parmi ces approches, la méthode SoftMax réside en un tirage aléatoire suivant une distribution de Gibbs. La probabilité de sélectionner une action i est alors de :

$$p_i(t) = \frac{e^{\frac{\hat{\mu}_i}{\tau}}}{\sum_{k=1}^K e^{\frac{\hat{\mu}_k}{\tau}}} \quad (4.3)$$

où $\hat{\mu}_i$ correspond à la moyenne des gains reçus après la sélection de l'action i , et $\tau \in \mathbb{R}^+$ est un paramètre de température illustrant le compromis exploration/exploitation, dont la valeur est laissée au choix de l'utilisateur. De la même manière que la stratégie *ϵ -decreasing*, une extension des méthodes SoftMax sont les méthodes SoftMax-*decreasing* qui remplacent le paramètre de température τ par une valeur décroissante $\tau_t = \frac{\tau_0}{t}$. Une variante très répandue est la méthode *Exp3*, dont l'idée principale est de diviser le gain reçu par la probabilité de l'action choisie. La probabilité de choisir l'action i devient alors :

$$p_i(t) = (1 - \gamma) \frac{w_i(t)}{\sum_{k=1}^K w_k(t)} + \frac{\gamma}{K} \quad (4.4)$$

avec $w(1) = 1$, $w_i(t+1) = w_i(t)e^{\gamma \frac{\mu_i(t)}{p_i(t)K}}$ si l'action i a été choisie au tour t ($I_t = i$), $w_i(t+1) = w_i(t)$ dans le cas contraire. $\gamma \in (0, 1]$ est un paramètre dont la valeur est laissée au choix de l'utilisateur. Si la méthode Exp3 offre des garanties théoriques plus solides (se référer à (Auer et al., 2002b) pour une preuve théorique d'une borne supérieure du regret $R_t \leq \sqrt{2TK \ln(K)}$, T étant l'horizon), les quelques applications ayant cherché à comparer les méthodes SoftMax et Exp3 n'ont pas abouti à des différences significatives (Vermorel and Mohri, 2005). Dans le cadre de séquences de décision courtes (horizon T petit), ces deux méthodes proposent également des comportements très semblables.

L'ensemble des techniques présentées jusqu'ici n'utilisent aucune connaissance a priori. Une légère variation de la méthode Exp3, nommée Exp4 et également proposée par (Auer et al., 2002b), permet de prendre en compte une expertise. Ainsi, à chaque itération t , l'agent reçoit les avis de N experts $\xi^{(1)}(t), \dots, \xi^{(N)}(t)$, où $\xi^{(j)}(t) \in [0, 1]^K$ est un vecteur dont les composantes $\xi_i^{(j)}$, $i = 1, \dots, K$ représentent les probabilités que l'expert j ait choisi l'action i . L'agent choisit alors l'action i avec la probabilité :

$$p_i = (1 - \gamma) \frac{\sum_{j=1}^N w_j(t) \xi_i^{(j)}(t)}{W(t)} + \frac{\gamma}{K} \quad (4.5)$$

où $W(t) = \sum_{j=1}^N w_j(t)$. Le poids accordé à chaque expert est mis à jour suite à la réception du gain :

$$w_j(t+1) = w_j(t) e^{\frac{\gamma \xi^{(j)} \mu(t)}{K}}, \quad (4.6)$$

où $\mu(t)$ est le vecteur dont les composantes $\mu_i(t)$, $i = 1, \dots, K$ sont égales à $\frac{\mu_i(t)}{p_i(t)}$ si $I_t = i$, 0 sinon.

Une différence majeure ici vient du fait que les poids w ne s'appliquent plus aux actions mais aux experts. Dans Exp4, il ne s'agit donc plus de déterminer l'action la plus pertinente mais l'expert le plus pertinent. C'est pourquoi Exp4 fait partie des méthodes utilisées dans le cadre des bandits contextuels, que nous évoquerons plus précisément dans le chapitre 6⁷. À noter qu'Exp4 nécessite, pour réduire au maximum le regret, un nombre d'experts très supérieur au nombre d'actions possibles ($N \gg K$), et s'applique majoritairement dans le cadre de longues séquences de décision (T grand, e.g. la publicité ciblée⁸).

Dans le cadre des environnements d'apprentissage, l'algorithme Exp4 a été utilisé par (Clement et al., 2013, 2014) pour la sélection automatique d'activités pédagogiques. Le contexte étudié est semblable au nôtre : il s'agit d'étudier

7. Chaque expert donne à chaque tour une distribution de probabilité sur chacun des choix possibles. Un expert peut donc être vu comme une politique liée à un contexte spécifique, d'où l'utilisation possible de Exp4 ou de variantes de Exp4 dans le cadre de bandits contextuels (Beygelzimer et al., 2011).

8. La notion de contexte est très importante dans le cadre de la publicité ciblée puisque le regret tel qu'il est défini dans la version simple du problème de bandit, comparant la stratégie utilisée avec la meilleure stratégie sélectionnant toujours la même action, devient alors totalement inopportun.

la sélection automatique d'actions dans un contexte dégradé, c'est à dire avec un modèle de l'apprenant réduit. La volonté est également de considérer l'individualité de l'apprenant (ses besoins propres), par opposition aux adaptations basées sur des modèles d'apprenants qui considèrent des populations d'apprenants, faisant alors l'hypothèse que tout apprenant, dans une représentation donnée, a les mêmes besoins. Dans (Clement et al., 2014), trois algorithmes sont comparés : l'algorithme expert, l'algorithme ZPDES⁹ qui s'inspire des théories relatives à la zone proximale de développement (voir chapitre 1) et de l'estimation du progrès (Oudeyer and Kaplan, 2009), et l'algorithme RiARiT¹⁰. L'algorithme expert sélectionne l'activité depuis des règles expertes fixées a priori, de façon analogue à la méthode que nous proposons dans le chapitre 2. L'algorithme ZPDES calcule le gain d'une action en évaluant la progression de l'apprenant lors des activités suivantes. N'ayant aucun modèle de l'apprenant, cet algorithme évalue la progression comme la faculté de l'apprenant à réussir un exercice, sans avantage de précision quant aux dimensions de l'apprentissage mobilisées dans l'activité. L'ensemble des activités "explorables" est réduit aux activités connexes dans la zone proximale de développement définie par un expert. L'algorithme RiARiT se base également sur une connaissance experte illustrant la zone proximale de développement à travers un arbre d'activités, mais repose en plus sur un modèle de l'apprenant qui, comme le modèle proposé dans le chapitre 2, décompose l'apprentissage en dimensions, ce qui permet d'affiner à la fois la zone proximale de développement et le calcul du gain d'une activité (cette affinement réduisant néanmoins la personnalisation de l'algorithme puisque les paramètres ajoutés sont communs à tous les apprenants). La comparaison entre ces trois algorithmes est proposée suivant deux approches : une première en utilisant des apprenants simulés, contexte dans lequel l'algorithme RiARiT offre les meilleures performances ; la seconde en conditions réelles où 400 participants (des élèves âgés de 7 à 8 ans) apprennent à compter de l'argent. Dans cette dernière configuration, l'algorithme ZPDES offre les meilleures performances. Les auteurs expliquent la supériorité de l'algorithme ZPDES par sa faculté à davantage personnaliser la sélection en comparaison avec l'algorithme RiARiT, bien que ce dernier soit plus précis.

4.2.4 Conclusions

Ces dernières années ont vu la recrudescence des travaux portant sur l'adaptation automatique des règles de sélection d'actions pédagogiques dans le cadre des environnements d'apprentissage. Des techniques initialement utilisées dans le web, telles que les systèmes de recommandation, sont désormais envisageable, notamment dans le cadre des MOOCs qui permettent la collecte de quantités importantes de données. Les systèmes de recommandation nécessitent par contre de pouvoir se baser sur une expertise, ou sur les avis des apprenants.

9. Zone of proximal development and empirical success algorithm.

10. Right Activity at the Right Time.

En outre, de nombreux défis, tels que la prise en compte du contexte et des besoins spécifiques de certains apprenants, restent ouverts.

L'apprentissage par renforcement rassemble un ensemble de méthodes qui, intuitivement, se prêtent à l'adaptation d'environnements d'apprentissage, puisque l'agent apprenant, ici l'environnement, apprend une politique de décision au fur et à mesure qu'il interagit avec l'utilisateur, ici l'apprenant. Toutefois ces méthodes demandent une quantité importante d'interactions avant de pouvoir converger vers une politique acceptable (méthodes sans modèles), ou nécessitent que la structure sous-jacente, contenant les états et les transitions entre les états, soit bien définie.

Finalement les problèmes de bandits, qui simplifient le problème posé par les modèles par renforcement pour travailler spécifiquement sur le compromis entre exploration et exploitation des actions possibles, modélisent de manière plus fidèle le problème de l'adaptation sans connaissances a priori.

4.3 Adaptation dynamique des règles de sélection

L'adaptation dynamique consiste à pouvoir réviser les règles de sélection de comportements en fonction des précédentes interactions entre l'apprenant et l'environnement. Contrairement à un ensemble fixe de règles, l'adaptation dynamique permet la prise en compte du caractère individuel de l'apprenant. Dans cette partie, nous rappelons brièvement le modèle présenté dans le chapitre 2. Nous discutons ensuite de la pertinence des différentes approches présentées dans la section 4.2. Nous présentons finalement notre approche, modélisant la sélection d'actions pédagogiques comme un problème de bandit.

4.3.1 Rappels

Rappelons brièvement le méta-modèle présenté en détails dans le chapitre 2 :

L'apprenant est représenté par le vecteur $\mathbf{s}^{(u)}(t) \in S$, S étant l'espace de représentation de l'apprenant ;

L'environnement est représenté par l'ensemble $A \in \{1, \dots, K\}$ des actions pédagogiques qu'il peut proposer à l'apprenant

La politique de sélection π est une application de S dans A qui, pour toute situation d'apprentissage $\mathbf{s}^{(u)}(t) \in S$, associe une action $I_t \in A$;

La politique de sélection optimale π^* est la politique particulière qui associe à toute situation l'action dont l'influence sur l'apprentissage sera optimal.

L'objectif de la sélection d'action est donc de trouver une politique π approchant π^* . Si l'expertise est suffisamment riche et précise, il est possible de considérer que la configuration donnée par l'expert est suffisamment proche de la configuration optimale. Cette hypothèse est celle que nous avons privilégiée

dans le chapitre 3. Dans le cadre de domaines mal définis, dont fait partie l'apprentissage de gestes, cette hypothèse ne se trouve vérifiée que dans des cas marginaux. Il est alors nécessaire de faire évoluer la configuration π au fur et à mesure des interactions entre l'apprenant et l'environnement, en espérant la faire converger au plus près de la configuration optimale π^* , d'où l'intérêt des approches par renforcement décrites dans la section 4.2.2.

