

وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY

UNIVERSITE BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des Sciences de L'Ingéniorat

Département d'Informatique

THESE

Présentée en vue de l'obtention

du diplôme de Doctorat 3^{ème} Cycle

**Un Système Bio-Inspiré pour l'Adaptation
Pédagogique dans un EIAH**

Domaine : Mathématiques et Informatique

Filière : Informatique

Spécialité : Ingénierie des Connaissances -IC-

Par : **Melle DEBBAH Amina**

Devant le Jury

Président Mr Tahar BENSEBAA Professeur Université Badji Mokhtar-Annaba

Rapporteur Mme Yamina MOHAMED BEN ALI Professeur Université Badji Mokhtar-Annaba

Examineurs Mr Yacine LAFIFI MCA Université 8 Mai 1945 - Guelma

Mme Hassina SERIDI Professeur Université Badji Mokhtar-Annaba

A mes parents,

A l'ouverture d'esprit ...

Remerciements

J'adresse mes premiers remerciements à Ma directrice de thèse Mohamed Ben Ali Yamina, Professeur à l'Université Badji Mokhtar Annaba, pour m'avoir dirigé tout au long de cette thèse. Je l'a remercie pour sa présence et disponibilité, de ces effort afin de guider ce travail. J'ai beaucoup bénéficié de ces conseils, de la documentation qu'elle m'a fournie et de ces suggestions pertinentes pour la mise au point de ce travail.

Je tiens aussi à remercier Bensebaa Taher, Professeur à l'université Badji Mokhtar Annaba, de m'avoir fait l'honneur de présider le jury de ma soutenance.

Je tiens aussi à remercier Seridi Hassina, Professeur à l'université Badji Mokhtar-Annaba, Laafifi Yacine, Maitre de conférences à l'université de Guelma de m'avoir fait l'honneur d'accepter d'examiner cette thèse.

Mes remerciements les plus chaleureux sont adressés à mes parents ma chère mère et mon cher père qui m'ont soutenu tout au long de ces années, aussi mes deux frères. Je remercie toute personne ayant aidé de près ou de loin à la réalisation de ce travail.

Que cette thèse soit un témoignage de ma gratitude envers mes parents et tous les membres de ma famille qui m'ont apporté leur soutien moral tout au long de cette thèse !

المخلص

الموضوع المعالج في هذه الأطروحة يندرج ضمن الإشكالية العامة للبيئات التفاعلية للتعليم البشري. فهو يعالج بشكل خاص تشخيص التعلم في عملية التعليم، تشخيص التعلم يلعب دورا جد كبير، و هذا لأنه طريقة جيدة لنقل التعليم

الهدف الرئيسي هو تصميم رسم بياني للأغراض البيداغوجية من أجل إقتراح مواد دراسية مشخصة، هذا هو، في الحقيقة التوليد التلقائي للأغراض البيداغوجية المضبوطة على المتعلم الخاص، إعتبارا من الرسم البياني للأغراض البيداغوجية و إعتقادا على إحتياجاته، و تفضيلاته، و معرفته. الموارد البيداغوجية المسماة حاليا أغراض بيداغوجية مفهومة بإستعمال معايير و مقاييس للفوقية التعليمية. هذه الأغراض البيداغوجية، التي تمثل المحتوى التعليمي، يتم تجميعها إعتبارا من حبيبات بيداغوجية و من بعد متحدة، من أجل تأليف الدورات الفردية للتكوين تدعى مناهج علمية

نقترح في إطار هذا العمل، مقاربة جديدة تستند على حساب الحمض النووي في رسم بياني للأغراض البيداغوجية مستندة على مدى التمكن، التفضيلة التعليمية المستعملة. تمثيل الملف الشخصي للمتعلم يصنع لاحقا، نقترح طريقة مستندة على حساب الحمض النووي من أجل توليد التسلسل الشخصي للأغراض البيداغوجية

في المخلص، إعتبارا من مجموعة الأغراض البيداغوجية و وصفهم الدالي و الملف الشخصي للمتعلم، النظام المترقب يجب أن يكون قادرا على توليد الأغراض المكيفة (دروس) مع السطح البيئي للتصفح على الإنترنت

الكلمات الرئيسية

الأغراض البيداغوجية، سلسلة مخصصة، الملف الشخصي للمتعلم، حساب الحمض النووي

Abstract

The subject matter of this thesis is part of a general problem of the Interactive Environments for Human Learning. It specifically addresses the personalization of learning. In the learning process, customizing learning plays a very important role because it is a good way of learning delivery.

The main goal is to design a graph of learning objects to the proposal custom routes. This is, in fact, the automatic generation of learning objects for particular learner, from a graph of learning objects and according to their needs, preferences and knowledge. Educational resources currently called learning objects are indexed using standards and educational metadata standards like SCORM and LOM. These learning objects, which are content to learn, are assembled from teaching grains and combined to form individual training course called Curriculum.

We propose in the context of this work, a new approach based on the calculation of weights in the learning object graph based on the mastery rate, the learning preferences used. The representation of the learner profile is done subsequently. We propose a method based on the DNA computing to generate personalized sequence of learning objects.

In summary, from a set of learning objects, their semantic descriptions and learner profile, the expected system must able to generate suitable learning objects (courses) with a navigation interface on the web.

Keywords: personalized Sequence, learning objects, LOM, SCORM, DNA Computing, learner profile.

Résumé

Le sujet traité par cette thèse s'intègre dans une problématique générale des Environnements Informatiques pour l'Apprentissage Humain (EIAH). Il traite plus particulièrement la personnalisation de l'apprentissage. Dans le processus d'apprentissage, la personnalisation de l'apprentissage joue un rôle très considérable car c'est un bon moyen de transmission de l'apprentissage.

Le but principal est de concevoir un graphe d'objets pédagogique pour la proposition des parcours personnalisés. Il s'agit, en fait, de la génération automatique d'objets pédagogiques adaptés à un apprenant particulier, à partir d'un graphe d'objets pédagogiques et en fonction de ses *besoins*, ses *préférences* et ses *connaissances*. Les ressources pédagogiques appelés actuellement *objets pédagogiques* sont indexés en utilisant des normes et des standards de métadonnées éducatifs comme LOM et SCORM. Ces objets pédagogiques, qui constituent le contenu à apprendre, sont assemblés à partir de grains pédagogiques puis associées pour constituer des parcours individuels de formation appelé Curriculum.

Nous proposons dans le cadre de ce travail, une nouvelle approche basée sur le calcul de poids dans le graphe d'objet pédagogique basée sur le *taux de maîtrise*, la *préférence d'apprentissage* utilisée. La représentation du profil de l'apprenant se fait Par la suite, nous proposons une méthode basée sur le calcul de l'ADN afin de générer la séquence personnalisée des objets pédagogiques.

En résumé, à partir d'un ensemble d'objets pédagogique, de leurs descriptions sémantiques et du profil de l'apprenant, le système attendu doit être capable de générer des objets adaptés (cours) avec une interface de navigation sur le Web.

Mots clés : EIAH, Séquence personnalisée, objets pédagogiques, LOM, SCORM, calcul de l'ADN, profil de l'apprenant.

Table des matières

Introduction générale	1
1. Contexte général du travail.....	1
2. Problématique et approches.....	1
2.1 Problématique	1
2.2 Approche	2
2.3 Objectifs	3
3. Organisation du manuscrit.....	3
Chapitre 1	5
La personnalisation de l'apprentissage dans les EIAH	5
1.1 Introduction	5
1.2 Qu'est ce que personnaliser dans un processus d'apprentissage	5
1.2.1 Définitions de la Personnalisation.....	5
1.2.2 Définition de l'individualisation.....	6
1.2.3 La Différence entre individualisation et personnalisation	7
1.3 Apprentissage personnalisé	8
1.3.1 Phases de l'apprentissage personnalisé.....	9
1.3.2 Composantes de l'apprentissage personnalisé	9
1.4 Apprentissage individualisé.....	10
1.4.1 Les étapes de l'apprentissage individualisé	10
1.5 A quoi sert la personnalisation et pour qui ?.....	11
1.6 Que peut-on personnaliser dans un EIAH ?.....	12
1.7 Qui personnalise l'EIAH ?	14
1.7.1 Personnalisation par l'apprenant	14
1.7.2. Personnalisation par l'enseignant.....	15
1.7.3. Personnalisation par le système.....	18
1.7.4 Synthèse sur les acteurs de la personnalisation	19
1.8 Conclusion.....	20
Chapitre 2	21
L'objet pédagogique et l'apprentissage	21
2.1 Introduction	21
2.2 Objets Pédagogiques, métaphores et esquisse de définition.....	21
2.3 Classes d'objets Pédagogiques	23

2.3.1 Les Unités d'Apprentissage	23
2.3.2 Les Ressources Pédagogiques	24
2.3.3 Les Activités Pédagogiques	24
2.4 Objet et Objet pédagogique	25
2.5 Technologies associées aux objets pédagogiques	26
2.5.1 Les métadonnées	26
2.5.1.1 Importance des métadonnées.....	27
2.5.1.2 Intérêt des métadonnées	27
2.5.2 Technologies associées	28
2.5.2.1 Web sémantique	29
2.5.2.2 Ontologie.....	29
2.5.2.3 Langages du web sémantique.....	30
2.6 Standards et normes pour les objets pédagogiques	32
2.6.1 L'exemple LOM.....	32
2.6.2 L'exemple SCORM.....	34
2.6.3 L'exemple IMS-LD (Learning Design).....	36
2.8 Conclusion.....	39
Chapitre 3	41
La personnalisation de l'apprentissage et le calcul évolutionnaire.....	41
3.1 Introduction	41
3.2 Le calcul évolutionnaire	41
3.2.1 L'optimisation par les essais particuliers (EOP)	42
3.2.2 Les algorithmes génétiques (AGs)	45
3.2.3 Les algorithmes de colonies de fourmis	48
3.3 Personnalisation de l'apprentissage basée calcul évolutionnaire	49
3.3.1 La personnalisation basée ACO	50
3.3.2 La personnalisation basée EOP	52
3.3.3 La personnalisation basée AG.....	53
3.4 Discussion	56
3.5 Conclusion.....	58
Chapitre 4	59
Le calcul de l'ADN.....	59
4.1 Introduction	59
4.2 L'ADN composant génétique	59

4.3 Les outils pour le calcul de l'ADN	62
4.3.1 Les enzymes.....	63
4.3.2 Les polymérase	63
4.3.3 Ligase	65
4.3.4 Enzymes de restrictions	66
4.3.5 Electrophorèse.....	67
4.3.6 PCR (Polymerase Chain Reaction)	69
4.4 De l'ADN biologique vers les ordinateurs à ADN	71
4.5 Exemples de résolutions de Problèmes combinatoires.....	73
4.5.1 Le problème du chemin hamiltonien.....	73
4.5.2 Le problème du sac à dos.....	75
4.5.3 Le problème SAT	75
4.6 Synthèse	76
4.7 Conclusion.....	76
Chapitre 5	78
Modélisation de la personnalisation de l'apprentissage : les objets pédagogiques.....	78
5.1 Modèle global de la méthode proposée	78
5.2 Conception des objets pédagogiques.....	80
5.3 Conception du profil de l'apprenant.....	81
5.3 Stratégies de conception d'un graphe.....	83
5.5 Calcul du poids proposé	85
5.6 Conclusion.....	87
Chapitre 6	88
Personnalisation de l'apprentissage et calcul ADN	88
6.1 Introduction	88
6.2 Conception d'une séquence personnalisée	88
6.2.1 La recherche d'une séquence dans un graphe	89
6.2.2 Le calcul de la fonction objectif.....	89
6.3 Scénario de la personnalisation de l'apprentissage	89
6.3.1 Construction du profil de l'apprenant.....	92
6.3.2 Caractéristiques d'un objet pédagogique	92
6.4 Application de la méthode du calcul de l'ADN.....	92
6.4.1. Étape 1 : codage	94
6.4.2 Étape 2 : construction de brins d'ADN d'arcs	96

6.4.3 Étape 3 : Ligature	98
6.4.4 Étape 4 : Amplification (T2, T3).....	101
6.4.5 Étape 5 : Amplification (T3, T4).....	102
6.4.6 Étape 6 : la séparation (T4, T5).....	102
6.5 PERSO : système pour la personnalisation de l'apprentissage PERSO	103
6.5.1 Démarche pédagogique de PERSO	103
6.5.2 Architecture de PERSO	103
6.5.2.1 Architecture logicielle.....	103
6.5.2.2 Architecture fonctionnelle.....	104
6.5.3 Expérimentation PERSO	106
6.5.3.1 Résultats.....	107
6.5.3.2 Discussion des résultats	110
6.6 Conclusion.....	111
Conclusion générale et Perspectives	112
Bibliographie	116
Webographie	132
Annexe A	133
(1). Quelques interfaces du système PERSO.....	133

Liste des Figures

FIG. 1.1– Etapes de l'apprentissage individualisé.	10
FIG.1.2– Rôle de la personnalisation selon les acteurs.	11
FIG1.3– Écrans d'ANDES permettant à l'apprenant de sélectionner le problème qu'il va résoudre.....	14
FIG.1.4– Menu de LILIMATH permettant à l'apprenant d'obtenir une série d'exercices en fonction d'un thème.....	15
FIG.1.5–Interface de description du contenu pédagogique à utiliser dans REDEEM.	16
FIG.1.6– Interface de définition d'une stratégie pédagogique dans REDEEM	16
FIG.1.7– Éditeur de méta-stratégies dans EON.	17
FIG.2.1– Architecture de l'Unité d'apprentissage OUN-EML.....	25
FIG.2.2 – Structure générale de LOM (LOM).....	34
FIG.2.3 – Vue sur le contenu partageable de SCORM.	35
FIG.2.4– Architecture de la spécification IMS-LD, Niveaux A, B et C.	37
FIG.3.1 – Fourmis réelles suivent un chemin entre le Nid et la Nourriture.	48
FIG.3.2 – Exemple de graphe pour la personnalisation de l'apprentissage.....	52
FIG.4.1 – Les bases de l'ADN.....	60
FIG.4.2 – La structure double hélice d'ADN.....	61
FIG.4.3 – Enzymes autour de l'ADN durant la réplication et la synthèse protéique..	63
FIG.4.4– L'action de ligase.....	65
FIG.4.5– Actions des enzymes de restriction	66
FIG.4.6 – Electrophorèse.....	67
FIG.4.7– Le principe de l'électrophorèse en gel.....	68
FIG.4.8 – Le principe du PCR	71
FIG.5.1– Modèle général de la méthode.	79
FIG.5.2– Composants du profil apprenant.....	82
FIG.5.3– Solutions d'un problème d'optimisation combinatoire.	84
FIG.5.4 – Les relations entre objets pédagogique.....	85
FIG.5.5– Exemple d'un graphe basé poids.....	87
FIG.6.1– Scénario de la personnalisation.	91
FIG.6.2– Représentation du profil apprenant après détermination du but	92
FIG.6.3– La représentation des métadonnées d'objets pédagogiques.....	93
FIG.6.4– Organigramme représentant les étapes de l'algorithme du calcul de l'ADN	94

FIG.6.5– Architecture logiciel de PERSO.....	104
FIG.6.6– Architecture fonctionnelle de PERSO.....	106
FIG.6.7– Interface principale du système PERSO.....	107
FIG.6.8– Graphe des objets pédagogique du module algorithmique.....	108
FIG.6.9– Variation des poids en fonction des paramètres PA, et taux de maîtrise TM.....	109
FIG.A.1– Exemple d’objets pédagogiques.....	133
FIG.A.2– Formulaire des préférences.....	134
FIG.A.3– Exemple d’un QCM.....	134

Liste des tableaux

TAB.1.1– Comparaison entre la personnalisation et l’individualisation.....	7
TAB.1.2–Les différents acteurs dans la personnalisation.....	11
TAB.2.1–Synthèse sur les trois standards.	39
TAB.3.1– Algorithme d’optimisation par essaim particulaire.....	45
TAB.3.2– Algorithme génétique.	47
TAB.3.3–Algorithme de colonie de fourmis.	49
TAB.4.1–Comparaison entre ordinateurs à ADN et ordinateurs conventionnels.	73
TAB.5.1– L’algorithme général de la méthode proposée.....	79
TAB.5.2–Valeurs des propriétés possible pour un objet pédagogique	81
TAB.5.3–Valeurs des propriétés des préférences.	83
TAB.6.1–Codes ADN des objets pédagogiques.	95
TAB.6.2–Codes ADN des poids.	96
TAB.6.3–Codes ADN des arcs.	97
TAB.6.4–Résultats de l’opération de ligase.	97
TAB.6.5–Résultats du Réaction en chaîne par Polymérase (PCR)(1).	101
TAB.6.6–Résultats du Réaction en chaîne par Polymérase (PCR)(2).	102
TAB.6.7–Résultats de la séparation.	103
TAB.6.8–Taux de maîtrise des objets présenté par PERSO.....	108
TAB.6.9–Préférences des objets présentés par PERSO.	108
TAB.6.10–Paramètres de distributions des taux de maîtrise par PERSO.	110

Table des acronymes

EIAH Environnements Informatiques pour l'Apprentissage Humain

OP Objet pédagogique

LO Learning Object

LOM Learning Object Metadata

IEEE Institute of Electrical and Electronics Engineers

SCORM Sharable Content Object Reference Model

LTSC Learning Technology Standards Committee

AES Algorithmes Evolutionnaires

SCO Sharable Content Object

CO Content Organization

AGs Algorithmes Génétiques

SE Stratégie d'évolution

PSC Problème de Satisfaction de Contraintes

QCM Questions à choix Multiples

OEP Optimisation par Essaims Particulaires

ACO Ant Colony Optimization

ADN Acide désoxyribonucléique

PCR Polymerase Chain Reaction

ADL Advanced Distributed Learning

RDF (Resource Description Framework)

W3C World Wide Web Consortium

Introduction générale

1. Contexte général du travail

Ce travail de recherche concerne les EIAH (Environnements Informatiques pour l'Apprentissage Humain) et plus particulièrement la personnalisation de l'apprentissage, cette tâche qui a comme spécificité d'être complexe et difficile de mettre en œuvre est réalisée au sein de notre système en utilisant la méthode du calcul de l'ADN. En effet, plusieurs méthodes ont été utilisées afin d'assurer la tâche de la personnalisation de l'apprentissage, on note parmi eux les essais particuliers, et les algorithmes génétiques. Dans les EIAH les résultats de la personnalisation de l'apprentissage doivent être adéquats pour chaque apprenant afin d'assurer un apprentissage adaptatif et individuel. Dans le but de proposer pour chaque apprenant l'enchaînement adéquat des cours, il faut que celui-ci soit conforme à certains aspects de l'apprenant. Ces aspects là représentent les facteurs psycho-cognitif, le style d'apprentissage, etc. Il existe dans la littérature plusieurs façons pour modéliser un apprenant, on note :

- Les connaissances : qui représentent toutes les informations qui impliquent un domaine particulier et que l'apprenant connaît. On peut distinguer les connaissances procédurales et les connaissances.
- Les préférences : qui représentent l'ensemble des éléments préférés par l'apprenant par exemple : une image préférée, une vidéo, un texte préféré. . .

2. Problématique et approches

2.1 Problématique

La notion de la personnalisation de l'apprentissage fait l'objet de plusieurs travaux de recherche de plusieurs communautés scientifiques; en effet il s'agit d'un problème

pluridisciplinaire qui fait inclure des recherches fondamentales (pédagogiques) et des recherches appliquées (techniques). Cependant du point de vue machine il s'agit de modéliser, optimiser, et planifier des chemins individuels en émergeant les caractéristiques humaines des apprenants et la complexité du phénomène d'apprentissage. La problématique de notre thèse concerne la personnalisation des objets d'apprentissage constituant les cours proposés. Qu'est le chemin optimal à proposer à un apprenant donné ? Qu'elles sont les différentes caractéristiques de l'apprenant qu'il faut prendre en considération pour construire son profil ? Qu'elle est la méthode utilisée pour réaliser ce travail ?

2.2 Approche

Face au grand nombre d'apprenants, possédant chacun un profil différent, et au grand nombre de cours qui sont enregistrés et mis-à-jour, la personnalisation de l'apprentissage devient indispensable dans chaque système d'apprentissage. La difficulté à laquelle est confrontée la personnalisation qui est le sujet de notre thèse est de pouvoir proposer à chaque apprenant le chemin optimal qui pourra l'amener à atteindre son objectif d'apprentissage. Cette difficulté réside plus particulièrement dans le fait de pouvoir résoudre un problème NP complet. On suppose qu'on a un graphe basé cours, que les nœuds de ce graphe sont les coûts et les valeurs des arcs sont la moyenne des taux de maîtrise de chaque cours pour un apprenant donné. Vu les différents profils des apprenants, donc chaque apprenant possède un parcours individuel durant son apprentissage. On suppose qu'on a un grand nombre de cours, l'apprenant est confrontés à eux, donc pour améliorer ses connaissances dans un d'eux, il a de fortes chances à subir soit une surcharge cognitive et/ou une désorientation, ce qui mène par la suite à une démotivation de la part de cet apprenant. Ce problème d'optimisation peut être résolu en appliquant la méthode du calcul de l'ADN. Face au problème de personnalisation de l'apprentissage, plusieurs systèmes ont été développés afin de pouvoir optimiser le chemin d'apprentissage pour les apprenants. Ces systèmes ont utilisés pas mal de techniques bio-inspirés. En effet pour atteindre nos objectifs fixés, nous avons utilisé la méthode du calcul de l'ADN afin de résoudre ce problème. L'architecture principale de notre système est composée d'un module profil apprenant, et un répertoire d'objets pédagogiques. Plusieurs outils ont été réalisés afin de proposer des séquences de cours adaptées aux besoins spécifiques de chaque apprenant. C'est dans ce cadre que nous proposons un outil de personnalisation de l'apprentissage afin d'aider les apprenants dans leurs parcours d'apprentissage via le système

en leur proposant des séquences personnalisées de cours selon leurs caractéristiques intrinsèques. L'outil pourra s'intégrer par la suite, comme une composante faisant partie d'un processus d'adaptation dans un EIAH donné.

2.3 Objectifs

- Rendre l'apprenant plus motivé et engagé dans le processus d'apprentissage.
- Améliorer l'interactivité de l'apprenant.
- guider l'apprenant dans son parcours d'apprentissage.

3. Organisation du manuscrit

Ce mémoire est structuré en trois parties. La première partie est réservée à l'état de l'art de la thématique considérée. La deuxième partie traite notre contribution pour la granularité des contenus. La troisième partie est consacrée à la conception et à la réalisation d'un système qui met en pratique l'approche proposée dans la deuxième partie.

L'ensemble de ce manuscrit peut être lu de façon indépendante. Nous notons cependant que les trois premiers chapitres présentent des concepts et des définitions qui peuvent faciliter la compréhension de la démarche que nous avons poursuivie.

–Chapitre 1 : Présente brièvement l'état d'art de la personnalisation de l'apprentissage dans les (Environnement Informatiques pour l'apprentissage Humain). Plusieurs aspects sont considérés, à savoir : les définitions de la personnalisation et de l'individualisation ainsi que la différence entre eux, le rôle des différents acteurs dans le processus de personnalisation.

–Chapitre 2 : S'intéresse aux objets pédagogiques dans le processus d'apprentissage. Ainsi les différentes définitions des objets sont établies, ensuite un tour sur les technologies associées aux objets pédagogiques est réalisé. Plusieurs standards et normes pour les objets pédagogiques sont discutés.

–Chapitre 3 : Dans le chapitre 3, nous allons présenter les fondements théoriques des algorithmes évolutionnaires. Ainsi nous allons discerner les techniques existantes qui ont été

utilisées pour faire de la personnalisation de l'apprentissage. À la fin de ce chapitre une discussion des différents points fort et points faible de chaque approche.

– Chapitre 4 : Dans le chapitre 4, nous allons étaler et expliquer les outils du calcul de l'ADN avec l'explication des moyens biologiques qui les entourent, ainsi que les problèmes d'optimisation combinatoires qui ont été résolus avec le calcul de l'ADN. Une synthèse va être abordée à la fin.

– Chapitre 5 : Dans ce chapitre, nous abordons notre contribution par rapport au sujet de thèse en commençant par décrire un modèle général de notre méthode proposée, qui est basée essentiellement sur des graphes du modèle du domaine et graphe du modèle de l'apprenant.

– Chapitre 5 : Ensuite dans ce chapitre, nous allons détailler en plus les étapes de notre algorithme issu de la méthode du calcul de l'ADN proposée. Ces étapes engendrent des résultats qui vont être présentés dans des tables de résultats. Juste après on présente l'architecture client/serveur de notre système. Le fondement de l'architecture s'appuie sur une représentation d'un cours sous la forme d'un objet pédagogique intégrant la méthode du calcul de l'ADN. Il s'agit de concevoir un système ayant pour rôle d'individualiser l'apprentissage pour chaque apprenant en utilisant le modèle de l'apprenant, et le modèle du domaine. Aussi nous présenterons nos résultats expérimentaux du système réalisé PERSO.

En guise de conclusion, nous présentons un bilan de nos travaux de thèse, les résultats obtenus et nous traçons des perspectives futures à nos recherches qui vont nous permettre d'améliorer ce qui est proposé. Nous évoquons, également, notre contribution dans l'avancement de la compréhension de notre champ de recherche.

Chapitre 1

La personnalisation de l'apprentissage dans les EIAH

1.1 Introduction

Avec l'apparition des Nouvelles Technologies de l'information et des communications (NTICS), la notion de la personnalisation et de l'individualisation de l'apprentissage a été de plus en plus utilisée.

Dans un processus d'apprentissage, la tâche de la personnalisation est incontournable, car son fondement est en partie historique et social au sens où elle reflète d'une manière significative les aspirations des apprenants et leur intérêt croissant pour l'apprentissage. Cette personnalisation est considérée comme une tâche complexe et longue à mettre en œuvre, elle constitue un élément majeur dans un grand nombre d'activités pédagogiques. La personnalisation de l'apprentissage dans un Environnement Informatique pour l'Apprentissage Humain (EIAH) nécessite le rassemblement de plusieurs facteurs psychopédagogiques et peut être interprétée de bien des manières tout en soulevant des questions profondes quant aux objectifs et aux possibilités de l'enseignement.

Durant ce chapitre de l'état de l'art nous allons aborder la notion de la personnalisation et l'individualisation de l'apprentissage, qu'est ce que ces deux notions ? Pourquoi individualiser et personnaliser et pour qui ? Qu'elle est la relation entre la personnalisation et le processus de l'apprentissage ? Que peut-on personnaliser dans un EIAH ?

1.2 Qu'est ce que personnaliser dans un processus d'apprentissage

1.2.1 Définitions de la Personnalisation

Dans le dictionnaire personnaliser signifie « adapter quelque chose à quelqu'un ». Selon [Lefevre 2009] et dans le cadre de l'apprentissage humain, la personnalisation consiste à

adapter aux spécificités de chaque apprenant le contenu, mais aussi la manière de proposer un apprentissage. Cet apprentissage peut se faire dans le contexte classique de la classe ou à distance en autonomie ou sous la supervision d'un tuteur.

Selon ces définitions on peut dire que la personnalisation de l'apprentissage est le fait de proposer un cours sur mesure à partir d'un ensemble de propriétés de l'apprenant. Les propriétés peuvent être :

- Les connaissances de l'apprenant
 - Les préférences de l'apprenant
 - Le style cognitif
 - Les centres d'intérêts
 - Les besoins de l'apprenant
- La personnalisation est sur plusieurs idées
- La découverte des connaissances de l'apprenant sur le contenu à apprendre.
 - Guider les apprenants à prendre des décisions académiques.

1.2.2 Définition de l'individualisation

Le terme individualisation vient du mot « individu » qui est défini dans le dictionnaire comme « une entité autonome qui ne peut être ni partagée ni divisée sans perdre les caractéristiques qui lui sont propres ». L'individualisation dans le cadre de l'apprentissage humain signifie prendre en compte les caractéristiques individuel de chaque apprenant afin de le guider dans son parcours d'apprentissage selon ses besoins et ses objectifs.

Nous reprendrons volontiers ici la conception de l'individualisation développée par Meirieu [Meirieu 1994]. Il nous rappelle que le concept d'individualisation est relativement ancien (les premières expériences datent des années 20) et qu'il était au départ essentiellement basé sur le souci de respecter le rythme d'apprentissage des apprenants. La notion actuelle d'individualisation met l'accent sur la diversité des moyens d'apprendre adaptés à chacun. Il s'agit aujourd'hui de différencier, d'individualiser, toutes les dimensions de l'acte d'apprendre, c'est-à-dire de proposer des situations d'apprentissage/enseignement tantôt individuelles et tantôt collectives, de mettre en œuvre des méthodes pédagogiques diversifiées et de mettre à

disposition des outils eux aussi variés qui privilégient tantôt l'écrit, tantôt l'oral et qui exploitent les potentialités des TIC.

Avec les TIC, on dispose aujourd'hui d'une palette de moyens d'action de formation/enseignement/apprentissage particulièrement riche qui s'appuie à la fois sur des outils de formation et des outils de communication.

Dans le domaine des langues, de nombreuses expériences sont menées qui explorent et exploitent à des fins d'apprentissage les potentialités des outils de communication en ligne (les réseaux dont bien sûr Internet), qu'ils soient synchrones comme la visioconférence, le bavardage (*chat*, en anglais) ou asynchrones comme le courrier électronique, les forums. Du côté des produits de formation, on dispose soit de produits pédagogiques hors-ligne sur cédérom (ou sur disquettes) ou de produits en ligne sur des sites pédagogiques.

Cette variété de moyens permet d'imaginer toutes sortes d'environnements d'apprentissage, de dispositifs de formation/enseignement "sur mesure" au sens propre du terme, pour répondre aux attentes et aux besoins des apprenants, donnant l'opportunité à chacun de mettre en œuvre ses propres stratégies d'apprentissage.

1.2.3 La Différence entre individualisation et personnalisation

[Rodet 2009] affirme que l'individualisation est du « prêt à porter », et la personnalisation est du « sur mesure », à partir de cela qu'on peut déduire que la personnalisation est plus spécifique que l'individualisation.

TAB.1.1– Comparaison entre la personnalisation et l'individualisation.

Individualisation	Personnalisation
Les mêmes objectifs pour tous les apprenants	Objectifs différents pour chaque apprenant
Application de différentes stratégies didactiques pour accomplir la compétence	Application de différentes stratégies didactiques Pour promouvoir le potentiel personnel

Le programme d'enseignement est défini par le personnel éducatif	L'apprenant participe activement à la construction de son propre programme
La valorisation de la dimension cognitive de l'apprenant	La valorisation de toutes les dimensions de l'apprenant (émotionnel, social, expérience de vie, etc.)
La valorisation des connaissances préalables et des compétences formelles et non formelle	La valorisation des connaissances, compétences, vie et compétence de travail aussi informelle
L'auto direction de l'apprenant est une compétence accessoire	L'auto direction de l'apprenant est une compétence fondamentale
L'enseignant joue un rôle clé	Le tuteur joue un rôle clé

À la naissance de ces deux concepts dans le domaine de l'apprentissage et de l'enseignement, deux autres termes connexes font leur apparition, ces termes sont l'apprentissage personnalisé et l'apprentissage individualisé.