4.3.2 Formalisation du problème

L'impact d'une action i peut être considéré comme étant le gain (pour reprendre le vocabulaire utilisé par les approches par renforcement) résultant du choix de cette action. L'impact d'une action i mesure la variation de l'état de l'apprenant suite à la sélection de cette action. En reprenant la définition donnée dans le chapitre 2, on a $\mu_i^{(u)}(t) = D_S(\mathbf{s}^{(u)}(t + \Delta), \mathbf{s}^{(u)}(t))$. On rappelle que $\mu_i^{(u)}(t) = 0$ lorsque $I_t \neq i$, c'est à dire lorsque l'action i n'a pas été sélectionnée à l'itération t .

Comme nous l'avons vu dans le chapitre 3, la sélection d'actions, dans le cadre de notre plateforme d'apprentissage de la calligraphie, ne peut se faire qu'à l'issue d'un exercice, le module de reconnaissance ayant besoin de l'intégralité des données relatives à cet exercice pour fournir son analyse. Nous pouvons donc formuler le processus de sélection automatique d'actions par l'algorithme 2.

Algorithme 2 : Formalisation du problème de sélection automatique d'actions

```

t = 1;
I0 = choix aléatoire parmi les éléments de A;
tant que la session n'est pas terminée faire
    proposer l'exercice avec l'action It-1;
    tant que l'exercice n'est pas terminé faire
        | enregistrer la performance de l'apprenant;
    fin
    récupérer la situation  $\mathbf{s}^{(u)}(t)$ ;
    mesurer l'impact  $\mu_{I_{t-1}}$  de l'action It-1;
    sélectionner l'action It sensée maximiser  $\mu_{i_t} \forall i \in A$ ;
    t = t + 1;
fin

```

4.3.3 Sélection d'action

L'objectif de l'adaptation est de sélectionner l'action qui maximisera le gain, en matière d'apprentissage, pour l'apprenant. On peut relier cette problématique à l'ensemble des travaux énoncés dans la partie 4.2 :

Système de recommandation Dans une approche orientée recommandation, il s'agit d'établir un filtre qui permettra de prédire, en fonction de la représentation de l'apprenant et de la représentation des actions pédagogiques, l'action qui sera la plus pertinente.

Apprentissage par renforcement Dans une approche par renforcement, il s'agit de tirer profit des interactions passées pour adapter les règles de sélection.

A priori, ces deux points de vue sont compatibles avec le modèle établi dans le chapitre 2. Toutefois, les approches orientées recommandation, telles que le filtrage par contenu ou le filtrage collaboratif, sont fondées sur l'hypothèse que l'ensemble des utilisateurs appartenant au même profil auront le même besoin. Ces approches sont donc fondées sur l'hypothèse que la représentation de l'utilisateur est suffisamment fine pour regrouper les groupes d'apprenants ayant les mêmes besoins. Il est également nécessaire d'acquérir une quantité importante de données pour chaque type de profil afin de pouvoir recommander une action. Les approches par renforcement sont en ce sens plus souples, puisqu'elles permettent de partir d'une situation sans connaissance a priori, et devront trouver le compromis entre exploration et exploitation qui leur permettra, au fil du temps, de raffiner la politique de sélection. Néanmoins, les approches usuelles nécessitent de pouvoir collecter des quantités suffisantes de données (approches sans modèle) ou de pouvoir se baser sur un modèle bien défini (approches avec modèle).

Nous décidons donc de modéliser la sélection comme un problème de bandit, au sein duquel il s'agit de maximiser le gain au cours d'une séquence de décisions (ici une succession d'exercices à l'issue desquels il est nécessaire de sélectionner le feedback à proposer au tour suivant). Nous reprenons les approches Soft-Max, qui ont montré leur efficacité sur de courtes séquences de décision, et qui permettent d'éliciter le compromis exploration/exploitation par un paramètre, noté τ . Dans cette approche, la probabilité de sélectionner une action $i \in A$ tient compte de l'historique des gains précédemment obtenus. En décomposant le gain moyen dans l'équation (4.3) par

$$\hat{\mu}_i^{(u)}(t) = \frac{\sum_{l=1}^t \mu_i^{(u)}(l)}{t} \quad (4.7)$$

on obtient la relation suivante permettant de définir la probabilité de sélectionner l'action i à l'itération t :

$$p_{i,t} \propto e^{\frac{\sum_{l=1}^t \mu_i^{(u)}(l-1)}{t_i^\tau}}, \quad (4.8)$$

où t_i correspond au nombre d'exercices ayant vu l'action i être sélectionnée. L'égalité dans l'équation précédente est obtenu en normalisant par la somme des probabilités $p_{i,t+1}$ associées aux actions $i = 1, \dots, K$.

Le gain d'une action est défini dans le chapitre 2 comme la différence, dans le modèle de l'apprenant, entre les représentations de ce dernier avant et après avoir proposé l'action. Nous nous positionnons sur des périodes d'apprentissage

courtes, et nous utilisons un calcul de gain dit *immédiat*, c'est à dire la différence de représentation entre deux exercices consécutifs :

$$\mu_i^{(u)}(t) = \delta_{[I_t=i]} D_S(\mathbf{s}^{(u)}(t+1) - \mathbf{s}^{(u)}(t)) \quad (4.9)$$

avec, pour rappel, $\delta_{[I_t=i]}$ valant 1 si l'action i est sélectionnée à l'itération t , 0 sinon. À noter que de récents travaux portent sur des comparaisons de différents calculs de gains, notamment (Shen and Chi, 2016) qui comparent le gain immédiat (tel que nous l'utilisons) au calcul de gain *retardé*, qui considère un écart $\Delta \gg 1$ entre deux situations. Il pourrait être intéressant de considérer les résultats de ces travaux lors de futures expérimentations portant sur des périodes d'apprentissage plus longues.

Afin de différencier le gain d'une action sur chacune des dimensions du geste, nous représentons désormais le vecteur de gain $\boldsymbol{\mu}$ tel que

$$\boldsymbol{\mu}_i^{(u)}(t) = \{\mu_{i,0}^{(u)}(t), \dots, \mu_{i,m}^{(u)}(t)\}, \quad (4.10)$$

avec

$$\mu_{i,j}^{(u)}(t) = \delta_{[I_t=i]} D_{S_j}(s_j^{(u)}(t+1), s_j^{(u)}(t)) \quad (4.11)$$

représentant le gain immédiat relatif à la $j^{\text{ème}}$ dimension du modèle S .

Contrairement au problème de bandit classique, dans lequel aucun contexte n'est associé, il existe dans le problème de la sélection d'actions pédagogiques une relation entre le gain obtenu par une action et la valeur de l'information $s^{(u)}(t)$. On peut en effet admettre que le gain reçu suite à la sélection d'un feedback pour un apprenant novice sera différent que le gain reçu lorsque ce même feedback sera proposé au même apprenant lorsque ce dernier aura acquis un niveau confirmé. L'extension du problème de bandit permettant la prise en compte de contexte a constitué une direction importante de recherche ces dernières années (Langford and Zhang, 2008; Li et al., 2010; Slivkins, 2014). La majorité des travaux associés aux bandits contextuels considèrent des ensembles discrets de contextes dont les éléments sont indépendants. Ici, les valeurs de s_j sont continues sur $[0, 1]$, et il est possible de mesurer la similarité entre deux contextes. Il est donc possible, lors du processus de sélection, de pondérer les expériences antérieures par leur pertinence eu égard à la situation courante. On définit alors le gain pondéré $\omega_{i,j}^{(u)}(l)$ à l'itération t tel que

$$\omega_{i,j}^{(u)}(l) = \mu_{i,j}^{(u)}(l) \times \text{sim}(\mathbf{s}_j^{(u)}(t), \mathbf{s}_j^{(u)}(l)), \quad (4.12)$$

où $\text{sim}(\mathbf{s}_j^{(u)}(t), \mathbf{s}_j^{(u)}(l)) = 1 - D_{S_j}(\mathbf{s}_j^{(u)}(t), \mathbf{s}_j^{(u)}(l))$. Incorporée dans l'équation (4.8), la probabilité de sélectionner l'action i par rapport à ses gains relatifs à la dimension S_j est définie par :

$$p_{i,j,t} \propto e^{\frac{1}{\tau_i \cdot \tau} \sum_{l=1}^t \omega_{i,j}^{(u)}(l-1)} \quad (4.13)$$

où l'égalité dans la précédente équation est obtenue en normalisant par le terme $Z(S)$ définie par :

$$Z(S) = \sum_{k=1}^K e^{\frac{1}{\tau_k \cdot \tau} \sum_{l=1}^t \omega_{k,j}^{(u)}(l-1)} \quad (4.14)$$

Plusieurs stratégies peuvent alors être pour définir la probabilité $p_{i,t}$ suivant laquelle tirer l'action i :

1. Privilégier les actions très pertinentes sur une dimension :

$$p_{i,t} \propto \max_{j=1,\dots,m} p_{i,j,t}$$

2. Privilégier les actions pertinentes sur l'ensemble des dimensions :

$$p_{i,t} \propto \sum_{j=1}^m p_{i,j,t}$$

où les égalités sont obtenues en normalisant par la somme de probabilités $p_{i,t}$ pour toute action $i \in A$.

Il serait intéressant de comparer ces deux stratégies, et leurs effets sur l'apprentissage. Nous privilégierons ici la seconde stratégie qui, contrairement à la première, prend en compte l'ensemble des dimensions du modèle de l'apprenant. La pertinence de la première stratégie reste donc ouverte et peut constituer une perspective d'étude.

4.4 Conclusions

Dans ce chapitre, nous présentons les principaux travaux relatifs à l'adaptation automatique des environnements d'apprentissage. Les approches par renforcement et par recommandation sont les principaux modèles utilisés. Néanmoins ces modèles ont besoin d'être entraînés par de grandes quantités de données, ou se basent sur des modèles complexes représentant les transitions entre différents états d'apprentissage. Nous choisissons donc de modéliser le problème de sélection d'actions comme un problème de bandit, qui s'attache à trouver un compromis entre exploitation des actions connues et exploration de nouvelles actions pour maximiser le gain cumulé lors d'une séquence de décisions. Parmi les méthodes proposées pour résoudre le problème de bandit, nous nous positionnons dans le cadre des méthodes SoftMax, qui permettent de sélectionner une action en fonction de probabilité qu'elle obtienne un bon gain. Un paramètre de température τ est inclus dans le calcul des probabilités et illustre le compromis entre exploration et exploitation. Nous adaptons les méthodes SoftMax à la nature multidimensionnelle de notre gain (i.e. l'impact d'un feedback sur chacune des dimensions de la représentation de l'apprenant) en définissant pour chaque action un vecteur de probabilités, où chaque élément illustre la probabilité que cette action obtienne un bon gain sur une dimension particulière. Enfin, nous pondérons les gains antérieurs par leur pertinence par

rapport à la situation courante, cette pertinence étant évaluée par la similarité entre la situation courante et la situation lors de laquelle le gain a été obtenu.

Dans le prochain chapitre, nous présenterons une expérimentation comparant les deux configurations d'environnement présentés jusqu'alors : l'adaptation basée sur une expertise, et l'adaptation automatique. Nous souhaitons ainsi mesurer un éventuel apport d'une adaptation dynamique des règles de sélection d'actions, en terme d'apprentissage comme en terme d'acceptation de l'environnement par les apprenants.