1.3 Apprentissage personnalisé

Selon [Bentley et al. 2004] l'apprentissage personnalisé peut aider à répondre aux futurs besoins éducatifs et offrir de nouveaux moyens de favoriser la capacité d'apprentissage de chaque apprenant. L'apprentissage personnalisé prend en considération la compétence, la motivation, l'adaptabilité et l'activité de l'apprenant.

Personnaliser l'apprentissage permet d'encourager les apprenants à développer des buts clairs et attendus pour les accomplir afin de prendre de bonnes décisions dans un environnement rigoureux.

1.3.1 Phases de l'apprentissage personnalisé

La théorie de la personnalisation pousse les éducateurs à penser en dehors de la boîte en se focalisant sur les besoins des apprenants pour qu'ils soient impliqués dans la conception de leurs propres processus d'apprentissage [Campbell et al. 2007]

C'est pour cela l'apprentissage personnalisé passe par plusieurs phases dont on peut citer :

- Phase d'évaluation : l'enseignant et l'apprenant travaillent dans une manière formative pour identifier les faiblesses et les forces des apprenants.
- Phase d'enseignement et d'apprentissage : les enseignants et les apprenants sélectionnent les stratégies d'apprentissage.
- Phase du choix du Curriculum : les apprenants choisissent le curriculum, la création du chemin selon l'apprenant.
- Phase de départ radical des modèles typiques : construite sur le progrès de l'apprenant, cette phase fournit aux enseignants de la flexibilité pour choisir leurs propres stratégies d'enseignement.
- La phase de l'éducation au delà de la salle de classe : en utilisant les connections sociale et communautaires, les apprenants personnalisent leur environnement (avec l'aide de l'enseignant, si besoin) pour créer leur idéal environnement d'apprentissage.

On note aussi que l'apprentissage personnalisé est constitué de plusieurs composantes.

1.3.2 Composantes de l'apprentissage personnalisé

Le but de l'apprentissage personnalisé est d'élever les niveaux en centrant l'enseignement et l'apprentissage sur les aptitudes et les intérêts des apprenants.

- L'évaluation pour l'apprentissage et l'utilisation des données et du dialogue pour diagnostiquer les besoins d'apprentissage de chaque apprenant.
- Exigence de développer chez l'apprenant la compétence et la confiance en soi au moyen de stratégies d'enseignement et d'apprentissage qui se construisent à partir des besoins individuels
- Le choix des programmes scolaires sollicite les apprenants tout en les respectant.
- L'apprentissage personnalisé appelle une approche radicale de l'organisation scolaire

- L'apprentissage personnalisé signifie que la collectivité, les institutions locales et les services sociaux qui soutiennent l'école.

1.4 Apprentissage individualisé

L'apprentissage individualisé, ou l'éducation individualisée, est une méthode d'enseignement dans laquelle le contenu, la technologie, et les parties d'apprentissage sont basés sur les capacités et les intérêts des apprenants.

1.4.1 Les étapes de l'apprentissage individualisé

L'apprentissage individualisé est considéré comme très proche de l'apprentissage personnalisé, en effet, les buts de l'apprentissage personnalisé sont basés sur le design des curriculums et des standards qui peut être le même pour tous les apprenants, mais pour l'apprentissage individualisé, le profil et le plan pour chaque apprenant peut varier. Parce que chaque apprenant progresse à travers le matériel de l'apprentissage à des vitesses différentes, par rapport a ses besoins et capacités d'apprentissage.

Par exemple, un apprenant peut prendre beaucoup plus de temps pour progresser à travers un sujet donné, passe des sujets qui couvrent des informations déjà connues, ou bien répète des sujets dans lesquels il a plus besoin.

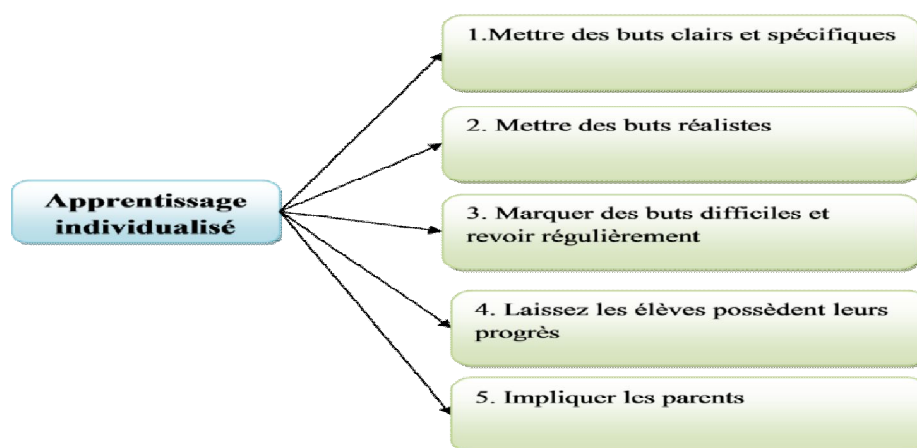


FIG.1.1– Etapes de l'apprentissage individualisé.

1.5 A quoi sert la personnalisation et pour qui ?

Il y'a deux raisons pour personnaliser : (1) encourager les apprenants à développer des buts clairs et attendues pour les accomplir afin de prendre de bonnes décisions dans un environnement rigoureux. (2) guider les apprenants à prendre des décisions académiques saines. La personnalisation est au cœur de la profession de l'enseignant qui a pour rôle d'entretenir le désir d'apprentissage pour les apprenants et de les aider à acquérir les capacités et la confiance nécessaire pour qu'il prennent eux même l'initiative de s'instruire.

TAB.1.2. Les différents acteurs dans la personnalisation.

Apprenant	Enseignant	Institution
Accroît l'intérêt que l'apprenant porte aux activités d'apprentissage.	Répondre aux besoins diversifiés des apprenants Motiver les apprenants à accomplir l'apprentissage	Améliorer la qualité d'apprentissage Juger l'efficacité d'apprentissage
Egaliser les possibilités d'apprendre	Faire preuve de solide esprit de communication et de collaboration	

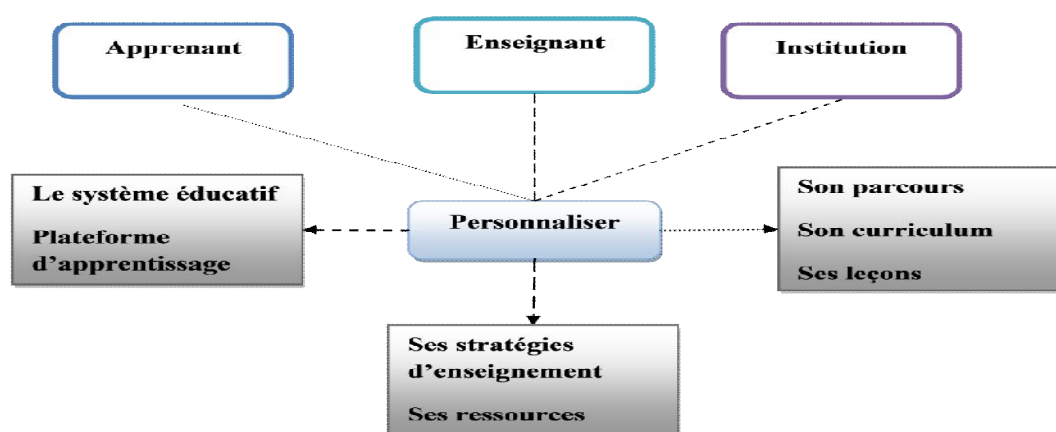


FIG.1.2– Rôle de la personnalisation selon les acteurs.

1.6 Que peut-on personnaliser dans un EIAH ?

La notion d'EIAH couvre une diversité de systèmes : *systèmes hypermédias*, *systèmes tuteurs intelligents*, *plates-formes de formation ouverte et à distance*, etc. C'est un domaine pluridisciplinaire qui allie les *sciences humaines*, la *psychologie cognitive*, les *sciences de l'éducation*, la *didactique*, l'*ergonomie*, l'*informatique* et l'*intelligence artificielle*. La conception de tels systèmes nécessite de prendre en compte une multitude de problématiques inter-reliées [Tchounikine 2002]. Le paramétrage d'un EIAH « concerne la possibilité de contrôler, de modifier à son gré les caractéristiques du logiciel, non seulement pour l'adapter à des situations d'apprentissage particulières, mais aussi pour en adapter le contenu, l'agencement ou le comportement pédagogique » [Hû 2001]. La personnalisation d'un EIAH peut donc porter sur plusieurs facettes de l'EIAH. Ces facettes regroupent les hypermédias adaptatifs et les tuteurs intelligents.

- **Les hypermédias adaptatifs**

[Brusilovsky 1999] présente l'ensemble des technologies utilisées pour adapter ces systèmes. Cet ensemble se focalise principalement sur les technologies et permet, à travers les éléments cibles de l'adaptation, de connaître les éléments personnalisables dans les systèmes étudiés. L'auteur regroupe les technologies permettant l'adaptation en deux grandes catégories : des technologies agissant sur le parcours des apprenants dans les systèmes et des technologies aidant à la résolution de problèmes.

Les technologies agissant sur le parcours des apprenants permettent : d'adapter l'ordre des leçons et des exercices à faire pour chaque apprenant, comme dans SIETTE [Ríos et al. 1999] ou ILESA [López et al. 1998] ; de guider l'apprenant en lui indiquant quel élément de la formation serait à faire en fonction de ses connaissances actuelles et de son but d'apprentissage, comme dans INTERBOOK [Brusilovsky et al. 1998] ou ELM-ART [Brusilovsky et al. 1996] ; de masquer ou rendre inactifs des liens qui mènent vers des informations que l'apprenant ne pourra pas encore comprendre, comme dans REMEDIAL MULTIMEDIA SYSTEM [Anjaneyulu 1997]; et de présenter la même information différemment selon l'apprenant en détaillant, masquant, résumant ou illustrant des parties de

la formation, comme dans PT [Kay et al. 1997], AHA [De Bra et al. 1998] ou METALINKS [Murray et al. 1998][Murray 2003].

Les technologies de cette catégorie agissent sur l'organisation des ressources pédagogiques, mais également sur le choix des ressources pédagogiques elles-mêmes, et sur l'interface du logiciel.

- **Les tuteurs intelligents**

Cependant, les technologies concernant l'aide à la résolution de problèmes permettent : de corriger et d'aider l'apprenant au cours de la résolution, comme dans LISP-TUTOR [Anderson et al. 1985] ; de déterminer, à partir de l'analyse des réponses de l'apprenant (si une réponse finale est correcte/incorrecte, si les connaissances utilisées sont correctes, etc.), les rétroactions à fournir et de mettre à jour le modèle de l'apprenant, comme dans PROUST [Johnson 1986] ; de fournir un exemple de résolution expliqué ou des problèmes résolus du même type, comme dans ELM-PE [Weber 1996] ou ELM-ART [Brusilovsky et al. 1996].

Les technologies de cette catégorie agissent sur les fonctionnalités du logiciel (aide, diagnostic, etc.) et permettent dans certains cas de modifier les rétroactions à fournir à l'apprenant.

En complément de cet ensemble centré sur les tuteurs intelligents et les hypermédias adaptatifs, on peut dire que la personnalisation peut porter sur :

- les **fonctionnalités**, en définissant celles disponibles (par exemple autoriser l'accès à l'aide, au diagnostic, aux compagnons, aux outils de dessins, de calculs) et selon quelle temporalité (constamment, à certains moments, avec un nombre d'accès autorisés, etc.) ;
- les **objets pédagogiques** proposés aux apprenants, en changeant l'accessibilité et le contenu des messages (adapter le vocabulaire ou la présentation de ceux-ci) ;
- l'**interface** du système, en choisissant la langue du système, en paramétrant la police (taille, style, etc.), en adaptant les couleurs (pour les dyslexiques par exemple), en activant la synthèse vocale, etc.

1.7 Qui personnalise l'EIAH ?

Nous venons de voir ce qui pouvait être personnalisé dans un EIAH. Nous nous intéressons à présent aux différents acteurs pouvant mettre en place une personnalisation : l'apprenant lui-même, l'enseignant et le module pédagogique d'un EIAH. Pour chaque type d'acteur, nous donnons des exemples avant de résumer les possibilités offertes, les forces, ainsi que les faiblesses de la personnalisation par chaque acteur donné.

1.7.1 Personnalisation par l'apprenant

Le principal acteur de l'apprentissage est l'apprenant lui-même. C'est pourquoi certains systèmes guident l'apprenant dans le choix des activités qu'il va effectuer, voire le laissent choisir seul ces activités [Lefevre 2009].

Un exemple de système permettant à l'apprenant de choisir précisément l'activité qu'il va effectuer est le système ANDES [VanLehn et al. 2005, Andes 2009], où l'apprenant choisit un thème, puis sélectionne un problème, parmi ceux associés à ce thème, d'après son énoncé. Sur la figure 1.3, l'écran de gauche contient la liste des thèmes proposés à l'apprenant et l'écran de droite contient la liste des problèmes associés au thème « Vectors ». Sur ce deuxième écran, l'apprenant peut sélectionner un problème pour avoir un aperçu de son énoncé. [Lefevre 2009].

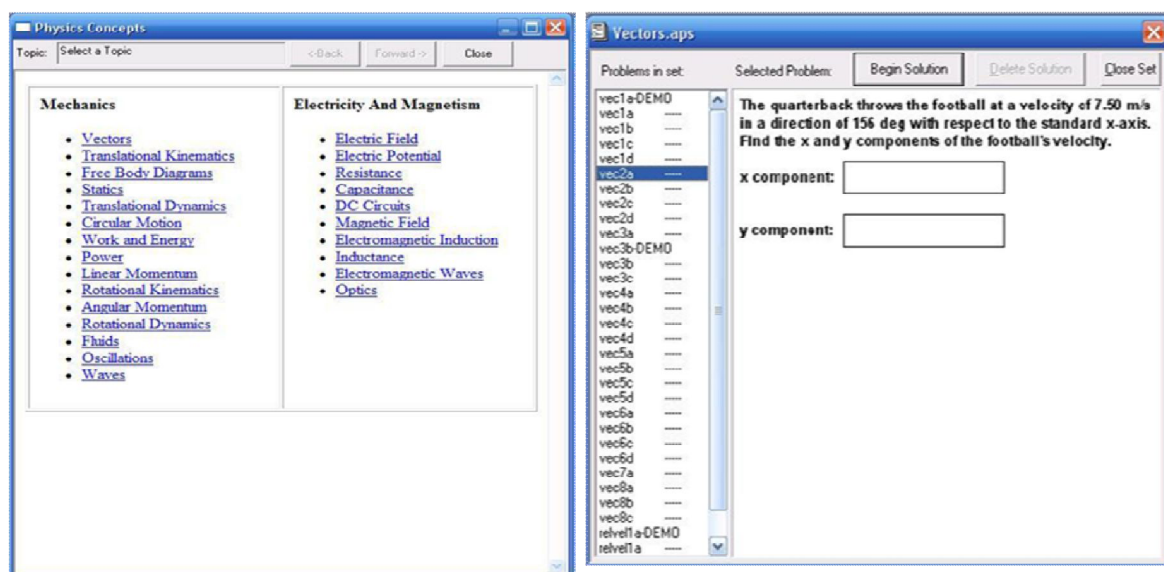


FIG.1.3– Écrans d'ANDES permettant à l'apprenant de sélectionner le problème qu'il va résoudre [VanLehn et al. 2005].