Chapitre 5

Adaptation automatique : Expérimentation

5.1 Introduction

Une première expérimentation, présentée dans le chapitre 3, a permis de mesurer l'apprentissage lors de courtes phases d'entraînement à l'aide d'un environnement proposant des retours adaptés à l'activité de l'apprenant. Cette adaptation de l'environnement se base sur un ensemble de règles délivrées par des experts. Plusieurs conclusions ont été tirées de cette première expérimentation : la diversité des retours proposés a permis aux apprenants de reproduire leurs performances suite à la phase d'entraînement, ainsi aucun syndrome de dépendance à l'environnement d'apprentissage n'a été détecté ; les participants n'étaient pas tous satisfaits de l'adaptation de l'environnement, certains participants indiquant que l'environnement comprenait et proposait des retours adaptés à leur activité, quand d'autres ne comprenaient pas pourquoi des aides qui leur semblaient inutiles continuaient à leur être proposées. Il apparaissait donc que les règles fournies par les experts ne pouvaient pas retranscrire le caractère individuel des besoins des apprenants. Enfin, le domaine d'apprentissage et les feedbacks implémentés avaient été pensés de telle façon qu'un ensemble de règles puisse être défini par les experts. La possibilité de retranscrire sous forme de règles l'influence potentielle de feedbacks plus complexes, dans le cadre de domaines d'apprentissage plus larges, apparaissait alors comme une limite de l'approche.

Pour toutes ces raisons, nous avons proposé dans le chapitre 4 une extension de notre modèle basée sur une approche de bandit permettant à l'environnement d'adapter automatiquement, et dynamiquement, les règles de décision au fil des interactions avec l'apprenant. Pour évaluer l'apport d'une telle approche, nous comparons deux configurations d'environnement : un environnement s'adaptant grâce à une base de règles expertes, et un environnement adaptant automatiquement et dynamiquement les règles de décision au fil des interactions, et ce sans connaissance a priori. Le positionnement sur une adap-

tation automatique des règles de sélection sans connaissance a priori nous permet en effet d'éliminer l'influence d'une expertise initiale sur les résultats. Il serait en effet complexe de distinguer, d'autant plus sur une période courte d'apprentissage, les bénéfices et limites liés à l'expertise initiale des bénéfices et limites liés à l'adaptation des règles de sélection.

5.2 Environnement CalliSmart (2)

L'environnement d'apprentissage étudié ici reprend l'environnement CalliSmart présenté dans le chapitre 3. Plusieurs modifications sont cependant apportées à la première version de l'environnement, afin que l'offre pédagogique de l'environnement (i.e. les feedbacks qu'il peut proposer) soit plus diversifiée. Nous présentons ici ces modifications en suivant le même schéma que celui proposé dans le chapitre 3 : nous présentons les dimensions d'erreurs qui constituent la caractérisation de l'activité de l'apprenant, puis nous présentons l'ensemble des feedbacks que l'environnement peut proposer. Ensuite, nous décrivons la manière dont l'environnement, à l'aide de techniques d'apprentissage automatique, évalue la probabilité que chaque erreur soit commise. Cette dernière partie ayant déjà été largement décrite dans le chapitre 3, nous ne présentons ici que les modifications apportées depuis la première version de CalliSmart. Enfin, la section 5.3 présente une expérimentation permettant de comparer l'adaptation basée sur l'expertise et l'adaptation automatique.

5.2.1 Caractérisation du geste

La caractérisation du geste suit ici la même approche que celle présentée dans la première version de l'environnement, l'activité étant caractérisée depuis des scores illustrant la probabilité que diverses erreurs, établies par des experts comme des erreurs fondamentales, aient été présentes dans la performance de l'apprenant. Nous nous concentrons ici sur deux types d'erreurs : la fluidité du geste, et l'inclinaison des éléments tracés. L'intérêt que nous portons à ces deux composantes particulières, provient du fait qu'elles s'inscrivent toutes deux dans une approche morphocinétique de l'activité gestuelle, contrairement à la première version de CalliSmart qui s'intéressait au geste depuis des aspects à la fois morphocinétiques et topocinétiques. Nous nous concentrons donc ici sur un aspect plus précis de l'activité. En outre, les deux dimensions étudiées sont constituantes du tracé mais sont, pour un pratiquant novice, contradictoires : il est aisé de tracer un élément parfaitement vertical, comme il est aisé de tracer d'une manière parfaitement fluide ; pourtant il est complexe d'acquérir un contrôle suffisant pour tracer un élément parfaitement vertical de manière fluide. Et c'est donc sur ce contrôle moteur que nous concentrons ici notre apprentissage. L'une des conséquences de ce choix de caractérisation concerne la granularité avec laquelle un exercice est considéré : alors que dans la première version de CalliSmart, nous considérions l'activité à l'échelle de l'exercice complet, afin de pouvoir apprécier la dimension de régularité (topo-

cinétique), nous pouvons ici travailler sur une échelle plus fine en considérant chaque élément du tracé individuellement.

5.2.2 Ensemble d’actions pédagogiques

Comme dans la première version de CalliSmart, les actions pédagogiques proposées par l’environnement sont réparties en deux groupes. Un feedback CR (connaissance du résultat) est proposé à la fin de chaque exercice, et permet à l’apprenant de pouvoir évaluer sa performance, et l’évolution de celle-ci. Les feedbacks supplémentaires sont des informations proposées en temps réel lors de la performance de l’apprenant, et le choix du type de feedback proposé lors d’un exercice est fait à l’issue de l’exercice précédent. Un seul type de feedback temps-réel est sélectionné à chaque exercice, afin de ne pas surcharger l’apprenant d’informations, les informations fournies en temps réel pendant une activité gestuelle impliquant déjà une charge cognitive importante (Danna and Velay, 2015).

5.2.2.1 Connaissance du résultat

Contrairement au feedback CR proposé précédemment (chapitre 3, figure 3.4), qui permettait d’informer l’apprenant sur la qualité globale de son exercice, il s’agit ici de pouvoir informer l’apprenant sur la qualité de chacun des éléments tracés lors de l’exercice. Pour cela, nous proposons à la fin de l’exercice deux pastilles de couleurs affichées sous chacun des éléments. La pastille marquée d’un “S” indique la qualité de l’activité en terme de fluidité, quand la pastille marquée d’un “T” indique la qualité du tracé en terme d’inclinaison. Les pastilles sont colorées suivant un gradient virant du rouge au vert suivant la valeur de la probabilité que l’erreur concernée ait été commise : une pastille verte indique une très faible probabilité d’erreur, alors qu’une pastille rouge illustre au contraire une forte probabilité d’erreur.

5.2.2.2 Feedbacks supplémentaires

Un ensemble de quatre types de feedback a été implémenté, en suivant les recommandations d’experts. Deux de ces feedbacks s’inscrivent clairement dans les dimensions morphocinétiques étudiées : les feedbacks illustrés par les figures 5.2 (que nous nommerons feedback *vélocité*) et 5.3 (que nous nommerons feedback *pression*) permettent ainsi d’informer l’apprenant sur les discontinuités respectivement liées à la dynamique et à l’appui de leur geste. Le feedback *vélocité* détecte les discontinuités dans la dynamique du geste en analysant la variation de vitesse sur une fenêtre glissante, et affiche un trait horizontal de couleur bleue lors d’une décélération brutale, et un trait horizontal rouge lors d’une forte accélération. Le feedback *pression* joue quant à lui sur la largeur du tracé, à la manière de l’encre qui bave lorsque la pression appliquée par l’outil (e.g. plume biseautée) sur le plan est trop importante. La couleur du tracé donne également une indication de la pression appliquée, depuis une pression

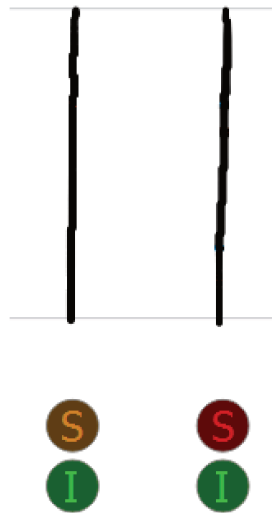


FIGURE 5.1 – Les pastilles de couleurs informent l’apprenant sur la qualité de chaque tracé

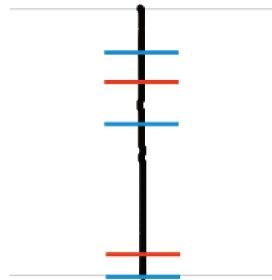


FIGURE 5.2 – Le feedback “vitesse” indique des discontinuités dans la dynamique du tracé : en rouge les accélérations, en bleu les ralentissements

trop faible illustrée par une couleur de trait bleue, vers une pression trop forte illustrée par un tracé rouge.

Les deux autres feedbacks, même s’ils ne fournissent pas d’explications aussi explicites, sont considérés par les experts comme utiles pour travailler les dimensions que nous considérons. Le feedback *peau d’oignon*, illustré par la figure 5.4, affiche l’image de l’élément précédemment tracé à la même position que le tracé en cours. Ce type de feedback, qui peut être utile puisqu’il permet de comparer deux performances successives, peut également troubler la performance en poussant à une tâche de mimétisme. Malgré cela, le feedback “peau d’oignon” reste une aide incontournable des cursus d’apprentissage de calligraphie.

Finalement le feedback *cible* (figure 5.5) consiste à afficher un point vert

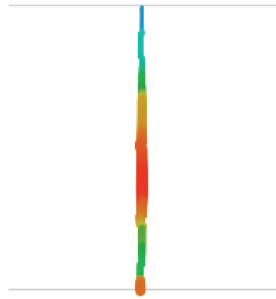


FIGURE 5.3 – Le feedback “pression” indique par la largeur du trait ainsi qu’un gradient de couleur la pression appliquée par l’outil sur le plan



FIGURE 5.4 – Le feedback “peau d'oignon” propose en grisé l’élément précédemment inscrit



FIGURE 5.5 – Le feedback cible permet de focaliser vers la suite du geste

agissant comme une cible, et situé quelques centimètres au delà de la position du tracé. Une telle cible permet de défocaliser l’attention de l’apprenant novice, ayant tendance à se focaliser, à tort, sur la position de l’outil plutôt que de se projeter sur la suite de son geste.

TABLE 5.1 – Caractéristiques calculées depuis les données extraites de l’activité de l’apprenant

Erreur d’inclinaison	Inclinaison cumulée Inclinaison globale Ratio entre taille de l’élément et écart vertical entre les portées
Erreur de fluidité	Nombre de pics de vitesse Vitesse moyenne Pression moyenne Écart-type des mesures de vitesse Écart-type des mesures de pression

5.2.3 Reconnaissance de gestes

Nous avons particulièrement insisté, dans le chapitre 3, sur les différentes méthodes appartenant au domaine de l’apprentissage supervisé, et permettant de prédire, depuis une observation, un ensemble d’étiquettes. Nous avons alors privilégié la technique du *label powerset*, qui consiste à transformer le problème multilabel en un problème multiclassé. Utilisant cette technique, et ayant émis le besoin de pouvoir interpréter les probabilités d’occurrence de chacune des erreurs, nous avons comparé les modèles probabilistes basés sur les modèles bayésiens naïfs et les régressions logistiques. N’ayant accès qu’à une quantité faible de données d’entraînement, nous avons alors privilégié l’utilisation des modèles bayésiens naïfs qui offraient des performances plus robustes.