D'autres systèmes ne laissent à l'apprenant que la possibilité d'un choix partiel concernant les activités. Ainsi, dans les systèmes APLUSIX [Bouhineau et al. 2005b, Bouhineau et al. 2005a] ou LILIMATH [LiliMath 2000], les apprenants choisissent un thème sur lequel ils souhaitent travailler (voir figure 1.4) et le système génère une série d'exercices sur ce thème.

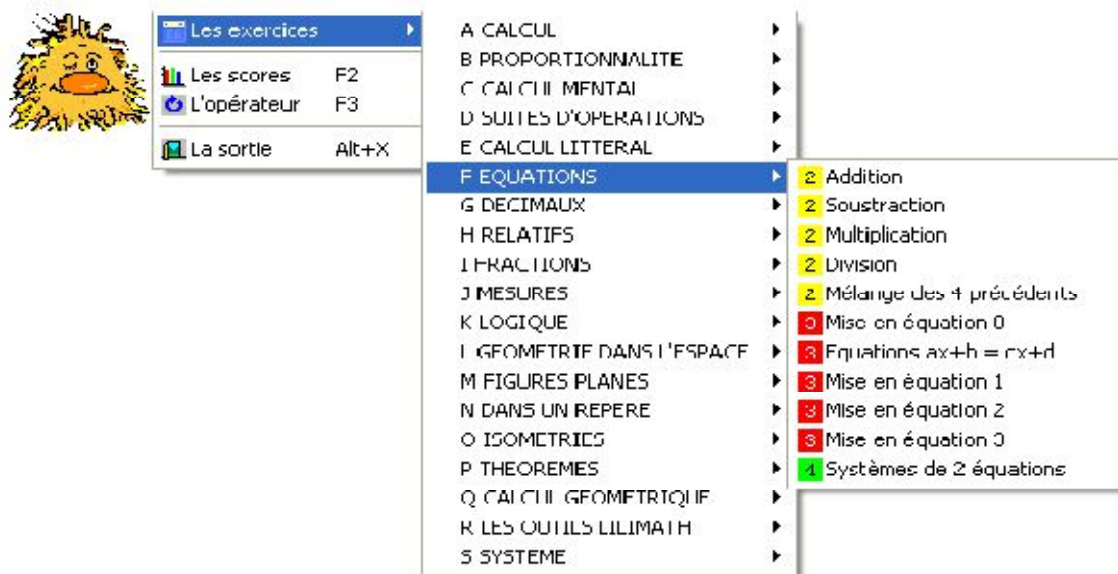


FIG.1.4– Menu de LILIMATH permettant à l'apprenant d'obtenir une série d'exercices en fonction d'un thème [LiliMath 2000].

1.7.2. Personnalisation par l'enseignant

Avant de présenter les approches permettant aux enseignants de personnaliser un EIAH, il est important de rappeler que les enseignants n'ont pas toujours eu un rôle prévu lors de l'utilisation des logiciels pédagogiques [Lefevre 2009]. Le désintérêt des enseignants vis-à-vis des EIAH est accentué par deux facteurs : leur manque d'expérience concernant l'utilisation des nouvelles technologies et le manque d'adéquation entre le contenu pédagogique du logiciel et les attentes des enseignants [Girard et al. 2007]. Suite à ce constat d'échec, les chercheurs en EIAH se sont attachés à créer des EIAH non seulement utiles pour l'apprentissage d'une discipline, mais également intégrables aux méthodes d'apprentissage traditionnelles des enseignants [Lefevre 2009]. Cette démarche nécessite d'attribuer un rôle à l'enseignant au sein des EIAH. Pour cela, plusieurs stratégies sont possibles : intégrer des enseignants lors de la conception des logiciels, développer des outils permettant aux

enseignants de créer eux-mêmes leurs logiciels pédagogiques ou développer une partie des logiciels spécifiquement pour l'enseignant [Jean-Daubias 2004].

La première stratégie, qui consiste à prendre en compte les enseignants lors de la conception, est essentielle, mais elle ne leur permet pas de personnaliser par la suite les logiciels créés, sauf s'ils possèdent une partie destinée à l'enseignant [Lefevre 2009].

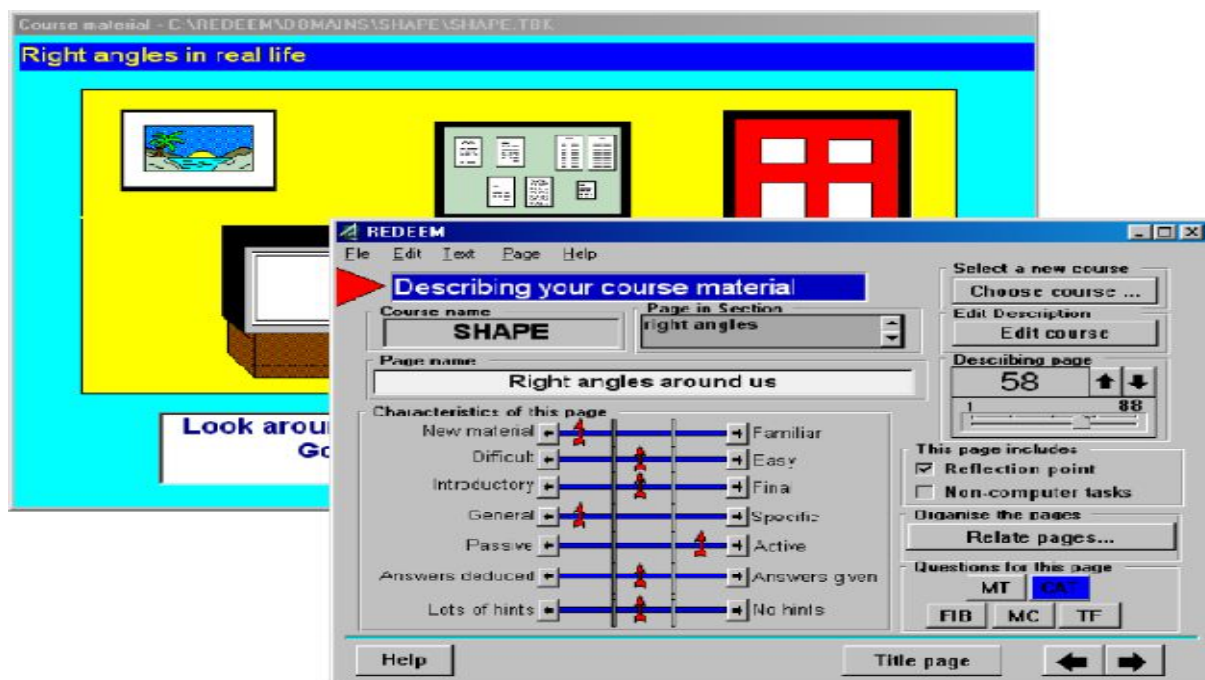


FIG.1.5–Interface de description du contenu pédagogique à utiliser dans REDEEM [Ainsworth 2000].

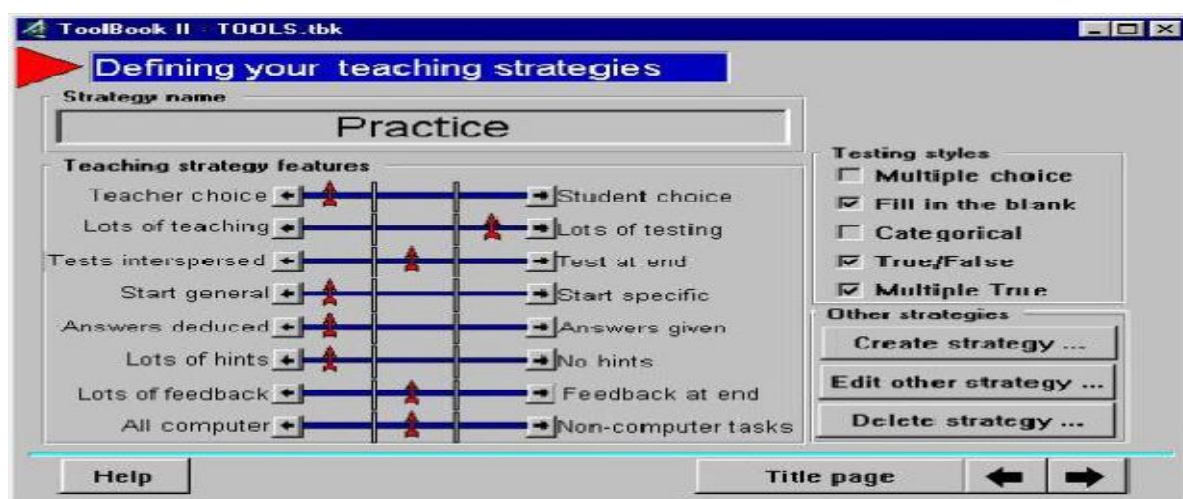


FIG.1.6– Interface de définition d'une stratégie pédagogique dans REDEEM [Ainsworth 2000].

Ainsi, le système REDEEM [Ainsworth 2000] permet aux enseignants d'intégrer le matériel pédagogique créé par un autre outil auteur afin de l'utiliser avec leurs stratégies pédagogiques. Pour cela, un enseignant décrit le matériel pédagogique qu'il souhaite utiliser (voir figure 1.6), forme des groupes avec ses apprenants, puis définit les stratégies pédagogiques qu'il veut mettre en place pour que le matériel pédagogique soit proposé de façon adéquate à chaque groupe d'apprenants (voir figure 1.7).

Le système EON [Murray 2003a] quant à lui, permet à l'enseignant de définir des méta-stratégies afin de choisir la stratégie pédagogique appropriée pour enseigner un type de connaissances. Chaque méta-stratégie est une règle « si-alors » avec pour prémisse une condition d'application et pour conclusion, des valeurs pour les paramètres du système. La figure 1.7 montre un exemple de méta-stratégie définie par un enseignant. Dans cet exemple, si la variable du modèle de l'apprenant « RecentWrong » a une valeur supérieure à 50%, c'est-à-dire si l'apprenant se trompe une fois sur deux, alors le nombre de conseils sera de deux, le niveau de difficulté sera croissant, etc.

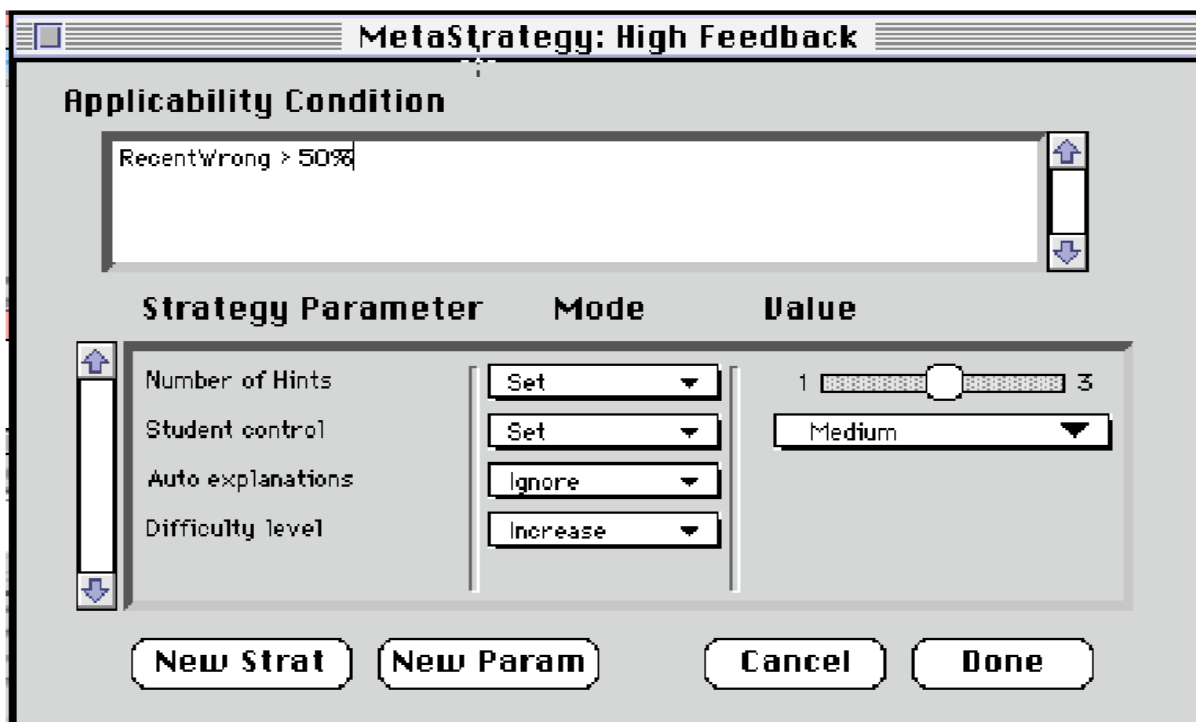


FIG.1.7– Éditeur de méta-stratégies dans EON [Murray 2003a].

La troisième stratégie pour attribuer un rôle à l'enseignant au sein des EIAH est de leur réserver une partie de l'EIAH. Nous allons à présent voir les différentes possibilités offertes aux enseignants pour adapter les EIAH possédant une partie qui leur est consacrée.

Une première possibilité réside dans la personnalisation des buts à atteindre par les apprenants. Dans DIVIDINGQUEST [Girard et al. 2007], l'enseignant fournit pour chaque apprenant son niveau habituel concernant l'apprentissage des divisions. Le système classe, à partir de cette information, l'apprenant dans un groupe auquel correspondent des buts pédagogiques à atteindre. Ces buts pédagogiques peuvent être modifiés par l'enseignant.

Une seconde possibilité consiste à permettre à l'enseignant d'agir directement sur les activités fournies aux apprenants et sur l'ordre dans lequel elles sont proposées aux apprenants. Ainsi, dans APLUSIX [Nicaud et al. 2003], l'enseignant peut créer des exercices, pour ensuite demander à ses élèves de les faire. Dans AMBRE-ENSEIGNANT [Duclosson et al. 2005, Jean-Daubias et al. 2009], l'enseignant crée des problèmes semi-automatiquement, c'est-à-dire en contraignant la génération des problèmes, puis définit, pour chaque apprenant, des séquences contenant les exercices qu'il souhaite que l'apprenant résolve et leur ordre. Enfin, dans ROBOTTEACH [Leroux 2002, Leroux 2006], l'enseignant définit une suite d'activités à faire grâce à une interface de gestion de séquences pédagogiques. Il peut compléter cette suite d'activités de créneaux où l'enseignant devra être appelé par l'apprenant pour interagir sur les activités passées ou à venir.

1.7.3. Personnalisation par le système

La personnalisation d'un EIAH peut être faite automatiquement par celui-ci afin que son contenu soit adapté à chaque apprenant. Cette personnalisation peut se faire de plusieurs manières, mais exploite toujours les informations que le système possède sur l'apprenant.

Les informations sur l'apprenant peuvent être représentées par des stéréotypes auxquels les apprenants sont associés [Girard et al. 2007] ou dans un modèle de l'apprenant contenu dans l'EIAH. Le modèle de l'apprenant peut être créé par l'apprenant lui-même, comme dans CREEK-TUTOR [Sørmo et al. 2002, Aamodt 2005], où l'apprenant dessine la carte des concepts qu'il possède sur un sujet particulier, ou calculé par le système, comme dans SQL-TUTOR [Mitrovic 1998].