La granularité plus fine avec laquelle nous considérons l’activité gestuelle, ainsi que la modification des dimensions étudiées, nécessitent une modification du calcul du vecteur de représentation, ce dernier représentant désormais un élément de tracé plutôt qu’un exercice complet. Le tableau 5.1 décrit les caractéristiques constituantes du vecteur de représentation d’un élément de tracé.

Puisqu’il existe désormais un vecteur de descripteurs (et donc une observation) par élément de tracé, et non plus par exercice, il devient possible d’entraîner l’algorithme de reconnaissance avec une quantité plus importante de données. Chaque exercice comprenant dix à quinze éléments, et ces éléments étant également étiquetés par les experts, nous multiplions d’autant la quantité de données d’entraînement (obtenant alors un jeu contenant environ 1300 observations étiquetées). Une comparaison entre les modèles bayésiens naïfs et les régressions logistiques sur ce nouveau jeu de données est illustrée par le tableau 5.2, lors d’une expérimentation en validation croisée à dix sous-ensembles. Le modèle basé sur les régressions logistiques obtenant ici en moyenne de meilleures performances ainsi qu’une plus faible variance, nous optons pour ce modèle au détriment du modèle bayésien naïf que nous utilisions précédemment.

TABLE 5.2 – Comparaison des aires sous la courbe ROC entre les modèles bayésiens naïfs (génératifs) et régressions logistiques (discriminants)

	Naïve Bayes	Régressions Logistiques
Inclinaison	0.8953 ± 0.1267	0.9263 ± 0.0471
Fluidité	0.8586 ± 0.1394	0.8701 ± 0.0937

5.3 Expérimentation

Après avoir mesuré l’impact sur l’apprentissage d’une adaptation basée sur des règles expertes lors de la précédente expérimentation, nous comparons ici une configuration semblable avec l’environnement présenté au chapitre précédent qui, sans connaissance a priori, adapte les règles de sélection. Outre l’évaluation de la progression des participants sur ces deux configurations d’environnement, nous souhaitons également observer l’adoption de ces environnements par les participants. Cette partie s’organise comme suit : nous commençons par présenter le protocole expérimental, où nous précisons notamment les règles fournies par les experts. Puis nous présentons et discutons les résultats.

5.3.1 Protocole

44 participants sont aléatoirement partagés en deux groupes : le groupe *expert*, et le groupe *automatique*. Le groupe expert travaille au sein d’un environnement qui sélectionne les actions pédagogiques depuis les règles fournies par les experts. Les règles sont les suivantes :

- le feedback “vélocité” correspond à une situation où la gestuelle est très saccadée ;
- le feedback “pression” correspond à une situation où la gestuelle est faiblement saccadée ;
- le feedback “peau d’oignon” correspond à une situation où la gestuelle est fluide ou faiblement saccadée, et où les éléments sont inclinés ou très inclinés ;
- le feedback “cible” correspond à une situation où la gestuelle est faiblement saccadée ou saccadée, et où les éléments sont faiblement inclinés ou inclinés.

Le groupe automatique travaille quant à lui au sein d’un environnement qui sélectionne les actions pédagogiques depuis les règles apprises au fur et à mesure des interactions, suivant le modèle décrit dans le chapitre 4 (section 4.3).

Tout participant débute par une période de familiarisation à l’utilisation de l’environnement, où il peut dessiner librement. La nature de l’exercice, et les dimensions évaluées, lui sont ensuite présentées. Une fois les participants prêts et l’exercice compris, un premier exercice de *pré-test* est proposé. Cet exercice, comme les exercices suivants, consiste à produire dix à quinze éléments. Aucun feedback n’est proposé lors de ce premier exercice, qui permet d’évaluer

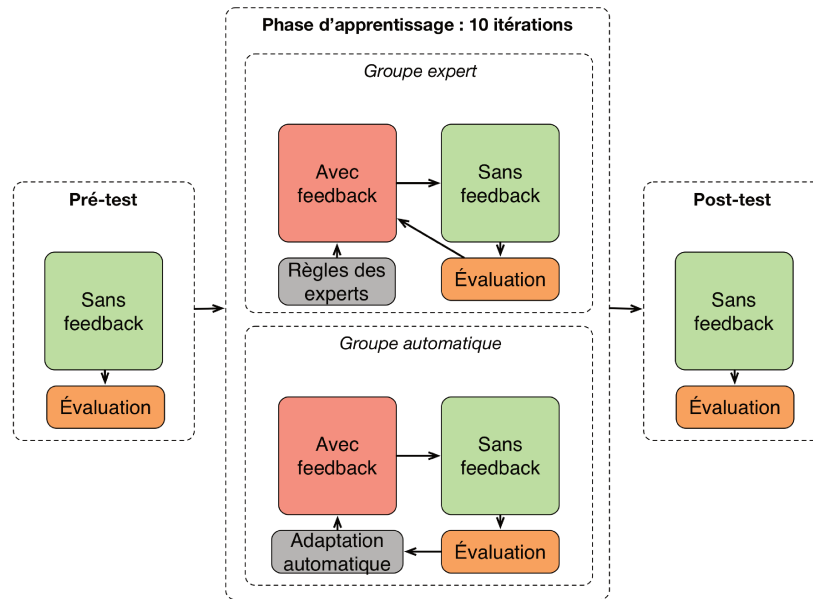


FIGURE 5.6 – Illustration du scénario expérimental pour les deux groupes considérés

le niveau initial des apprenants. Suite à ce pré-test, la phase d'apprentissage comprend dix itérations.

Chaque itération comprend deux exercices : un premier exercice où l'environnement propose l'un des feedbacks possibles, et qui permet au participant de travailler son geste, puis un second exercice sans feedback qui permet d'évaluer la performance. Cette méthode permet de réduire le biais qui pourrait être induit par la mesure de la performance sur des exercices au sein desquels des feedbacks sont proposés. L'évaluation d'un exercice sur une dimension est donnée par la moyenne des erreurs relatives à cette dimension sur chacun des éléments tracés. La performance globale correspond à la distance euclidienne entre la représentation de l'activité de l'apprenant et le vecteur représentant l'expertise.

Suite à ces dix itérations d'entraînement, un dernier exercice, le *post-test*, permet d'évaluer le niveau final des apprenants en conditions réelles (sans feedback). La figure 5.6 illustre ce scénario expérimental. Une fois ce scénario terminé, un questionnaire est proposé. Ce questionnaire est composé de deux questions : une première relative à l'impression du participant quant à sa progression, et une seconde relative à son impression vis-à-vis de la pertinence des feedbacks proposés par l'environnement.

	Groupe expert	Groupe automatique	p
	Moyenne (Écart-type)	Moyenne (Écart-type)	
Pre-test	0.34(0.12)	0.30(0.13)	0.57
Itération 2	0.30(0.17)	0.29(0.12)	0.73
Itération 4	0.28(0.19)	0.28(0.17)	0.92
Itération 6	0.29(0.21)	0.24(0.18)	0.38
Itération 8	0.30(0.22)	0.23(0.17)	0.19
Itération 10	0.30(0.25)	0.19(0.12)	0.07
Post-test	0.29(0.26)	0.20(0.13)	0.04*
Progression	$p = 0.53$	$p = \mathbf{0.03^*}$	

TABLE 5.3 – Comparaison entre les performances des deux groupes expérimentaux à chaque itération (tests t de Student), et entre les progressions globales (tests t de Student appariés)

5.3.2 Résultats et discussion

Les performances des deux groupes de participants au fil des itérations sont présentées par le tableau 5.3, à travers les moyennes et écarts-types des échantillons mesurés.

Les p -valeurs à chaque itération résultent de tests t de Student. Les p -valeurs relatives à la progression de chaque groupe entre le pré-test et le post-test résultent de tests t de Student appariés. Aucune différence significative n'est à noter entre les deux groupes lors du pré-test, la p -valeur étant égale à 0.57. Le groupe automatique, qui pour rappel travaille sur l'environnement avec adaptation automatique, surpasse de manière significative le groupe expert, qui travaille sur l'environnement s'adaptant depuis les règles expertes, lors du post-test ; cette différence étant illustrée par une p -valeur inférieure à 0.05 ($p = 0.04$).

La comparaison entre les performances en pré-test et en post-test de chacun des deux groupes fait apparaître une progression significative du groupe automatique ($p = 0.03$), alors que le groupe expert ne progresse pas significativement ($p = 0.53$). Un regard plus attentif aux différentes étapes de l'expérimentation nous indique que les performances entre les deux groupes sont très similaires sur les quatre premières itérations d'entraînement. À partir de la quatrième itération cependant, il apparaît que les participants du groupe automatique progressent de façon plus homogène que les participants du groupe expert. Les écarts-types et moyennes calculés depuis les échantillons mesurés dans le groupe expert ($\sigma = 0.21$ à 0.26 lors des itérations 6 à 10, pour une moyenne constante autour de 0.30) semble montrer qu'une partie seulement des participants progresse, pendant que l'autre, au contraire, voit ses performances se dégrader au fil des itérations. Ce résultat est cohérent avec les conclusions tirées de notre première expérimentation, et montre que les décisions prises depuis l'expertise ne sont pertinentes que pour une partie seulement des participants. Une explication possible de la dégradation des performances des autres participants, pour lesquels les choix de l'environnement ne s'avèrent pas per-

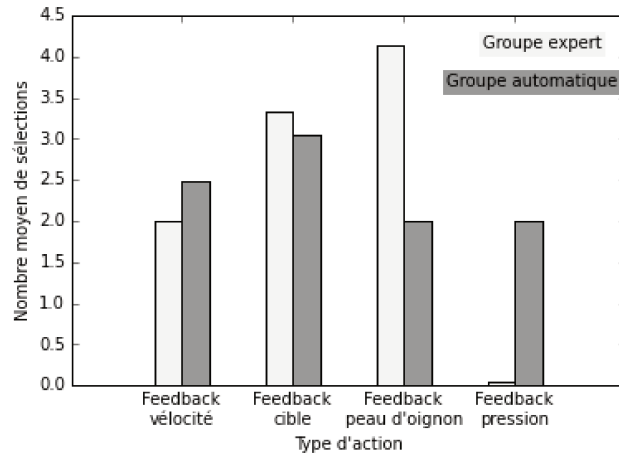


FIGURE 5.7 – Diagramme barre illustrant le nombre moyen de sélection des différents feedbacks par apprenant

tinents, peut être qu'ils se lassent et se désengagent de la tâche après que ne leur soient proposés que des retours qu'ils ne jugent pas pertinents, ou qu'ils ne comprennent pas. Les participants du groupe automatique, qui se voient proposer les feedbacks ayant eu l'influence la plus positive lors des précédentes itérations, conservent quant à eux une concentration et une motivation suffisante pour progresser.

La figure 5.7 illustre la fréquence de sélection de chacun des feedbacks par phase d'apprentissage. Il est tout d'abord remarquable que le groupe automatique se voit proposer une diversité plus importante de feedbacks, ce qui signifie qu'outre les sélections causées par les périodes d'exploration, chacun des quatre feedbacks s'est avéré être le plus performant pour certains apprenants. Au contraire, le feedback pression n'est jamais sélectionné dans l'environnement basé sur les règles expertes, aucun apprenant ne s'étant retrouvé dans une situation permettant cette sélection. Ces résultats illustrent, outre l'impossibilité de représenter le caractère individuel des besoins des apprenants depuis des règles expertes, la grande difficulté liée à la définition de règles prédisant l'influence de différents feedbacks dans le cadre de l'apprentissage de geste.

Les figures 5.8 et 5.9 présentent les résultats du questionnaire proposé suite au post-test. Les participants du groupe automatique ont un avis plus favorable quant à leur progression (6.3 en moyenne contre 5.0 pour les participants du groupe expert, figure 5.8). En outre, l'avantage significatif de groupe automatique en matière d'apprentissage se traduit également dans l'impression qu'ont gardée les participants de leurs performances ($p = 0.021$, figure 5.8). Les participants du groupe automatique ont également un avis plus favorable vis-à-vis de la pertinence des choix de l'environnement (5.6 en moyenne contre

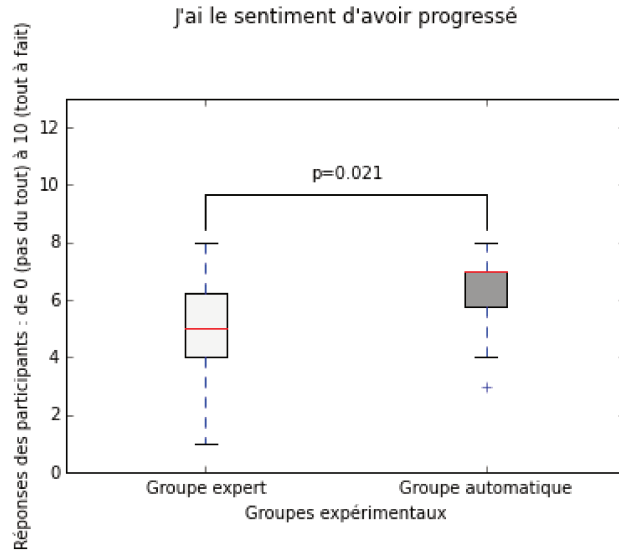


FIGURE 5.8 – Réponses des participants face à la proposition : “J’ai le sentiment d’avoir progressé”

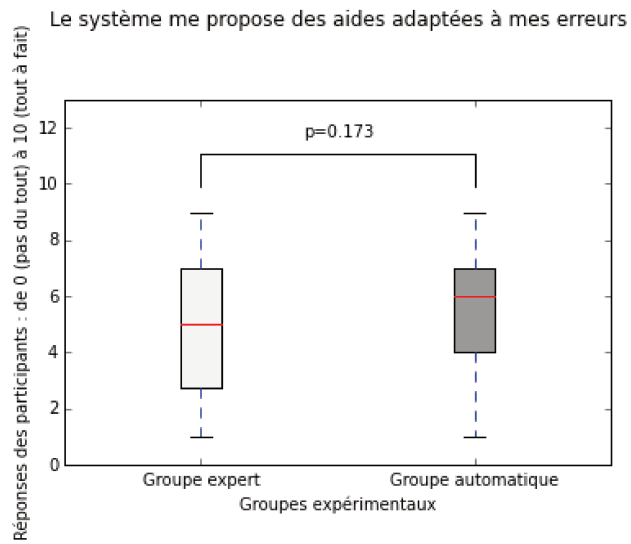


FIGURE 5.9 – Réponses des participants face à la proposition : “Le système me propose des aides adaptées à mes erreurs”

4.6 pour les participants du groupe expert), bien que la différence ne soit pas significative ($p = 0.173$).

5.4 Conclusions

Suite aux deux expérimentations, respectivement présentées dans les chapitres 3 et 5, il apparaît que proposer une diversité de feedbacks permet aux apprenants, dans le cadre de périodes courtes d'apprentissage (environ trente minutes), de retranscrire les performances établies lors de la phase d'entraînement en conditions réelles, et ainsi éviter le syndrome de dépendance à l'environnement souvent noté dans ce contexte.

Le positionnement que nous avons adopté en ce qui concerne la modélisation de l'activité de l'apprenant, que nous avons caractérisé depuis un ensemble de dimensions d'erreurs, nous permet de cibler la situation d'apprentissage, et ainsi de sélectionner l'action apparaissant comme la plus pertinente. Concernant cette notion de pertinence, il apparaît que la capacité de l'environnement à adapter les règles lui permettant de sélectionner une action pédagogique (dans notre cas un type de feedback visuel parmi ceux disponibles) soit un caractère essentiel à l'adoption de l'environnement par les apprenants, et à l'efficacité de ce dernier sur l'apprentissage.

Dans l'expérience présentée ici, l'environnement utilisé par le groupe automatique n'utilise aucune connaissance a priori, et nous testons ainsi l'adaptation dynamique dans sa configuration la plus naïve afin de ne pas biaiser l'étude par l'influence d'une expertise initiale. Cette configuration naïve montre ici de bons résultats. Le nombre d'actions possibles étant réduit, il est en effet possible d'explorer l'ensemble de ces actions dans un temps court. Il est toutefois évident que dans un cadre plus complexe, impliquant notamment de grands ensembles d'actions possibles et de longues périodes d'apprentissage, cette approche peut s'avérer sous-optimale. Une extension possible serait alors d'initialiser la politique de sélection, non pas aléatoirement mais depuis les règles expertes, ou de restreindre l'ensemble d'actions pédagogiques explorables, dans une approche similaire à celle proposée par (Clement et al., 2014) mobilisant le principe de zone proximale de développement.

Nous dressons dans le prochain chapitre les conclusions de nos travaux, et présentons les principales perspectives de travaux futurs.

Chapitre 6

Conclusions et perspectives

En conclusion de ce mémoire, nous proposons un résumé des travaux et des contributions au regard des problématiques générales présentées en introduction. Nous dressons ensuite les limites des approches proposées, et proposons des pistes de recherches futures.

6.1 Résumé et contributions

Nous avons présenté dans ce mémoire une approche permettant la création d'un environnement d'apprentissage de gestes. Les principaux défis inhérents à la création d'un tel environnement sont :

- la caractérisation du geste : comment le geste est-il considéré ? quelles sont les dimensions/aspects du geste qui puissent permettre son évaluation ?
- l'interprétation du geste : comment obtenir, depuis un geste, des valeurs quantitatives sur chacune des dimensions le caractérisant qui puissent le situer par rapport à l'objectif recherché ?
- la représentation du geste et de l'environnement : comment représenter le niveau de l'apprenant sur chacune des dimensions considérées ? Comment faire le lien entre la représentation de l'apprenant et l'ensemble des feedbacks qu'un environnement peut proposer ?
- l'adaptation de l'environnement : quelles sont les heuristiques pouvant permettre à un environnement de sélectionner, en fonction d'une situation d'apprentissage, le feedback le plus pertinent ?

6.1.1 Caractérisation

Nous proposons une caractérisation de l'activité gestuelle depuis des dimensions d'erreur. Dans les domaines relatifs à la résolution de problèmes, les modélisations du domaine et des perturbations représentent des ensembles disjoints (d'où l'intérêt, dans ces domaines, de combiner les deux modèles). Le domaine représente l'ensemble des solutions acceptables, et les perturbations

des erreurs types. Dans le cadre du geste, une représentation des perturbations permet également de représenter le domaine. Le lien entre ces deux notions est en effet plus ambigu : si l'on considère un ensemble complet de dimensions d'erreur, un geste expert est caractérisé par une absence d'erreur. Le geste peut donc être caractérisé par un score sur chacune des dimensions d'erreurs. On établit qu'un score de 0 représente l'expertise (pas d'erreur), et qu'un score de 1 représente le novice (erreur grossière).

Dans le cadre de l'apprentissage de la calligraphie, les deux activités que nous avons étudiées se sont focalisées sur les erreurs relatives à la morphocinétique (inclinaison et taille du tracé), topocinétique (régularité entre les éléments du tracé) et dynamique (fluidité) du geste.

6.1.2 Interprétation

La problématique de l'interprétation consiste à calculer les scores définis précédemment. S'inspirant des travaux de (Brusilovsky and Millán, 2007), nous proposons d'interpréter ce score comme la probabilité que l'apprenant ait commis une erreur grossière sur la dimension considérée. Les erreurs que nous considérons pouvant être caractérisées par des caractéristiques globales au tracé (caractéristiques présentées dans les chapitres 3 et 5), nous avons opté pour une représentation vectorielle du geste. Une représentation vectorielle du geste consiste à représenter chaque observation par un unique vecteur de caractéristiques qui décrit à lui seul cette observation toute entière.

Une phase de collecte de données menée à l'ESAD a permis la constitution d'un jeu d'exemples d'entraînement. Depuis ce jeu de données, nous avons comparé plusieurs méthodes de la littérature associée au domaine de la classification multilabel, qui consiste à étiqueter chaque observation (i.e. un geste) en l'associant à un ensemble de classes (i.e. des erreurs). Nous avons finalement validé la mobilisation de la méthode de *label powerset* utilisant des classifieurs bayésiens naïfs (dans la première expérience, chapitre 3) puis utilisant des régressions logistiques (dans la seconde expérience, chapitre 4).

6.1.3 Représentation

Nous proposons de représenter l'apprenant dans un espace métrique (S, D_S) , qui permet de pouvoir mesurer la similarité entre situations d'apprentissage. Ce mode de représentation permet de formaliser la notion d'influence d'un feedback comme la différence entre la situation de l'apprenant avant et après avoir proposé ce feedback. Il est également possible de traduire des règles de sélection de feedbacks en représentant, pour chaque décision possible, l'ensemble des situations qui doivent la provoquer.

6.1.4 Adaptation

Nous étudions dans ces travaux deux types d'adaptation : une adaptation basée sur des règles expertes, et une adaptation automatique basée sur les pré-

cédentes interactions entre l'apprenant et l'environnement. Dans une première expérimentation, nous avons mesuré les bénéfices de l'apport d'une diversité de feedbacks sur l'apprentissage de l'apprenant. Les résultats de cette étude indiquent que les apprenants à qui l'on propose une diversité de feedbacks montrent une meilleure progression. En outre, cette progression se traduit après la phase d'apprentissage dans un exercice proposé en conditions réelles (sans feedback). Il existe néanmoins entre les apprenants de grandes disparités de performances, et il semble notamment que, si certains apprenants sont satisfaits des choix de feedbacks proposés par l'environnement, ces choix ne reflètent pas les besoins de l'ensemble des participants.