Lorsqu'il est calculé par le système, le modèle de l'apprenant peut contenir plusieurs types d'informations. Tout d'abord, les informations peuvent être, tout simplement, les réponses de l'apprenant. Ainsi, le système IDEBUGGY [Burton 1982] génère un nouveau problème après

chaque erreur de l'apprenant. De façon analogue, APPELEC [Guéraud et al. 2007] contient 22 activités dont l'enchaînement est géré en fonction des réponses fournies progressivement par l'apprenant. Un autre type d'informations sur l'apprenant peut être le comportement de celui-ci lorsqu'il interagit avec le logiciel. Ainsi, le didacticiel ESCARRE [Hofmann 1990] s'adapte aux réactions de l'apprenant. De même, le simulateur TELEOS [Teleos 2003], [Mufti-Alchawafa et al. 2004] interprète le comportement de l'apprenant pendant le processus de résolution des problèmes pour lui fournir un feedback approprié. Enfin, les informations sur l'apprenant peuvent être relatives à ses connaissances ou compétences. Ainsi, le logiciel MOREMATH [Bull et al. 2003] construit un modèle des connaissances de l'apprenant pour, entre autres, lui fournir une séquence de révision adaptée (contenant des cours et des exercices). Les modèles de l'apprenant peuvent contenir d'autres informations (comme les préférences de l'apprenant), nous avons présenté uniquement les trois types d'informations les plus couramment utilisées.

1.7.4 Synthèse sur les acteurs de la personnalisation

Nous avons indiqué les avantages et les inconvénients de la personnalisation par chacun des acteurs. Tout d'abord, lorsque l'apprenant personnalise lui-même un EIAH, cela lui permet de se l'approprier pleinement et de choisir les thèmes qu'il souhaite aborder. Il peut ainsi, selon les cas, choisir un cours à suivre, un exercice à résoudre ou un thème à aborder. L'inconvénient de cette approche est qu'elle est difficilement intégrable dans le cadre d'un enseignement en classe. Un enseignant ne pourra pas contrôler ce que fait chaque apprenant, et ainsi être sûr que celui-ci fera les activités nécessaires pour maîtriser une notion qu'il souhaite lui faire acquérir.

À l'inverse, lorsque l'enseignant personnalise les EIAH avant qu'ils soient utilisés par l'apprenant, il peut maîtriser le contenu du logiciel pour chaque élève ou groupe d'élèves. Le problème est que cette personnalisation n'est possible que sur un nombre restreint de systèmes. De plus, dans la majorité des EIAH concernés, aucune aide n'est fournie à l'enseignant sur la personnalisation à mettre en place pour un élève donné, en exploitant par exemple les réponses de l'apprenant ou un modèle des connaissances de celui-ci. Enfin, lorsque la personnalisation de l'EIAH est faite automatiquement par le système, elle l'est en fonction de la connaissance que le système a de l'apprenant, mais elle n'est pas forcément adaptée aux buts pédagogiques des enseignants.

1.8 Conclusion

La personnalisation automatisée dans les EIAH est un ensemble de méthodes et de techniques qui utilisent les différentes technologies de l'information et de la communication. Ces techniques sont principalement plus efficaces que celle du tuteur humain. Le processus de personnalisation, en plus de pouvoir répondre aux besoins et aux intérêts des apprenants, leur apprend à gérer leur propre apprentissage et prendre le contrôle et la propriété de celui-ci. Ce n'est pas quelque chose qui est fait pour eux seulement, mais quelque chose qu'ils participent à faire pour eux-mêmes. Pour les enseignants, l'apprentissage personnalisé est sur la facilitation plus de la diffusion.

Chapitre 2

L'objet pédagogique et l'apprentissage

2.1 Introduction

Les objets pédagogiques, objets d'apprentissage ou 'Learning Objects' sont actuellement au cœur de nombreuses applications de l'ingénierie pédagogique. Les travaux récents concernant les objets pédagogiques s'intéressent à la standardisation de ces derniers en s'appuyant sur les métadonnées pour décrire leurs contenus afin d'assurer une interopérabilité et une réutilisation de ceux-ci. D'une manière générale, cette standardisation se base sur différentes orientations de recherche, soit décrire l'objet pédagogique comme des entités dont le système dispose et dont les manipulations se basent sur la spécification des métadonnées, soit s'orienter vers des langages de modélisation pédagogique pour représenter les unités d'apprentissage. En effet, il nous semble devoir encore faire un état des lieux sur les travaux effectués dans le cadre de la spécification des objets pédagogiques. Nous allons essayer dans ce chapitre d'exposer les différentes notions sur l'objet pédagogique à savoir les définitions, les classes ainsi que les différentes technologies associées.

2.2 Objets Pédagogiques, métaphores et esquisse de définition

Le terme objet pédagogique ou Learning object (LO) en anglais, est exploité, aujourd'hui par la plupart des acteurs du domaine de 'E-Learning' pour les anglo-saxons et 'EIAH' pour les francophones mais avec différents sens. Un accord sur la définition de ce concept n'est pas encore établi. Avant de citer les différentes définitions de l'objet pédagogique, on donnera dans un premier temps les trois métaphores exploitées par les chercheurs en 'Learning design'. La métaphore la plus courante est celle du jeu de LEGO. Ce qui est sous-entendu par cette métaphore, c'est que chaque granule (brique) est combinable avec n'importe quel autre

granule, de n'importe quelle façon. Combiner les différentes granules est une activité intuitive. La métaphore du jeu de LEGO est courante mais un peu simpliste pour représenter la complexité de l'assemblage des objets pédagogiques.

Une autre métaphore, un peu moins utilisée que la précédente, se base sur la structure de la matière. Les objets pédagogiques sont alors considérés comme des atomes. La différence entre cette métaphore et celle de jeu de LOGO, c'est que chaque atome ne peut pas se combiner avec n'importe quel autre atome de n'importe quelle façon et qu'il faut des relations de liens pour les assembler. Selon Wiley [Wiley 2002], cette métaphore permet de rendre compte du fait que tout objet pédagogique ne peut s'accorder qu'avec un nombre restreint d'autres objets et seulement avec certains types. Cependant d'après [Paquette et al. 2002], cette représentation ne tient pas compte des acteurs qui conçoivent et utilisent ces objets, elle ne rend pas en compte des aspects dynamiques et du cycle de vie des objets pédagogiques. C'est pourquoi, ils préfèrent la notion de corps vivants composés de cellules, structurées en organes.

La troisième métaphore concerne une métaphore théâtrale, il s'agit de considérer l'activité comme étant une pièce de théâtre. L'activité représentée comme étant un objet pédagogique ou unité d'apprentissage se déroule selon un scénario de pièce de théâtre. Il y a différents acteurs qui interagissent avec l'apprenant qui est considéré comme l'acteur principal jouant le premier rôle dans l'activité [Aouag 2008].

- **Définition IEEE**

Plusieurs définitions ont été données par l'équipe de travail IEEE. Les deux définitions proposées, citées par [Bourda 2001] sont les suivantes :

- Un objet pédagogique est défini comme toute entité, numérique ou non, qui peut-être utilisée pour l'enseignement ou l'apprentissage.
- Un objet pédagogique est défini comme toute entité numérique ou non qui peut être utilisée, réutilisée ou référencée pendant des activités d'apprentissage assisté par ordinateur (enseignement intelligent assisté par ordinateur, environnements d'enseignement interactifs, systèmes d'enseignement à distance, environnements d'apprentissage collaboratifs).

Dans la première définition tout est objet pédagogique, la seconde définition désigne les entités utilisées dans le cadre des systèmes d'apprentissage assisté par ordinateur. L'objectif du groupe de travail des IEEE connus sous le nom LTSC (Learning Technology Standards Committee) est de proposer des normes, des procédures et des manuels pour les composants logiciels, des outils et des techniques ainsi que des méthodes de conception qui aident au développement, au déploiement, à la maintenance et à l'interopérabilité des composants et des systèmes pédagogiques informatisés. L'interopérabilité signifie que plusieurs systèmes, qu'ils soient identiques ou radicalement différents, puissent communiquer sans ambiguïté et opérer ensemble.

2.3 Classes d'objets pédagogiques

Selon [Pernin 2003], il existe trois principales classes d'objets pédagogiques : Les Unités d'Apprentissage, les Activités Pédagogiques et Les Ressources Pédagogiques.

2.3.1 Les Unités d'Apprentissage

Elles permettent de structurer la formation et de l'organiser dans l'espace et dans le temps. Le concept des unités d'apprentissage représente un ensemble d'activités organisé autour d'un scénario ayant pour objectif d'entraîner des apprentissages chez les apprenants. Ces activités sont réalisées par des acteurs tenant un rôle précis et se déroulent dans un environnement caractérisé par un ensemble de services, d'outils et de ressources pédagogiques [Aouag 2008]. Une des principales propriétés d'une unité d'apprentissage réside dans le fait qu'elle ne peut être décomposée sans perdre son sens et son efficacité d'apprentissage. En pratique, il peut s'agir de cours, de programmes d'études, de groupe de travail, de travaux pratiques, d'une leçon etc. Le modèle EML intègre des préoccupations pédagogiques en définissant précisément pour chaque unité d'apprentissage les objectifs et les pré-requis pédagogiques.

2.3.2 Les Ressources Pédagogiques

Dans la logique des EMLs, une unité d'apprentissage n'implique pas forcément l'usage d'une ressource pédagogique. Elle peut consister en une série d'activités, comme par exemple des conversations entre les apprenants et leur tuteur, des débats entre apprenants... Une ressource pédagogique multimédia (RPMu) est un contenu, à vocation de formation, incorporé dans une interface utilisateur. La structure physique d'une RPMu est représentée en code informatique de caractère multimédia. Une RPMu contient tout ce dont elle a besoin pour fonctionner à la fois sur le plan pédagogique et sur le plan technique. Elle a donc un fonctionnement autonome et elle est opérationnelle dans un environnement pour lequel elle peut émettre des contraintes ; elle est portable et réutilisable dans différents ensembles de RPMu construits à leur gré par les formateurs. Une RPMu est composite ; elle est composée de différents fichiers multimédias (texte, son, image, animation...) ; elle comporte tous les fichiers nécessaires à son fonctionnement en particulier tous les fichiers de codage de l'interface utilisateur.

Afin d'être intégrée à un dispositif dédié au travail collaboratif en ligne ou dédié à la formation ouverte et à distance, la RPMu devra respecter le standard SCORM 1.2 édité par l'ADL (Advanced Distributed Learning).

2.3.3 Les Activités Pédagogiques

Elles définissent les modalités précises d'acquisition, de validation, de communication d'une ou plusieurs connaissances. Parallèlement à l'approche centrée sur les ressources, de nouveaux modèles de spécification des systèmes d'apprentissage électroniques émergent progressivement, en particulier autour de la notion de Langage de Modélisation Pédagogique.

– **Définition** l'activité pédagogique peut se définir par un objectif pédagogique et une stratégie pédagogique complétés éventuellement par une série de paramètres : temps, lieu de travail de l'apprenant, ressources à utiliser, accompagnement déployé. . . La lecture, l'observation, la découverte, la recherche, l'analyse, la conception, la résolution de problèmes, les exercices, l'évaluation personnelle sont des exemples d'activités pédagogiques. « Le

glossaire de la e formation, Institut de la Gestion Publique et du Développement Economique’, Mars 2006.»

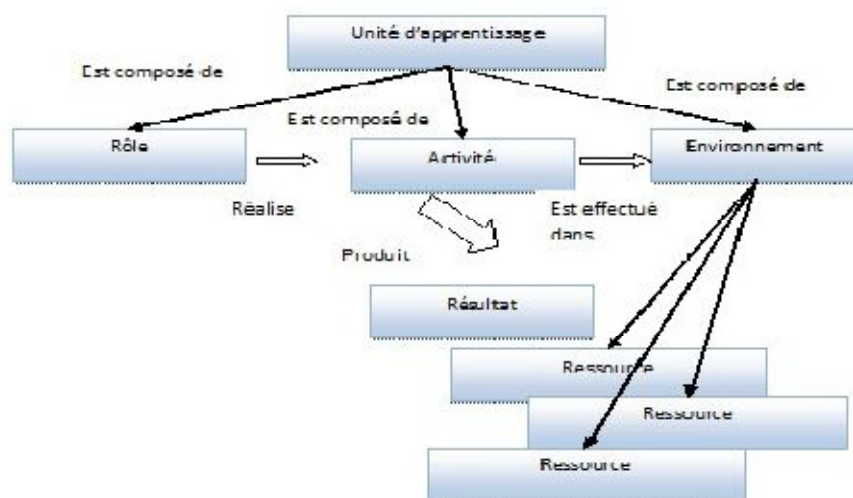


FIG 2.1– Architecture de l'Unité d'apprentissage OUN-EML [Pernin 2003].

Le modèle de base, développé par Koper [Koper 2001], repose sur les concepts fondamentaux d'activités et d'unités d'apprentissage. L'approche EML situe l'activité au cœur de l'apprentissage et non plus les ressources qui représentent seulement des éléments utiles pour le déroulement des activités. Les EML visent à proposer un modèle pour la mise en place de situations d'apprentissage, en s'appuyant sur les activités. Les ressources pédagogiques ne sont pas vues nécessairement comme des objets préexistants, mais peuvent également être le résultat d'une activité, comme par exemple dans les cas où les apprenants doivent réaliser un rapport après une activité de recherche.

2.4 Objet et Objet pédagogique

Une première proposition comportant les éléments pouvant servir à identifier les aspects d'un objet pédagogique est celle de la théorie de Brousseau [Brousseau 1986] ; il s'agit de distinguer entre situation didactique et situation pédagogique. L'idée de base est que les objets pédagogiques seraient partie intégrante du milieu auquel l'apprenant est confronté ; pour qu'un objet puisse jouer un rôle pédagogique, il faut qu'il intègre des intentions didactiques. De nombreuses situations d'apprentissage, peuvent comporter des objets sans intention

didactique, comme par exemple : une ressource documentaire : photo, bouton, vidéo... (Ce que l'on appelle les instruments techniques d'interface) Pour mettre en place une situation d'apprentissage, de tels objets ne peuvent suffire. Le milieu doit inclure des objets pédagogiques correspondant à des intentions didactiques. Il faut distinguer les objets élémentaires qui n'ont pas d'intention didactique, des objets "composites" comportant des objets élémentaires en fonction d'intentions didactiques. Pour être réutilisé, un tel objet "composite" doit être inséré dans un milieu où les ressources et les intentions didactiques sont conciliables, ce qui limite considérablement les possibilités d'adaptation. "Un milieu sans intentions didactiques (c'est-à-dire non volontairement organisé pour enseigner un savoir) est insuffisant à induire chez un sujet toutes les connaissances que la société souhaite qu'il acquière." [Bessot 2002].

L'objet pédagogique est caractérisé par les connaissances mises en jeu lors du processus d'acquisition des connaissances de l'apprenant. Réutilisation, souplesse et mécanismes de décomposition sont subséquentement utilisés pour structurer les connaissances décrivant le déroulement de l'activité didactique, selon Willey [Willey 2002] les objets pédagogiques sont des éléments d'un nouveau type d'instructions basées sur l'ordinateur réunies dans le paradigme de l'orienté objet. Il s'agit de poser ici un double questionnement, d'un côté les considérations techniques pour la structuration de l'activité didactique. D'un autre côté les apports de cette structuration par rapport aux attentes pédagogiques et aux objectifs d'apprentissage afin de pouvoir inclure tous les éléments d'adaptation. Un des points clés dans la réutilisation des objets pédagogiques réside dans leur capacité à s'adapter à des intentions didactiques différentes.