Nous nous intéressons ensuite aux travaux permettant de réviser automatiquement les règles de décision en fonction des interactions entre l'environnement et chaque apprenant. Un intérêt récent s'est porté à l'application des modèles issus des domaines des systèmes de recommandation et de l'apprentissage par renforcement dans le cadre d'environnements intelligents d'apprentissage. Nous présentons ces méthodes, leurs limites, et les défis inhérents à leur application dans le cadre d'environnements d'apprentissage, dans le chapitre 4. Nous proposons finalement une approche modélisant la sélection d'actions comme un problème de bandit, où il s'agit d'étudier le compromis entre exploration et exploitation des actions possibles afin de maximiser les gains obtenus lors d'une séquence de décision. Nous nous positionnons dans le cadre des approches SoftMax, qui visent à tirer aléatoirement une action en fonction de la probabilité qu'elle obtienne le meilleur gain, et où un paramètre de température permet de configurer la tendance du modèle à privilégier soit l'exploration, soit l'exploitation. Nous étendons ces approches pour qu'elles puissent prendre en compte un gain multidimensionnel (la progression de l'apprenant sur chacune des dimensions considérées). Enfin, nous proposons de pondérer les gains antécédents par leur pertinence par rapport à la situation lors de la décision, mobilisant pour cela les propriétés de similarité permises par la représentation en espaces métriques.

Nous comparons les deux approches proposées (adaptation depuis une expertise vs. adaptation automatique) dans une seconde expérimentation. Les résultats de cette étude illustrent les bénéfices liés à une adaptation prenant en compte les interactions antérieures, les participants ayant appris sur l'environnement "automatique" obtenant des résultats significativement meilleurs que les participants du groupe "expert". Bien sûr, ces résultats sont très dépendants de l'exhaustivité et de la précision du modèle de l'apprenant, ainsi que des règles de décision : si l'on imagine un contexte où le modèle de l'apprenant représente parfaitement ce dernier, et si les règles fournies par l'expert prennent en compte toutes les dimensions de ce modèle et sélectionnent alors l'action jugée comme la plus pertinente parmi un ensemble important de feedbacks possibles, alors l'approche "expert" obtiendra sûrement de meilleurs résultats que ceux obtenus lors de notre expérimentation. Toutefois, l'établissement de cette expertise sera un processus très long, et coûteux. En outre, il existera encore des spécificités liées au caractère individuel des besoins de l'apprenant qu'un ensemble de règles fixes ne pourront prendre en compte. Ces contraintes justifient les

recherches consacrées aux stratégies d’adaptation automatique (ces dernières pouvant néanmoins se baser sur une expertise partielle).

6.2 Limites et perspectives

6.2.1 Modélisation de l’activité

Les environnements intelligents d’apprentissage sont étudiés depuis plusieurs décennies, et leur application dans des domaines bien définis est aujourd’hui suffisamment mûre pour que ces environnements soient utilisés par des milliers d’élèves à travers le monde (Woolf, 2009). Les problématiques inhérentes à l’application des principes prônés par les environnements intelligents d’apprentissage dans le cadre de domaines mal définis, et a fortiori dans le cadre de l’apprentissage de gestes, sont néanmoins beaucoup plus récentes, et restent en majorité ouvertes. Considérant également la période d’apprentissage que nous considérons, nous avons volontairement réduit le nombre de dimensions caractérisant le geste. Cette limite n’altère pas les résultats de nos travaux, puisque l’important lors de la création d’un environnement d’apprentissage est de représenter le sous-ensemble du domaine sur lequel on souhaite que l’apprenant progresse (Fournier-Viger et al., 2013). Enrichir l’ensemble des dimensions d’erreur, tant en considérant des erreurs liées au geste de calligraphie (e.g. sur-compensation au niveau des doigts) que des facteurs liés à l’efficacité cognitive (e.g. attention, motivation de l’apprenant¹) constitue tout de même une perspective d’évolution de la plateforme CalliSmart.

6.2.2 Ensemble des actions pédagogiques

Enrichir l’offre pédagogique de CalliSmart, afin qu’il soit possible de proposer des feedbacks haptiques (Teo et al., 2002; Palluel-Germain et al., 2007; Bayart et al., 2005) ou sonores (Danna and Velay, 2015; Danna et al., 2015) constitue une seconde piste d’amélioration de la plateforme. Cette piste est actuellement étudiée au sein du projet Descript. En effet, les problématiques d’adaptation que nous avons étudiées sont motivées par la possibilité de proposer des feedbacks adaptés à l’apprenant. Il est donc essentiel, pour que la capacité de l’environnement à s’adapter ait un quelconque sens, que suffisamment de feedbacks différents soient proposés pour couvrir un grand nombre de situations d’apprentissage pour tout apprenant. L’ajout de feedbacks haptiques et sonores renforcera en ce sens l’intérêt pédagogique de l’environnement adaptatif CalliSmart.

6.2.3 Représentation des actions pédagogiques

Finalement, une perspective intéressante, et commune aux environnements intelligents classiques, est la prise en compte d’une représentation des actions

1. La détection de ces facteurs a notamment été étudiée dans le cadre d’environnements conversationnels (D’Mello and Graesser, 2010).

pédagogiques. Cette perspective nous apparaît d'autant plus importante que l'ensemble des feedbacks qu'un environnement peut proposer est riche et diversifié. Il n'est en effet plus possible, si l'ensemble des actions est suffisamment grand, d'explorer chacun des éléments dans un temps suffisamment court. Et il est alors intéressant de pouvoir prendre en compte des informations relatives à la similarité entre actions pour réduire l'ensemble à explorer (e.g. en privilégiant des feedbacks très différents de ceux proposés jusqu'alors, ou au contraire en privilégiant des feedbacks similaires aux feedbacks déjà proposés ayant reçu les meilleurs gains).

Dans les travaux relatés dans les précédents chapitres, nous avons modélisé l'ensemble des feedbacks par l'ensemble $A \in \{1, \dots, K\}$. Il n'existe alors pas de caractérisation de la nature du feedback, et chaque feedback est considéré individuellement. En outre, il est impossible de mesurer une quelconque similarité entre deux types de feedback. Un feedback peut pourtant être décrit par un ensemble de caractéristiques : sa modalité (e.g. information visuelle, sonore, haptique) ; son degré d'explicitation (i.e. depuis implicite vers explicite) ; la nature de l'information qu'il porte (e.g. connaissance des résultats, de la performance) ; et bien d'autres encore. On peut également affirmer que certains feedbacks sont plus similaires que d'autres.

Prendre en considération une représentation des actions reviendrait à modifier la nature de A . D'une manière analogue à l'approche nous ayant permis de représenter l'apprenant suivant plusieurs dimensions de son activité, il est possible d'identifier des dimensions caractérisant une action (i.e. les types de caractéristiques présentés précédemment). Une action est alors représentée suivant des dimensions A_p , $p = 1, \dots, q$. L'ensemble produit $A = A_1 \times A_2 \times \dots \times A_q$ contient alors toute action. Si l'on munit toute dimension A_p d'une distance D_{A_p} , l'espace métrique (A, D_A) , où $D_A = Q(D_{A_1}, \dots, D_{A_q})$, Q étant une norme L^p arbitraire sur \mathbb{R}^q , permet de représenter ces actions. À noter que définir la distance D_A comme une distance triviale² revient à adopter l'approche de représentation $A \in \{1, \dots, K\}$.

Il est alors nécessaire de proposer un algorithme de sélection permettant la prise en compte des représentations à la fois de l'apprenant et des actions. L'approche proposée par (Slivkins, 2014) dans le cadre des bandits contextuels nous paraît être une piste à considérer. Ces travaux prennent en effet en compte le produit des espaces métriques (P, D_P) , avec $P \subset S \times A$ et $D_P((s, a), (s', a')) = \min(1, D_S(s, s') + D_A(a, a'))$, et utilisent les probabilités topologiques de l'espace produit P .

6.2.4 Évolution de la plateforme CalliSmart et des modèles proposés

Nos travaux ont permis la conception d'un environnement intelligent d'apprentissage de la calligraphie. La plateforme CalliSmart sera prochainement

2. $D_A(a_1, a_2) = 1$ si $a_1 = a_2$, 0 si $a_1 \neq a_2$, sachant que $a_1 = a_2 \Leftrightarrow \forall p \in 1, \dots, q \ a_{1,p} = a_{2,p}$.

utilisée dans le cadre des formations proposées par l'ESAD. Outre les perspectives d'évolution de la plateforme présentées précédemment dans ce chapitre, la généralité des modèles proposés lors de ces travaux permettra leur prochaine utilisation dans le cadre de projets impliquant des interactions gestuelles dans un espace en trois dimensions au sein du CAVE de l'UTC. Un premier prototype d'environnement 3D pour l'apprentissage du *top spin* coup droit³ a notamment été conçu ces derniers mois dans le cadre du stage de Sarah Bali au sein de notre équipe.

3. Le top spin coup droit est l'un des coups les plus joués au tennis de table, et consiste à imprimer à la balle une forte rotation vers l'avant : <https://fr.wikipedia.org/wiki/Topspin>.

Bibliographie

- A. Aamodt and E. Plaza. Case-based reasoning : Foundational issues, methodological variations, and system approaches. *AI communications*, 7(1) : 39–59, 1994.
- V. Aleven. Using background knowledge in case-based legal reasoning : a computational model and an intelligent learning environment. *Artificial Intelligence*, 150(1) :183–237, 2003.
- J. R. Anderson. *The Architecture of Cognition*. Harvard University Press, Cambridge, MA, USA, 1983.
- J. R. Anderson and C. J. Lebiere. *The atomic components of thought*. Mahwah, NJ : Erlbaum, 1998.
- G. Arutyuny, V. Gurfinke, and M. Mirskii. Investigation of aiming at a target, 1968.
- P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3) :235–256, 2002a.
- P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1) :48–77, 2002b.
- R. S. Baker, A. T. Corbett, and V. Aleven. More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In *Intelligent Tutoring Systems*, pages 406–415. Springer, 2008.
- F. Bara, E. Gentaz, P. Colé, and L. Sprenger-Charolles. The visuo-haptic and haptic exploration of letters increases the kindergarten-children’s understanding of the alphabetic principle. *Cognitive development*, 19(3) :433–449, 2004.
- T. Barnes and J. Stamper. Toward automatic hint generation for logic proof tutoring using historical student data. In *International Conference on Intelligent Tutoring Systems*, pages 373–382. Springer, 2008.

- G.-M. Baschera and M. Gross. Poisson-based inference for perturbation models in adaptive spelling training. *International Journal of Artificial Intelligence in Education*, 20(4) :333–360, 2010.
- B. Bayart, A. Pocheville, and A. Kheddar. An adaptive haptic guidance software module for i-touch : example through a handwriting teaching simulation and a 3d maze. In *IEEE international workshop on haptic audio visual environments and their applications*, pages 6–pp. IEEE, 2005.
- C. R. Beal, I. Arroyo, P. R. Cohen, B. P. Woolf, and C. R. Beal. Evaluation of animalwatch : An intelligent tutoring system for arithmetic and fractions. *Journal of Interactive Online Learning*, 9(1) :64–77, 2010.
- J. Beck, B. P. Woolf, and C. R. Beal. Advisor : A machine learning architecture for intelligent tutor construction. *AAAI/IAAI*, 2000 :552–557, 2000.
- J. E. Beck and K.-m. Chang. Identifiability : A fundamental problem of student modeling. In *International Conference on User Modeling*, pages 137–146. Springer, 2007.
- J. E. Beck and J. Sison. Using knowledge tracing in a noisy environment to measure student reading proficiencies. *International Journal of Artificial Intelligence in Education*, 16(2) :129–143, 2006.
- H. Ben-Pazi, A. Ishihara, S. Kukke, and T. D. Sanger. Increasing viscosity and inertia using a robotically controlled pen improves handwriting in children. *Journal of child neurology*, 2009.
- N. A. Bernstein. *The co-ordination and regulation of movements*. Pergamon Press Ltd., 1967.
- F. Bevilacqua, B. Zamborlin, A. Sypniewski, N. Schnell, F. Guédy, and N. Ramamimanana. Continuous realtime gesture following and recognition. In *International Gesture Workshop*, pages 73–84. Springer, 2009.
- A. Beygelzimer, J. Langford, L. Li, L. Reyzin, and R. E. Schapire. Contextual bandit algorithms with supervised learning guarantees. In *AISTATS*, pages 19–26, 2011.
- K. K. Biswas and S. K. Basu. Gesture recognition using microsoft kinect®. In *Automation, Robotics and Applications (ICARA), 2011 5th International Conference on*, pages 100–103. IEEE, 2011.
- G. Bouchard and B. Triggs. The tradeoff between generative and discriminative classifiers. In *16th IASC International Symposium on Computational Statistics (COMPSTAT'04)*, pages 721–728, 2004.
- D. Bridge, M. H. Göker, L. McGinty, and B. Smyth. Case-based recommender systems. *The Knowledge Engineering Review*, 20(03) :315–320, 2005.

-
- J. S. Brown and R. R. Burton. Diagnostic models for procedural bugs in basic mathematical skills*. *Cognitive science*, 2(2) :155–192, 1978.
- P. Brusilovsky and E. Millán. User models for adaptive hypermedia and adaptive educational systems. In *The adaptive web*, pages 3–53. Springer-Verlag, 2007.
- S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *arXiv preprint arXiv :1204.5721*, 2012.
- C. J. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2) :121–167, 1998.
- R. Burke. Hybrid recommender systems : Survey and experiments. *User modeling and user-adapted interaction*, 12(4) :331–370, 2002.
- K. Carpentier. *Scénarisation personnalisée dynamique dans les environnements virtuels pour la formation*. PhD thesis, Université de Technologie de Compiègne, 2015.
- K. Carpentier, D. Lourdeaux, and I. M. Thouvenin. Dynamic selection of learning situations in virtual environment. In *ICAART (2)*, pages 101–110, 2013.
- B. Carr and I. P. Goldstein. Overlays : A theory of modelling for computer aided instruction. Technical report, DTIC Document, 1977.
- H. Cen, K. Koedinger, and B. Junker. Learning factors analysis—a general method for cognitive model evaluation and improvement. In *Intelligent tutoring systems*, pages 164–175. Springer, 2006.
- M. Chi, K. VanLehn, D. Litman, and P. Jordan. Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction*, 21(1-2) :137–180, 2011a.
- M. Chi, K. VanLehn, D. Litman, and P. Jordan. An evaluation of pedagogical tutorial tactics for a natural language tutoring system : A reinforcement learning approach. *International Journal of Artificial Intelligence in Education*, 21(1-2) :83–113, 2011b.
- K. Chrysafiadi and M. Virvou. Student modeling approaches : A literature review for the last decade. *Expert Systems with Applications*, 40(11) :4715–4729, Sept. 2013.
- W. J. Clancey. Use of mycin’s rules for tutoring. *Rule-Based Expert Systems*, pages 19–2, 1984.
- B. Clement, D. Roy, P.-Y. Oudeyer, and M. Lopes. Multi-armed bandits for intelligent tutoring systems. *arXiv preprint arXiv :1310.3174*, 2013.

- B. Clement, P.-Y. Oudeyer, D. Roy, and M. Lopes. Online optimization of teaching sequences with multi-armed bandits. In *Educational Data Mining 2014*, 2014.
- C. Conati. Intelligent tutoring systems : New challenges and directions. In *IJCAI*, volume 9, pages 2–7, 2009.
- A. T. Corbett and J. R. Anderson. Knowledge tracing : Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 4(4) :253–278, 1994.
- J. Danna and J.-L. Velay. Basic and supplementary sensory feedback in handwriting. *Frontiers in Psychology*, 6, Feb. 2015. ISSN 1664-1078.
- J. Danna, M. Fontaine, V. Paz-Villagrán, C. Gondre, E. Thoret, M. Aramaki, R. Kronland-Martinet, S. Ystad, and J.-L. Velay. The effect of real-time auditory feedback on learning new characters. *Human movement science*, 43 :216–228, 2015.
- D. Delignières, D. Nourrit, B. Lauriot, and I. Cadjee. L’acquisition des habiletés motrices complexes. *Science et motricité*, pages 82–82, 1998a.
- D. Delignières, D. Nourrit, R. Sioud, P. Leroyer, M. Zattara, and J.-P. Micallef. Preferred coordination modes in the first steps of the learning of a complex gymnastics skill. *Human Movement Science*, 17(2) :221–241, 1998b.
- M. C. Desmarais and R. S. Baker. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, 22(1-2) :9–38, 2012.
- S. D’Mello, A. Olney, C. Williams, and P. Hays. Gaze tutor : A gaze-reactive intelligent tutoring system. *International Journal of human-computer studies*, 70(5) :377–398, 2012.
- P. Doan, R. Frenoy, O. Gapenne, I. Thouvenin, P. Steiner, and D. Boutet. Descript, a multimodal adaptive learning platform focusing on attention training. In *the Seventh Conference of the International Society for Gesture Studies*, page 164, Paris, 2016.
- H. Drachsler, H. G. Hummel, and R. Koper. Personal recommender systems for learners in lifelong learning networks : the requirements, techniques and model. *International Journal of Learning Technology*, 3(4) :404–423, 2008.
- T. Dragon, B. P. Woolf, D. Marshall, and T. Murray. Coaching within a domain independent inquiry environment. In *International Conference on Intelligent Tutoring Systems*, pages 144–153. Springer, 2006.
- S. K. D’Mello and A. Graesser. Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction*, 20(2) :147–187, 2010.

-
- A. A. Efros, A. C. Berg, G. Mori, and J. Malik. Recognizing action at a distance. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 726–733. IEEE, 2003.
- D. Feygin, M. Keehner, and R. Tendick. Haptic guidance : Experimental evaluation of a haptic training method for a perceptual motor skill. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on*, pages 40–47. IEEE, 2002.
- W. T. Fitch and G. Kramer. Sonifying the body electric : Superiority of an auditory over a visual display in a complex, multivariate system. In *Sante Fe Institute Studies in the Sciences of Complexity*, volume 18, pages 307–307. Addison-Wesley Publishing, 1994.
- P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6) :381, 1954.
- P. M. Fitts and M. I. Posner. *Human performance*. Brooks/Cole, 1967.
- P. Fournier-Viger, R. Nkambou, and E. Mephu Nguifo. Learning procedural knowledge from user solutions to ill-defined tasks in a simulated robotic manipulator. *Romero, et al.(eds.) Handbook of Educational Data Mining*, pages 451–465, 2010a.
- P. Fournier-Viger, R. Nkambou, and E. M. Nguifo. Building intelligent tutoring systems for ill-defined domains. In *Advances in intelligent tutoring systems*, pages 81–101. Springer, 2010b.
- P. Fournier-Viger, R. Nkambou, E. M. Nguifo, A. Mayers, and U. Faghihi. A multiparadigm intelligent tutoring system for robotic arm training. *IEEE Transactions on Learning Technologies*, 6(4) :364–377, 2013.
- L. Fricoteaux. *Interaction informée par un système décisionnel pour un retour multimodal en environnement virtuel : application à la formation au transport fluvial*. PhD thesis, Université de Technologie de Compiègne, 2012.
- A. C. Graesser, P. Wiemer-Hastings, K. Wiemer-Hastings, D. Harter, T. R. G. Tutoring Research Group, and N. Person. Using latent semantic analysis to evaluate the contributions of students in autotutor. *Interactive learning environments*, 8(2) :129–147, 2000.
- J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1) :5–53, 2004.
- W. E. Hick. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4(1) :11–26, 1952.

- R. Hodhod and D. Kudenko. Interactive narrative and intelligent tutoring for ethics domain. *Intelligent Tutoring Systems for Ill-Defined Domains : Assessment and Feedback in Ill-Defined Domains.*, page 13, 2008.
- J. Holland, B. Martin, and A. Mitrovic. J-latte : a constraint-based tutor for java. 2009.
- K. G. Holt, J. Hamill, and R. O. Andres. Predicting the minimal energy costs of human walking. *Med Sci Sports Exerc*, 23(4) :491–8, 1991.
- P. Holt, S. Dubs, M. Jones, and J. Greer. The state of student modelling. In *Student modelling : The key to individualized knowledge-based instruction*, pages 3–35. Springer, 1994.
- A. Iglesias, P. Martínez, R. Aler, and F. Fernández. Learning teaching strategies in an adaptive and intelligent educational system through reinforcement learning. *Applied Intelligence*, 31(1) :89–106, 2009.
- L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning : A survey. *Journal of artificial intelligence research*, 4 :237–285, 1996.
- T. Käser, S. Klingler, A. G. Schwing, and M. Gross. Beyond knowledge tracing : Modeling skill topologies with bayesian networks. In *Intelligent Tutoring Systems*, pages 188–198. Springer, 2014.
- A. Kobsa. Generic user modeling systems. *User modeling and user-adapted interaction*, 11(1-2) :49–63, 2001.
- K. R. Koedinger, J. R. Anderson, W. H. Hadley, and M. A. Mark. Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8 :30–43, 1997.
- R. Kumar, C. P. Rosé, Y.-C. Wang, M. Joshi, and A. Robinson. Tutorial dialogue as adaptive collaborative learning support. *Frontiers in artificial intelligence and applications*, 158 :383, 2007.
- J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields : Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the eighteenth international conference on machine learning, ICML*, volume 1, pages 282–289, 2001.
- S. Lallé. *Assistance à la construction et à la comparaison de techniques de diagnostique des connaissances*. PhD thesis, Université de Grenoble, 2013.
- J. Langford and T. Zhang. The epoch-greedy algorithm for multi-armed bandits with side information. In J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis, editors, *Advances in Neural Information Processing Systems 20*, pages 817–824. Curran Associates, Inc., 2008.