2.5 Technologies associées aux objets pédagogiques

Les éléments liés à la terminologie et à la caractérisation des OPs sont déterminés plus haut, nous présentons dans cette section les technologies émergentes associées au développement des OPs, notamment les *métadonnées*, le *web sémantique* et les *ontologies*.

2.5.1 Les métadonnées

2.5.1.1 Importance des métadonnées

Les métadonnées permettent d'identifier tout type de contenu, et plus spécifiquement des contenus éducatifs, avec un format commun décrivant le type de contenu, son auteur, la meilleure manière de l'utiliser, etc. [Arnaud 2004]. Elles sont généralement stockées dans des bases de données et sont liées aux objets qu'elles décrivent.

Selon Brand, Frank et Meyers [Brand et al. 2003], nous trouvons quatre principales classes. Les *métadonnées descriptives* permettant de décrire une ressource pédagogique pour des fins telles que la recherche et l'identification. Elles peuvent inclure des éléments tels que titre, résumé, l'auteur, mots-clés, etc. Les *métadonnées structurelles* indiquent la composition d'une ressource pédagogique. Par exemple, comment les pages sont ordonnées pour former des chapitres. Les *métadonnées administratives* qui permettent la gestion des ressources pédagogiques par exemple la date de création, type de fichier, accessibilité, etc. Cette catégorie inclut aussi les *métadonnées de préservation* qui permettent l'archivage numériques et les *métadonnées de droit de gestion* qui traitent les droits de propriété intellectuelle, etc. [Niso 2004].

2.5.1.2 Intérêt des métadonnées

Les métadonnées ont un rôle important à jouer dans la sélection et la réutilisation adaptative des objets pédagogiques. Leurs utilisations dans les SHAD, permet de décrire les différents OPs et de les stocker dans des entrepôts pour qu'ils puissent être réutilisés dans différents contenus et partagés entre les différents apprenants des systèmes d'apprentissage. Chaque OP doit être décrit par un fichier de métadonnées qui définit son titre, la langue, description, mot-clé, format, type de ressource d'apprentissage, niveau d'interactivité et de difficulté, etc. Lorsqu'un OP est cherché, ses caractéristiques sont comparées avec les préférences des apprenants, afin que le contenu le plus approprié possible soit livré à l'apprenant.

Niso [Niso 2004] énumère quatre raisons qui justifient la création des métadonnées :

- **La recherche des ressources** : les métadonnées permettent d'identifier les ressources, ressortir l'ensemble des ressources similaires, distinguer les ressources dissimilaires et donner des informations sur l'emplacement des ressources ;

- *L'organisation des ressources numériques* : l'utilisation des métadonnées permet d'organiser les ressources numériques selon des critères bien définis (type de ressource, objet, l'emplacement, etc.), ce qui permet d'optimiser la recherche ;

L'identification numérique : la plupart des schémas de métadonnées contiennent des éléments tels que des numéros pour identifier l'objet de ces métadonnées. Une ressource peut être identifiée par son emplacement (exemple URL). Toutefois, cette donnée peut changer. Mais l'ajout des métadonnées à cet identifiant peuvent agir comme un ensemble de données d'identification et différencier une ressources d'une autre ;

- *L'archivage et préservation* : les métadonnées sont essentielles pour que les ressources survivent et soient accessible dans le futur. L'archivage et préservation exigent des éléments spéciaux pour suivre l'historique d'une ressource (d'où elle vient et comment elle a changé au fil du temps), détailler ses caractéristiques physiques et de documenter ses comportements afin de l'adapter aux technologies futurs.

Pour que les métadonnées puissent remplir leur rôle et faciliter l'accès aux ressources en ligne, il faut impérativement qu'un standard stable soit utilisé afin que fournisseurs de ressources et apprenants puissent utiliser le même référentiel [Bourda 2002].

La définition d'un standard pour décrire les métadonnées constitue une première étape nécessaire en vue de faciliter la recherche et la réutilisation des OPs [Malaxa 2005]. Plusieurs travaux ont été réalisés dans cette direction et ont donné lieu à des standards de description tels que le standard LOM (que nous allons détailler dans la quatrième partie du présent chapitre). Dans ce qui suit nous allons présenter quelques nouvelles technologies associées aux OPs à savoir le web sémantique et les ontologies.

2.5.2 Technologies associées

Dans cette partie, nous allons présenter les principales technologies associées aux objets pédagogiques notamment le web sémantique et les ontologies ainsi que les concepts qui leurs sont liés.

2.5.2.1 Web sémantique

Le Web sémantique est une manière de définir ou de publier des données sur le Web de telle sorte que les ordinateurs puissent aussi les comprendre et les traiter d'une manière significative.

Tim Berners-Lee [Berners-Lee et al. 2001], fondateur du concept, définit le web sémantique comme étant une extension du Web qui vise à enrichir les ressources disponibles sur le toile, avec des descriptions sémantiques de leur contenu avec comme but de les rendre exploitables par des programmes informatiques.

De cette définition, on peut dire que le web sémantique à la différence du web traditionnel, permet aux machines de rechercher et d'identifier des OPs en utilisant des données formalisées spécifiant la sémantique de ces OPs. Ainsi, le web sémantique permet de répondre aux problématiques liées à l'identification sémantique des OPs, les relations qui les lient entre eux, etc.

Le web sémantique appliqué au domaine d'apprentissage est considéré comme une très importante initiative touchant l'éducation. Cette fusion a donné naissance à un nouveau volet du web sémantique appelé *web sémantique éducatif*. Ce volet consiste, essentiellement, à ajouter une nouvelle couche sémantique pour les ressources d'apprentissage afin de faciliter leurs récupérations et leurs réutilisations [El Mezouary 2011].

2.5.2.2 Ontologie

L'ontologie constitue la base de représentation du web sémantique. D'où la nécessité d'étudier ce concept d'un point de vue technologique. Plusieurs définitions du terme «ontologie» sont citées dans la littérature. Pour le domaine de l'informatique, ce terme a commencé à être employé en intelligence artificielle au début des années 90. Nous allons présenter ici quelques unes. Selon Gruber [Gruber 1993], une ontologie est *une spécification explicite d'une conceptualisation*. Pour Brost [Brost 1997], une ontologie est définie comme étant *une spécification formelle d'une conceptualisation partagée*. Quant à Chandrasekaran [Chandrasekaran et al. 1999], une ontologie est *une théorie du contenu sur les sortes d'objets, les propriétés de ces objets et leurs relations possibles dans un domaine spécifié de connaissances*.

Ce que nous pouvons retenir de ces définitions, c'est que l'ontologie permet de spécifier un processus de classification des concepts d'un domaine donné et les relations qui lient ces concepts. Les concepts permettent de comprendre le domaine et la classification des données. Tandis que les propriétés permettent de déterminer les relations entre les différents concepts.

2.5.2.3 Langages du web sémantique

Depuis l'avènement du Web sémantique, plusieurs outils et langages de formalisation ont été mis en place. C'est dans ce contexte que le World Wide Web Consortium (W3C) a proposé plusieurs formalismes de langages qui sont basés sur des spécifications du langage de balisage XML. Dans cette section, nous allons nous intéresser plus particulièrement aux langages RDF, RDFS et OWL.

- **Le langage RDF**

Le langage RDF (Ressource Description Framework) a été créé en 1999 par le (W3C). Il s'agit d'un des formalismes standards de représentation de connaissances sur le Web et se présente sous une forme particulière du langage XML avec un langage sémantique spécifique. L'objectif du langage RDF est de rendre plus pertinent le traitement automatique des informations contenues sur le Web. Ce langage, en outre des mots clés contenus dans un document, fournit une information plus sémantique aux outils de traitement de l'information. Un autre objectif de RDF est de pouvoir manipuler et classer des métadonnées Web, afin de fournir des informations sur les ressources Web et les systèmes qui les utilisent. Ailleurs, RDF permet de développer des modèles d'information ouverts et applicables à n'importe quel application. RDF contribue aussi à l'optimisation de la coopération entre applications, en permettant de combiner les données de plusieurs applications, pour générer de nouvelles informations.

Nous notons, toutefois, que le langage RDF n'est pas assez robuste pour exprimer tous les aspects sémantiques liés à un domaine en particulier. En effet, le langage RDF ne permet pas de définir un vocabulaire spécifique dédié à un contexte d'utilisation particulier. De plus, l'organisation des ressources ne permet pas d'utiliser des moteurs d'inférence pour déduire de nouveaux triplets, ils doivent donc être explicitement répertoriés, ce qui peut être laborieux même avec un outil permettant d'éditer des fichiers RDF [Cueilliez 2006].

- **Le langage RDFS**

Pour pallier aux limites de RDF, Le W3C a proposé un nouveau langage appelé *RDF Schéma (RDFS)*. RDF Schéma permet de donner du sens aux propriétés associées à une ressource et de formuler des contraintes sur les valeurs associées à une propriété afin de lui assurer aussi une signification.

Le langage RDFS utilise les notions de classe et de sous-classe et définit un lien sémantique entre ces deux notions, ce qui permet de hiérarchiser les ressources. RDFS permet aussi d'exprimer des contraintes sur les éléments du vocabulaire qui permettent d'éviter les erreurs d'association sémantique entre classes.

En fait RDF-Schema ne permet pas de représenter la connaissance de manière aussi formelle qu'en utilisant des ontologies. Pour cela, le langage doit être enrichi et de nouvelles contraintes doivent être ajoutées. Dans la section suivante, nous allons présenter le langage de représentation des ontologies OWL.

- **Le langage OWL**

Le langage OWL est conçu comme extension des langages RDF et RDF Schema. Ce langage intègre en plus, des outils de comparaison, des propriétés et des classes tels que l'identité, l'équivalence, le contraire, la cardinalité, la symétrie, la transitivité, la disjonction, etc.

Le langage OWL permet, ainsi, une plus grande capacité d'interprétation du contenu web et de raisonnement que RDF et RDFS. Ceci est dû essentiellement au vocabulaire qui est plus large et à la vraie sémantique formelle.

Le langage OWL offre trois sous-langages d'expression croissante conçus pour des communautés de développeurs et d'utilisateurs spécifiques [W3C 2004] :

- **Le langage OWL Lite** : ce langage concerne les utilisateurs ayant principalement besoin d'une hiérarchie de classifications et de mécanismes de contraintes simples. Par exemple, quoique OWL Lite gère des contraintes de cardinalité, il ne permet que des valeurs de cardinalité de 0 ou 1 ;

- **Le langage OWL DL** : ce langage concerne les utilisateurs souhaitant une expressivité maximum sans sacrifier la complétude des calculs et la décidabilité que tous ces calculs seront

terminés dans un temps fini. OWL DL doit son nom à sa correspondance avec la logique de description ;

- **Le langage OWL Full** : ce langage est destiné aux utilisateurs souhaitant une expressivité maximum et la liberté syntaxique de RDF sans garantie de calcul en étendant la signification le vocabulaire prédéfini (RDF ou OWL). Dans ce qui suit et comme nous l'avons souligné auparavant, nous allons présenter les normes et standards liés aux OPs et qui sont les plus cités dans nos références étudiées.

2.6 Standards et normes pour les objets pédagogiques

La normalisation pour la formation en ligne répond aux propriétés des OPs présentées dans le deuxième paragraphe du présent chapitre. Une étude approfondie des différentes normes et standards du e-learning dépasse largement le cadre du présent document. De ce fait, nous allons étudier, dans ce qui suit, les normes les plus connues que nous avons manipulé lors de nos travaux de recherches. Ainsi, nous allons voir successivement le standard LOM, IMS LD et enfin SCORM.

2.6.1 L'exemple LOM

Le groupe de travail P1484.12 des IEEE comportant la contribution de ARIADNE et IMS (IMS) visent un consensus pour leur indexation. La version 6 du LOM (Learning Object Metadata) définit une soixantaine de descripteurs regroupés en neuf catégories Yolaine (2001)

- General : caractéristiques indépendantes du contexte comme Identifier (un identificateur global unique) ou Title (le nom de la ressource) ou Language (la langue utilisée principalement par la ressource pour communiquer avec l'utilisateur).
- LifeCycle : caractéristiques relatives au cycle de vie, comme Version ou Status (Draft, Final, Revised, Unavailable).
- Metametadata : caractéristiques de la description elle-même comme Contribute (personnes ayant participé à l'élaboration des métadonnées).
- Technical : caractéristiques techniques comme Format (du logiciel nécessaire pour accéder à la ressource).
- Educational : caractéristiques pédagogiques.

- (a) Interactivity Type : le type d'interaction entre la ressource et l'utilisateur typique (Active, Expositive, Undefined) ;
 - (b) Learning Resource Type : le type pédagogique (Exercice, Simulation.. .), peut être présent plusieurs fois ;
 - (c) InteractivityLevel : degré d'interactivité ;
 - (d) SemanticDensity : densité sémantique (Very Low, Low, Medium, HighVery High) ;
 - (e) Intended end user role : utilisateur de la ressource ;
 - (f) Context : environnement d'utilisation de la ressource ;
 - (g) Typical Age Range : âge de l'utilisateur ;
 - (h) Difficulty : difficulté de la ressource ;
 - (i) Typical Learning Time : temps approximatif ou typique pour travailler avec la ressource ;
 - (j) Description : commentaires sur l'utilisation de la ressource ;
 - (k) Language : la langue de l'utilisateur.
- Rights : caractéristiques exprimant les conditions d'utilisation comme Cost (ressource payante,ou non).
 - Relation : caractéristiques exprimant les liens avec d'autres ressources comme kind (nature de la relation).
 - Annotation : commentaires sur l'utilisation pédagogique de la ressource.
 - Classification : caractéristiques de la ressource décrites par des entrées dans des systèmes de classification.

Le modèle LOM comporte trois niveaux d'agrégation, en effet, le plus bas niveau concerne des données brutes ou des Fragments, le niveau 2 concerne l'ensemble d'atomes, par exemple : un document HTML comprenant des images, une conférence, une leçon, ..., et le niveau 3 concerne l'ensemble de ressources : un site web avec un sommaire, un cours entier, ...

Ce standard ne prend pas position sur la taille des granules indexables. Le problème qui se pose ici est différent, car il s'agit d'indexer non seulement un livre entier, mais aussi ses chapitres, en effet, on peut par contre citer ici quelques problèmes qui peuvent être rencontrés en utilisant ce standard à savoir la complexité technique du modèle, et la difficulté de remplir les 79 champs du modèle. Insuffisance en termes d'aspects pédagogiques. Les 9 catégories de LOM renferment implicitement des modèles pédagogiques sous-jacents, approuvés de facto. Et aussi le manque de définition de ce qu'est une ressource pédagogique.

Les LOM s'appliquent indistinctement à des objets qui ont des natures et des fonctionnalités très différentes (ressources, activités et unités d'apprentissage) ; cela empêche leurs descriptions en tant que tels Excès de généralité.

Les LOM offrent des champs de données à utiliser comme fourre tout (Ex : la catégorie "classification" permet à chaque institution de classer l'objet suivant son propre catalogue). Le rôle de descripteur n'est pas assez poussé pour prendre en compte des ressources plus interactives : Comme celles mises en œuvre dans l'évaluation des connaissances.