-
- N.-T. Le, F. Loll, and N. Pinkwart. Operationalizing the continuum between well-defined and ill-defined problems for educational technology. *IEEE Transactions on Learning Technologies*, 6(3) :258–270, 2013.
- L. Li, W. Chu, J. Langford, and R. E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670. ACM, 2010.
- B. López. Case-based reasoning : a concise introduction. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 7(1) :1–103, 2013.
- C. Lynch, K. Ashley, V. Alevan, and N. Pinkwart. Defining ill-defined domains ; a literature survey. In *Proceedings of the workshop on intelligent tutoring systems for ill-defined domains at the 8th international conference on intelligent tutoring systems*, pages 1–10, 2006.
- S. Manitsaris, A. Glushkova, F. Bevilacqua, and F. Moutarde. Capture, modeling, and recognition of expert technical gestures in wheel-throwing art of pottery. *Journal on Computing and Cultural Heritage (JOCCH)*, 7(2) :10, 2014.
- M. J. Mayo. *Bayesian student modelling and decision-theoretic selection of tutorial actions in intelligent tutoring systems*. PhD thesis, University of Canterbury. Computer Science, 2001.
- S. Mitra and T. Acharya. Gesture recognition : A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(3) :311–324, 2007.
- A. Mitrovic and S. Ohlsson. Evaluation of a constraint-based tutor for a database. *International Journal of Artificial Intelligence in Education*, 10 : 238–256, 1999.
- A. Mitrovic, K. R. Koedinger, and B. Martin. A comparative analysis of cognitive tutoring and constraint-based modeling. In *User Modeling 2003*, pages 313–322. Springer, 2003.
- A. Mitrovic, A. Weerasinghe, V. Dinatrova, R. Mizoguchi, and A. Graessner. Revisiting ill-definedness and the consequences for itss. *Artificial intelligence in education : Building learning systems that care from knowledge representation to affective modelling*, pages 375–382, 2009.
- S. Moritz and G. Blank. Generating and evaluating object-oriented designs for instructors and novice students. *Intelligent Tutoring Systems for Ill-Defined Domains : Assessment and Feedback in Ill-Defined Domains.*, page 35, 2008.
- T. Murray and I. Arroyo. Toward measuring and maintaining the zone of proximal development in adaptive instructional systems. In *International Conference on Intelligent Tutoring Systems*, pages 749–758. Springer, 2002.

- S. Narciss, S. Sosnovsky, L. Schnaubert, E. Andrès, A. Eichelmann, G. Gogvadze, and E. Melis. Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education*, 71 : 56–76, 2014.
- A. Newell and P. S. Rosenbloom. Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*, 1 :1–55, 1981.
- K. M. Newell. Constraints on the development of coordination. *Motor development in children : Aspects of coordination and control*, 34 :341–360, 1986.
- A. Y. Ng and M. I. Jordan. On discriminative vs. generative classifiers : A comparison of logistic regression and naive bayes. In *Advances in Neural Information Processing Systems*, pages 841–848, 2002.
- S. Ohlsson. Constraint-based student modeling. In *Student modelling : the key to individualized knowledge-based instruction*, pages 167–189. Springer, 1994.
- R. Oppermann and R. Rasher. Adaptability and adaptivity in learning systems. *Knowledge transfer*, 2 :173–179, 1997.
- P.-Y. Oudeyer and F. Kaplan. What is intrinsic motivation? a typology of computational approaches. *Frontiers in neurorobotics*, 1 :6, 2009.
- J. Paillard. Les bases nerveuses du contrôle visuo-manuel de l’écriture [the neural bases of the visuo-manual control of handwriting]. *L’écriture : Le cerveau, l’œil et la main (pp. 245e266)*. Turnhout, France : Brepols, 1990.
- J. Paillard. Les bases nerveuses du contrôle visuo-manuel de l’écriture [the neural bases of the visual-manual control of handwriting]. In *L’écriture : Le cerveau, L’œil Et La Main [Writing : Brain, Eye, and Hand]*, pages 23–52. C. Sirat, J. Irigoien, and E. Poulle (éd.), 1991.
- R. Palluel-Germain, F. Bara, A. H. de Boisferon, B. Hennion, P. Gouagout, and E. Gentaz. A visuo-haptic device-telemaque-increases kindergarten children’s handwriting acquisition. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC’07)*, pages 72–77. IEEE, 2007.
- P. I. Pavlik Jr, H. Cen, and K. R. Koedinger. Performance factors analysis—a new alternative to knowledge tracing. *Online Submission*, 2009.
- M. J. Pazzani and D. Billsus. Content-based recommendation systems. In *The adaptive web*, pages 325–341. Springer, 2007.
- B. Plimmer, P. Reid, R. Blagojevic, A. Crossan, and S. Brewster. Signing on the tactile line : A multimodal system for teaching handwriting to blind children. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 18 (3) :17, 2011.

-
- A. Quattoni, S. Wang, L.-P. Morency, M. Collins, and T. Darrell. Hidden conditional random fields. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(10) : 1848–1852, 2007.
- L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2) :257–286, 1989.
- E. Rich. User modeling via stereotypes*. *Cognitive science*, 3(4) :329–354, 1979.
- R. Ronsse, V. Puttemans, J. P. Coxon, D. J. Goble, J. Wagemans, N. Wenderoth, and S. P. Swinnen. Motor learning with augmented feedback : modality-dependent behavioral and neural consequences. *Cerebral cortex*, 21(6) :1283–1294, 2011.
- A. W. Salmoni, R. A. Schmidt, and C. B. Walter. Knowledge of results and motor learning : a review and critical reappraisal. *Psychological bulletin*, 95(3) :355, 1984.
- R. A. Schmidt and T. Lee. *Motor control and learning*. Human kinetics, 1988.
- S. Shen and M. Chi. Reinforcement learning : The sooner the better, or the later the better? In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, UMAP 2016, pages 37–44, New York, NY, USA, 2016. ACM.
- V. J. Shute. Stealth assessment in computer-based games to support learning. *Computer games and instruction*, 55(2) :503–524, 2011.
- V. J. Shute and R. Glaser. A large-scale evaluation of an intelligent discovery world : Smithtown. *Interactive Learning Environments*, 1(1) :51–77, 1990.
- R. Sigrist, G. Rauter, R. Riener, and P. Wolf. Augmented visual, auditory, haptic, and multimodal feedback in motor learning : a review. *Psychonomic bulletin & review*, 20(1) :21–53, 2013.
- H. A. Simon. The structure of ill structured problems. *Artificial intelligence*, 4(3-4) :181–201, 1973.
- P. Simonet. *Apprentissages moteurs : processus et procédés d’acquisition*. Vigot, 1985.
- R. E. Slavin. Research on cooperative learning : Consensus and controversy. *Educational leadership*, 47(4) :52–54, 1990.
- A. Slivkins. Contextual bandits with similarity information. *Journal of Machine Learning Research*, 15(1) :2533–2568, 2014.
- M. A. Smith, A. Ghazizadeh, and R. Shadmehr. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol*, 4(6) :e179, 2006.

- Y. Soullard. *Classification et détection de figures chartistes par apprentissage statistique*. PhD thesis, Université Pierre et Marie Curie-Paris 6, 2013.
- N. Søvik and H.-L. Teulings. Real-time feedback of handwriting in a teaching program. *Acta psychologica*, 54(1-3) :285–291, 1983.
- P. Suraweera and A. Mitrovic. Kermit : A constraint-based tutor for database modeling. In *Intelligent Tutoring Systems*, pages 377–387. Springer, 2002.
- D. Suthers, J. Connelly, A. Lesgold, M. Paolucci, E. Toth, J. Toth, and A. Weiner. Representational and advisory guidance for students learning scientific inquiry. *Smart machines in education : The coming revolution in educational technology*, pages 7–35, 2001.
- R. S. Sutton and A. G. Barto. *Reinforcement learning : An introduction*. MIT press Cambridge, 1998.
- G. A. ten Holt, M. J. Reinders, and E. Hendriks. Multi-dimensional dynamic time warping for gesture recognition. In *Thirteenth annual conference of the Advanced School for Computing and Imaging*, volume 300, 2007.
- C. L. Teo, E. Burdet, and H. Lim. A robotic teacher of chinese handwriting. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on*, pages 335–341. IEEE, 2002.
- J. R. Tetreault and D. J. Litman. Using reinforcement learning to build a better model of dialogue state. In *EACL*, 2006.
- J. R. Thomas. Acquisition of motor skills : Information processing differences between children and adults. *Research Quarterly for Exercise and Sport*, 51 (1) :158–173, 1980.
- I. M. Thouvenin. *Interaction et connaissance : construction d’une expérience dans le monde virtuel*. PhD thesis, Habilitation à diriger des recherches, Université de Technologie de Compiègne, 2009.
- G. Tsoumakas and I. Katakis. Multi-label classification : An overview. *Dept. of Informatics, Aristotle University of Thessaloniki, Greece*, 2006.
- G. P. Van Galen. Handwriting : Issues for a psychomotor theory. *Human movement science*, 10(2-3) :165–191, 1991.
- G. Veletsianos and G. S. Russell. Pedagogical agents. In *Handbook of research on educational communications and technology*, pages 759–769. Springer, 2014.
- K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic, and E. Duval. Context-aware recommender systems for learning : a survey and future challenges. *IEEE Transactions on Learning Technologies*, 5(4) : 318–335, 2012.

-
- B. Vereijken, H. Whiting, and W. Beek. A dynamical systems approach to skill acquisition. *The Quarterly Journal of Experimental Psychology*, 45(2) : 323–344, 1992.
- J. Vermorel and M. Mohri. Multi-armed bandit algorithms and empirical evaluation. In *European conference on machine learning*, pages 437–448. Springer, 2005.
- L. S. Vygotsky. *Mind in society : The development of higher mental process*, 1978.
- E. Walker, A. Ogan, V. Alevan, and C. Jones. Two approaches for providing adaptive support for discussion in an ill-defined domain. *Intelligent Tutoring Systems for Ill-Defined Domains : Assessment and Feedback in Ill-Defined Domains*, 1, 2008.
- J. Wang, Z. Liu, Y. Wu, and J. Yuan. Mining actionlet ensemble for action recognition with depth cameras. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 1290–1297. IEEE, 2012.
- A. Weerasinghe and A. Mitrovic. Facilitating adaptive tutorial dialogues in eer-tutor. In *Artificial Intelligence in Education*, pages 630–631. Springer, 2011.
- M. J. Weil and S. J. C. Amundson. Relationship between visuomotor and handwriting skills of children in kindergarten. *American Journal of Occupational Therapy*, 48(11) :982–988, 1994.
- J. L. Wheeler and J. W. Regian. The use of a cognitive tutoring system in the improvement of the abstract reasoning component of word problem solving. *Computers in Human Behavior*, 15(2) :243–254, 1999.
- B. White and J. Frederiksen. Developing metacognitive knowledge and processes : The key to making scientific inquiry and modeling accessible to all students. *An Overview of the ThinkerTools Inquiry Project*, 1995.
- B. P. Woolf. *Building intelligent interactive tutors student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann Publishers/Elsevier, Amsterdam ; Boston, 2009.
- J. Wu, G. Pan, D. Zhang, G. Qi, and S. Li. Gesture recognition with a 3-d accelerometer. In *International Conference on Ubiquitous Intelligence and Computing*, pages 25–38. Springer, 2009.
- M.-L. Zhang and Z.-H. Zhou. A review on multi-label learning algorithms. *IEEE transactions on knowledge and data engineering*, 26(8) :1819–1837, 2014.