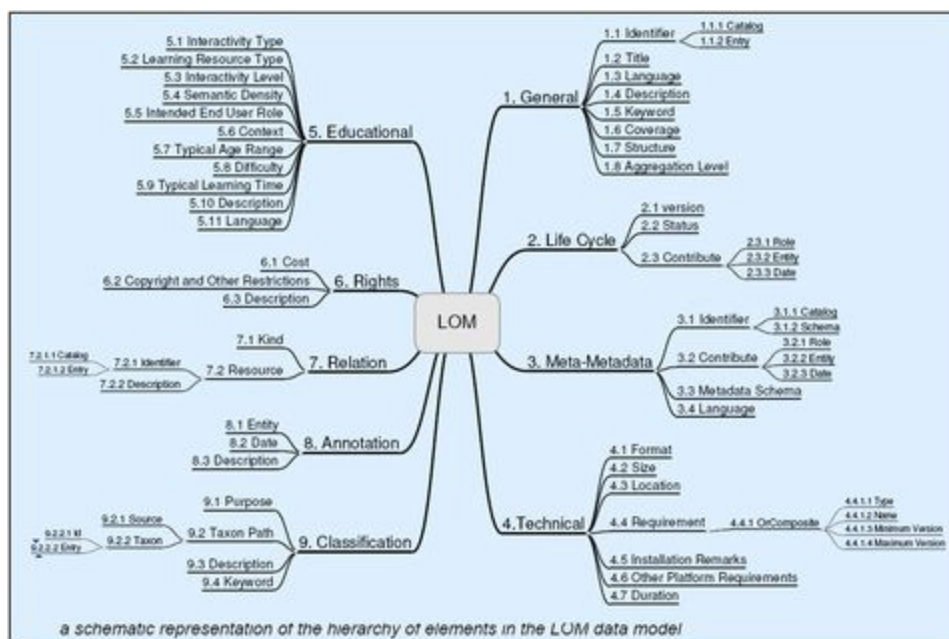


FIG.2.2– Structure générale de LOM (LOM).

2.6.2 L'exemple SCORM

SCORM est un standard d'origine américaine développé par le consortium ADL (Advanced Distributed Learning). Il se présente sous la forme d'un ensemble de spécifications et de normes issues du travail d'autres organisations telles qu'IMS (IMS) et AICC (Aviation Industry CBT Committee).

Le modèle SCORM doit répondre aux différentes exigences : Puce Accessibilité : capacité de repérer des composants d'enseignement à partir d'un site distant, d'y accéder et de les distribuer à plusieurs autres sites.

– Adaptabilité : capacité à personnaliser la formation en fonction des besoins des personnes et

Organisations.

- Durabilité : capacité de résister à l'évolution de la technologie sans nécessiter une reconception, une reconfiguration ou un recodage
- Interopérabilité : capacité d'utiliser dans un autre emplacement et avec un autre ensemble d'outils ou sur une autre plateforme des composants d'enseignement développés dans un site, avec un certain ensemble d'outils ou sur une certaine plateforme.
- Réutilisabilité : souplesse permettant d'intégrer des composants d'enseignement dans des contextes et des applications multiples.

Ces spécifications techniques sont actuellement regroupées selon trois catégories :

- le modèle d'agrégation de contenu (MAC) qui décrit l'organisation des contenus,
- l'environnement d'exécution qui permet de jouer les contenus utilisant le modèle d'agrégation de contenu,
- le séquençement et la navigation (SN) qui reprennent la spécification IMS Simple Sequencing 24 et décrivent l'utilisation des contenus en une activité séquentielle.

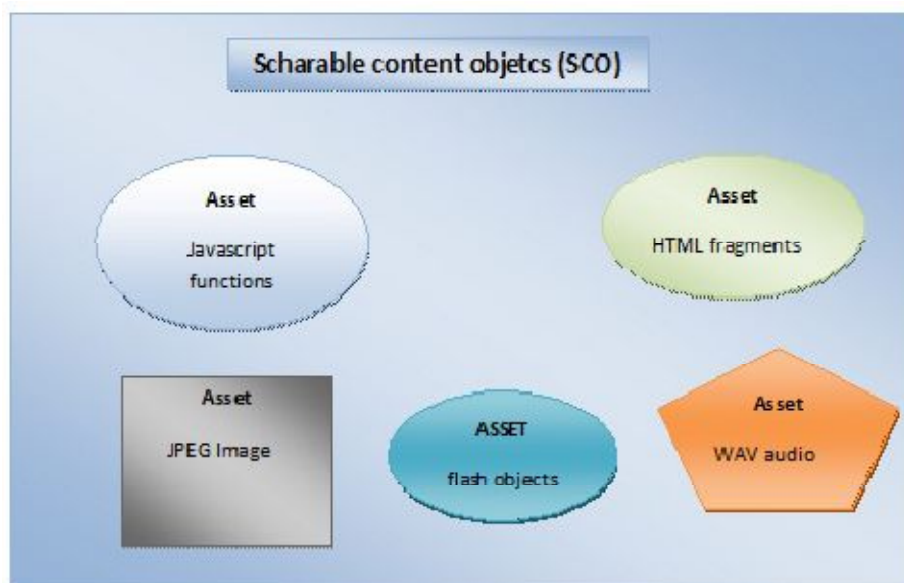


FIG.2.3– Vue sur le contenu partageable de SCORM.

Il nous semble que ces spécifications permettent une meilleure gestion technique des objets pédagogiques, cependant des problèmes d'ordre pédagogiques persistent. Comment expliciter les intentions pédagogiques dans ce modèle ? Et comment les adapter selon les profils d'apprenants ? Nous essayons de dépasser ces difficultés et répondre à ces questions dans le

cadre de notre travail de recherche. Il s'agit de représenter les intentions pédagogiques comme des propriétés de l'objet pédagogique que l'on propose. Cette technique permet de rendre l'objet pédagogique capable de s'adapter à des intentions pédagogiques différentes qui doivent avoir un format spécifiques facilitant leurs utilisations dans le système. La personnalisation de l'objet pédagogique se fait en fonction de l'objectif d'apprentissage à atteindre, représenté comme une liste d'intentions d'enseignement, et du profil de l'apprenant.

Notons aussi qu'un autre problème complexe peut être se poser dans ce contexte concernant l'opérationnalisation d'une théorie pédagogique. Il s'agit de recueillir des corpus psychologiques de recherche à utiliser dans le système pour rendre la théorie opérationnelle. Toute fois, sachant que les objectifs de ces projets restent sur le stade de proposition d'un système d'enseignement générique, par conséquent les recherches doivent concentrer sur la méthode de travail. Quant à notre intervention sur ce point, nous nous sommes proposés de définir une méthode de travail, se basant sur le principe de design pédagogique, pour l'utilisation de concepts issus des théories de sciences humaines.

2.6.3 L'exemple IMS-LD (Learning Design)

Une première synthèse des travaux sur les EML a été réalisée par le groupe de travail CEN/ISSS sur les technologies d'apprentissage (WS-LT). Les résultats de ces travaux ont été réutilisés par le groupe de travail "Learning Design" du consortium IMS qui a débouché en novembre 2002 sur une proposition de standardisation, IMS-Learning Design. Ce modèle statique de design d'apprentissage (LD) se concentre sur trois entités : rôles, activités et environnement.

- Rôles tels qu'apprenant et facilitateur (professeur, tuteur, animateur) sont joués par des personnes décrites par leurs propriétés.
- Activités réalisées par des Rôles, elles font partie d'une structure arborescente appelée Méthode, décomposée en cheminements alternatifs, eux-mêmes décomposés en actes séquentiels, chaque acte est composé de partitions définies chacune par un rôle et son activité.

L'ensemble de partitions forment le scénario de l'acte qui est généralement composé d'une structure d'activités, qui contient d'autres structures d'activités jusqu'aux activités d'apprentissage ou activités de support terminales.

Environnements regroupe tous les types d'Objets d'Apprentissage (OA) ou de services utilisés par les rôles dans les activités, ainsi que les résultats produits par les rôles dans les activités. L'activation d'une Unité d'Apprentissage, se passe à travers l'élément « Méthode » qui est considéré comme central dans ce modèle. Cet élément unique et ses sous-éléments décrivent le processus d'apprentissage et contrôlent le comportement de l'unité d'apprentissage comme un tout, coordonnant les activités des acteurs dans leurs rôles variés ainsi que l'utilisation qu'ils font des ressources.

Les « Méthode », « Cheminement », « Acte » et « Partition (couple rôle-activité) » sont tous imbriqués l'un dans l'autre, tel que présenté dans la figure 2.4.

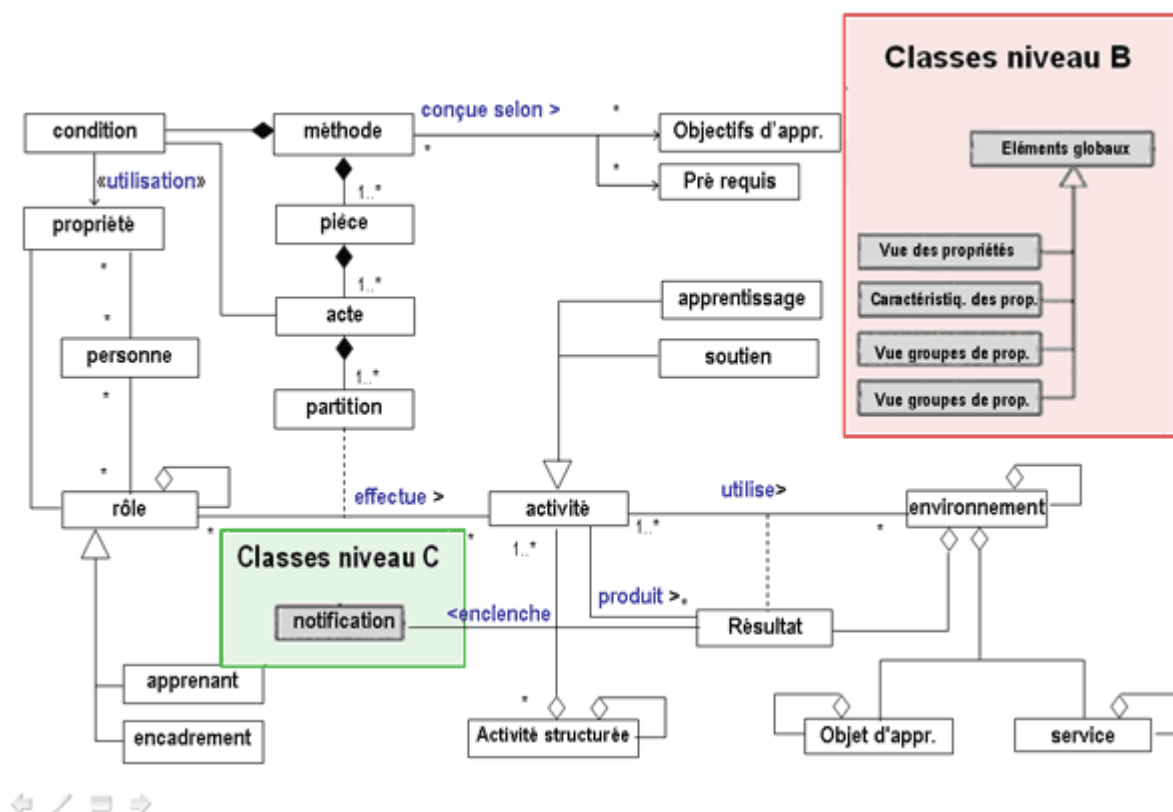


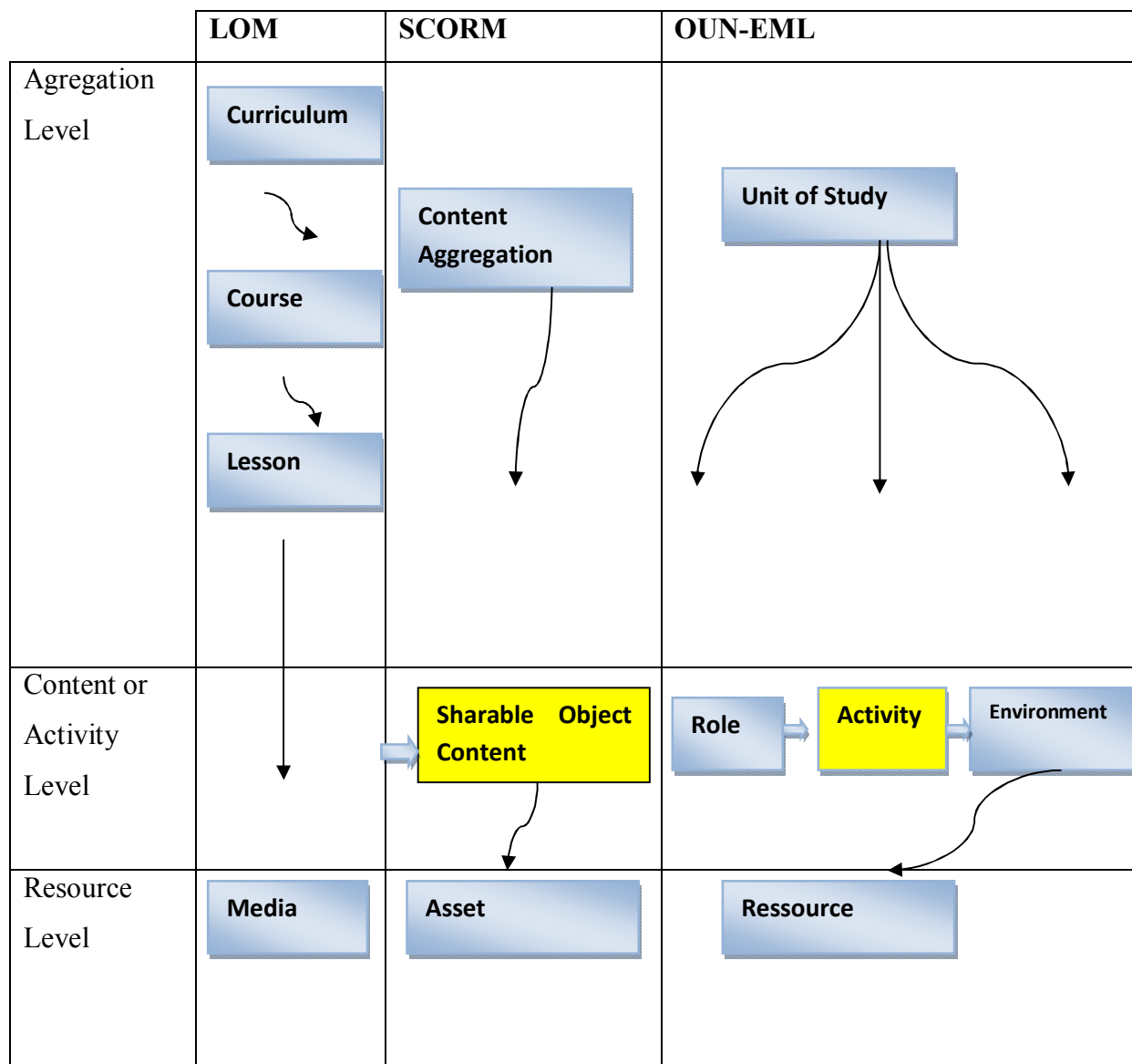
FIG.2.4– Architecture de la spécification IMS-LD, Niveaux A, B et C [Burgos 2005].

2.7 Synthèse

Tous les standards qui sont en cours de développement visent la spécification des normes, des procédures et des manuels pour les composants logiciels, les outils et les techniques, ainsi que des méthodes de conception qui aident au développement, au déploiement, à la maintenance, à l'accessibilité, à l'adaptabilité et à l'interopérabilité des composants et des systèmes pédagogiques informatisés. L'interopérabilité signifie que plusieurs systèmes, qu'ils soient identiques ou radicalement différents, puissent communiquer sans ambiguïté et opérer ensemble. Le modèle conceptuel utilisé dans chaque projet de recherche fait apparaître les différents niveaux de conception de l'objet pédagogique. Nous rappelons ici que l'objet pédagogique que l'on propose appartient à la classe des activités pédagogiques. Notre objectif est de concevoir un objet pédagogique qui prend en compte le processus d'acquisition des connaissances. Il s'agit de spécifier des artefacts informatiques complexes, constituant l'objet pédagogique, intégrant le matériel pédagogique qui favorisent l'apprentissage et permettent à l'apprenant de réaliser les tâches prescrites correspondant aux intentions d'enseignement.

Ces artefacts représentent les instruments pédagogiques conçus selon l'approche instrumentale. Il s'agit de prendre en considération les deux côtés (technologique et rationnel). Ainsi, la structuration de l'objet pédagogique doit correspondre aux attentes pédagogiques et aux objectifs d'apprentissage afin de pouvoir inclure tous les éléments d'adaptation dans les instruments représentés par les composantes structurelles de l'objet pédagogique.

TAB.2.1– Synthèse sur les trois standards [Pernin 2003].



2.8 Conclusion

La question de la création de ressources numériques et l'utilisation des métadonnées pour individualiser les enseignements est particulièrement d'actualité. Ainsi, comme présenté dans ce chapitre, plusieurs travaux ont été consacrés à la proposition des outils et des méthodes pour cerner ces deux problématiques.

Ce chapitre a essayé de faire une mise au point sur la terminologie utilisée dans les EIAH, notamment les termes «ressource pédagogique» et «objet pédagogique». Par ailleurs, nous

nous sommes intéressés à la question des métadonnées, aux technologies associées et leurs importances pour la personnalisation de l'apprentissage, tout en passant en revue quelques standards et normes pour la création des contenus pédagogiques. Ainsi, nous avons présenté les définitions les plus répandues dans la littérature des OPs.

Après, nous nous sommes focalisés sur les objets pédagogiques en présentant leurs caractéristiques, propriétés, principes de conception et technologies liées. Nous avons par la suite, passé en revue les métadonnées, les technologies associées et leurs apports pour les systèmes de personnalisation de l'apprentissage. En suite, nous avons fait le tour des standards les plus répandus pour la création et l'indexation des contenus pédagogiques. A la fin de ce chapitre, nous avons présenté une synthèse des standards utilisés pour la création et l'indexation des OPs.

Chapitre 3

La personnalisation de l'apprentissage et le calcul évolutionnaire

3.1 Introduction

Les Environnements Informatiques pour l'Apprentissage Humain (EIAH) peuvent (1) fournir le contenu d'apprentissage ou services; (2) contrôler l'expérience d'apprentissage et le progrès ; et (3) gérer le contenu d'apprentissage, les services et les cours associés [Chu et al. 2009]. La personnalisation est une caractéristique importante du siècle actuel [Hauger et al. 2007], elle est clairement établie dans la navigation Web et e-commerce, où les services sont organisés en fonction du profil et des intérêts de l'utilisateur. La Personnalisation dans l'apprentissage est considérée comme le centre de la scène pour une expérience d'apprentissage efficace [Hauger et al. 2007], elle consiste à adapter la connaissance selon les préférences, les capacités de et les exigences de l'apprenant. Plusieurs approches ont été appliquées pour faire la personnalisation de l'apprentissage, parmi eux les approches du calcul évolutionnaire qui ont montrés leur efficacité pour faire de la personnalisation de l'apprentissage.

Dans les sections suivantes, nous allons voir les différentes approches du calcul évolutionnaire pour la personnalisation de l'apprentissage.

3.2 Le calcul évolutionnaire

Le calcul évolutionnaire est une catégorie d'algorithmes qui sont développés selon le théorème darwinien (la survie du plus apte) et prends la métaphore de l'évolution naturelle [Hwang et al. 2008]. Les Algorithmes évolutionnaires (AES) sont principalement utilisés pour l'optimisation, la modélisation et simulation des problèmes [Eiben et al. 2007]. Ce qui distingue les techniques des algorithmes évolutionnaires, c'est qu'elles sont basées sur les

principes darwiniens de la sélection naturelle. Diverses espèces ont historiquement réussi à évoluer pour s'adapter à des environnements divers changeants. Sur la base de cette métaphore, si ces techniques ont été fournis avec de nouveaux environnements, alors une adaptation des solutions évolutives serait observée [Eiben et al. 2007].

Une autre caractéristique qui distingue les techniques du calcul évolutionnaire, c'est qu'elles sont des techniques de recherche basées sur la Population. La capacité de ces algorithmes à maintenir une population diversifiée de solutions n'aide pas seulement à traiter des espaces de recherche grands et discontinus, mais aussi favorise l'évasion de l'optima locaux [Eberhart et al. 2007, Eiben et al. 2007]. Les techniques du calcul évolutionnaire sont connues comme des techniques flexibles et robustes avec une performance satisfaisante [Eiben et al. 2007]. Le calcul évolutionnaire inclut les algorithmes génétiques (AGs) [Hollande 1975], Stratégies d'évolution (SE) [Rechenberg 1965] [Shwefel 1965], programmation évolutive (PE) [Fogel et al. 1965], et la Programmation Génétique (PG) [Koza 1992]. Plus tard, l'optimisation par colonie de fourmis (en Anglais ACO) [Dorigo 1992] et l'Optimisation par essais particuliers (OEP) [Eberhart et al. 1995] ont été introduits comme nouveaux membres dans la famille des algorithmes du calcul évolutionnaire. Selon [Chu et al. 2009] et [Hwang et al. 2008], les méthodes des AGs et OEP sont considérées comme les techniques les plus populaires du calcul évolutionnaire.

3.2.1 L'optimisation par les essais particuliers (EOP)

Description

L'optimisation par essaim particulière (OEP) est une technique, proposée par Kennedy et Eberhart [Kennedy et al. 1995], qui s'inspire du comportement social des animaux lorsqu'ils se déplacent en essaim. L'algorithme utilise la collaboration entre des agents simples pour tenter de localiser l'optimum dans un espace de recherche. On entend par là que chaque particule utilise non seulement sa propre expérience, mais aussi l'expérience globale de l'essaim de particules. Cette méthode, qui est relativement récente, s'est avérée compétitive vis-à-vis de méthodes déjà éprouvées [Eberhart et al. 1998] et elle est actuellement en plein essor. L'OEP, comme toutes les autres métaheuristiques, présente le désavantage de comporter un nombre élevé de paramètres. Les performances de la méthode face à un problème donné étant fortement liées aux valeurs de ses paramètres de réglage, il est souvent

difficile et long de trouver les valeurs optimales de chacun des paramètres. C'est pourquoi, de nombreuses recherches sont menées en vue de proposer des méthodes comportant un nombre minimal de paramètres. L'objectif ultime est de pouvoir proposer un algorithme dont les valeurs des paramètres s'adaptent automatiquement en cours d'exécution de l'algorithme. L'algorithme ne serait alors qu'une "boîte noire" à laquelle l'utilisateur n'aurait à procurer que le critère d'arrêt et la fonction objectif.

L'essaim de particules correspond à une population d'agents simples, appelés particules. Chaque particule est considérée comme une solution du problème, où elle possède une position (le vecteur solution) et une vitesse. De plus, chaque particule possède une mémoire lui permettant de se souvenir de sa meilleure performance (en position et en valeur) et de la meilleure performance atteinte par les particules « voisines » (informatrices) : chaque particule dispose en effet d'un groupe d'informatrices, historiquement appelé son voisinage. Un essaim de particules, qui sont des solutions potentielles au problème d'optimisation, « survole » l'espace de recherche, à la recherche de l'optimum global. Le déplacement d'une particule est influencé par les trois composantes suivantes :

- Une composante d'inertie : la particule tend à suivre sa direction courante de déplacement ;
- Une composante cognitive : la particule tend à se diriger vers le meilleur site par lequel elle est déjà passée ;
- Une composante sociale : la particule tend à se fier à l'expérience de ses congénères et, ainsi, à se diriger vers le meilleur site déjà atteint par ses voisins.

Formalisation

Dans un espace de recherche de dimension D , la particule i de l'essaim est modélisée par son vecteur position $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ et par son vecteur vitesse $\vec{v}_i = (v_{i1}, v_{i2}, \dots, v_{iD})$. La qualité de sa position est déterminée par la valeur de la fonction objectif en ce point. Cette particule garde en mémoire la meilleure position par laquelle elle est déjà passée, que l'on note $\vec{pbest}_i = (pbest_{i1}, pbest_{i2}, \dots, pbest_{iD})$. La meilleure position atteinte par les particules de l'essaim est notée $\vec{gbest}_i = (gbest_{i1}, gbest_{i2}, \dots, gbest_{iD})$. Nous nous référons à la version globale de l'EOP, où toutes les particules de l'essaim sont considérées comme voisines de la particule i , d'où la notation \vec{Gbest} (global best).

Au départ de l'algorithme, les particules de l'essaim sont initialisées de manière aléatoire/régulière dans l'espace de recherche du problème. Ensuite, à chaque itération, chaque particule se déplace, en combinant linéairement les trois composantes citées ci dessus.

$$v_{i,j}^{t+1} = wv_{i,j}^t + c_1 r_{1,i,j}^t [pbest_{i,j}^t - x_{i,j}^t] + c_2 r_{2,i,j}^t [gbest_{i,j}^t - x_{i,j}^t] \quad (3.1)$$

$$j \in \{1, 2, \dots, D\}.$$

$$x_{i,j}^{t+1} = x_{i,j}^t + v_{i,j}^{t+1}, j \in \{1, 2, \dots, D\} \quad (3.2)$$

En effet, à l'itération $t + 1$, le vecteur vitesse et le vecteur position sont calculés à partir de l'équation (3.1) et de l'équation (3.2), respectivement.

Où w est une constante, appelée coefficient d'inertie; c_1 et c_2 sont deux constantes, appelées coefficients d'accélération ; r_1 et r_2 sont deux nombres aléatoires tirés uniformément dans $[0, 1]$, à chaque itération t et pour chaque dimension j . Les trois composantes mentionnées ci-dessus (i.e. d'inertie, cognitive et sociale) sont représentées dans l'équation (3.1) par les termes suivants :

1. $wv_{i,j}^t$ correspond à la composante d'inertie du déplacement, où le paramètre w contrôle l'influence de la direction de déplacement sur le déplacement futur ;
2. $c_1 r_{1,i,j}^t [pbest_{i,j}^t - x_{i,j}^t]$ correspond à la composante cognitive du déplacement, où le paramètre c_1 contrôle le comportement cognitif de la particule ;
3. $c_2 r_{2,i,j}^t [gbest_{i,j}^t - x_{i,j}^t]$ correspond à la composante sociale du déplacement, où le paramètre c_2 contrôle l'aptitude sociale de la particule. Une fois le déplacement des particules effectué, les nouvelles positions sont évaluées et les deux vecteurs $\vec{P}best$ et $\vec{G}best$ sont mis à jour , à l'itération $t + 1$, suivant les deux équations (3.3) (dans le cas d'une minimisation) et (3.4) (dans une version globale de EOP), respectivement. Cette procédure est présentée dans l'Algorithme 1 dessous, où N est le nombre de particules de l'essaim.

$$\vec{p}best_i(t+1) = \begin{cases} \vec{p}best_i(t), & \text{si } f(\vec{x}_i(t+1)) \geq f(\vec{p}best_i(t)) \\ \vec{x}_i(t+1) & \text{sinon} \end{cases} \quad (3.3)$$

$$\vec{G}best(t+1) = \arg \min_{\vec{p}best_i} f(\vec{p}best_i(t+1)), 1 \leq i \leq N. \quad (3.4)$$

TAB.3.1– Algorithme d'optimisation par essaim particulaire [El Dor 2012].

Algorithme 1 : Algorithme d'optimisation par essaim particulaire

Début

Initialiser aléatoirement N particules : position et vitesse ;

Evaluer les positions des particules ;

Pour chaque particule i , $\vec{p}best_i = \vec{x}_i$;

Calculer $\vec{G}best$ selon (3.4)

tant que le critère d'arrêt n'est pas satisfait **faire**

Déplacer les particules selon (3.1) et (3.2)

Evaluer les positions des particules

Mettre à jour

$\vec{p}best_i$ et $\vec{G}best$ selon (3.3) et (3.4)

Fin

3.2.2 Les algorithmes génétiques (AGs)

Description

Les algorithmes génétiques (voir table 3.2) sont sans doute la classe la plus représentative et la plus populaire parmi les algorithmes évolutionnaires. Ils sont basés sur les travaux de Holland [Holland 1992] et de Goldberg [Goldberg 1989], et ils ont été utilisés au début avec des représentations binaires, où les opérateurs de croisement et de mutation jouent un rôle majeur. Un algorithme génétique est un modèle d'apprentissage machine, qui est inspiré du processus de l'évolution observé dans la nature. Pour ce faire, un système artificiel contenant

une population d'individus représentée par des chromosomes est créée. Des processus observés dans la nature montrent que différents individus sont en compétition pour l'accès aux ressources dans leur environnement. Des individus sont meilleurs que d'autres. Ces derniers individus ont plus de chances de survivre et par conséquent de transmettre leurs gènes à la génération suivante. Dans la nature, on observe que l'encodage des informations génétiques (génome), qui résulte d'une reproduction non sexuelle, produit des enfants génétiquement proches de leurs parents. Par contre, une reproduction sexuelle permet la création d'enfants radicalement différents de leurs parents, tout en ayant les mêmes caractéristiques communes à l'espèce. D'une façon simplifiée, le mécanisme au niveau moléculaire est le suivant : deux chromosomes se réunissent, et un échange d'information génétique est effectué. Ce dernier processus est appelé le croisement, puisque les composantes génétiques sont croisées d'un chromosome à l'autre. L'opération de croisement se passe dans un environnement dans lequel le choix d'un partenaire ou l'opération de sélection est une fonction de correspondance, par exemple on choisit le plus compétitif dans son environnement. Quelques algorithmes génétiques utilisent une fonction de sélection simple (par exemple probabiliste) : dans ce cas, le croisement et la reproduction non sexuelle (où les composantes génétiques restent intactes) sont négligés. D'autres méthodes d'implantation choisissent un modèle où quelques individus sont sélectionnés au hasard dans des sous-groupes et les plus aptes sont conservés. Ce type de sélection est observé dans la nature.

Les deux processus qui contribuent le plus à l'évolution sont le croisement et la correspondance basée sur la sélection / reproduction. Il y a eu des preuves mathématiques qui montrent que le processus de reproduction basé sur la correspondance est en fait sous-optimal. La mutation joue aussi un rôle dans le processus de l'évolution, mais l'importance de son rôle reste un sujet de débat (certains la considèrent comme un opérateur de second plan, tandis que d'autres lui attribuent un rôle dominant dans le processus de l'évolution). Nous ne pouvons pas dire que les algorithmes génétiques (AGs) (une simulation du processus génétique) ne sont pas une recherche aléatoire. Les AGs utilisent un processus stochastique, mais les résultats obtenus sont meilleurs que des résultats aléatoires. Les AGs sont utilisés dans de nombreux domaines d'application. Par exemple, les problèmes d'optimisation multidimensionnelle ; dans ce type de problème, les chaînes de caractères du chromosome sont utilisées pour coder les différents paramètres à optimiser. En pratique, le modèle génétique peut être implémenté en utilisant une matrice de bits pour représenter les chromosomes. Les simples opérations de manipulation de bits permettent d'implémenter le

croisement, la mutation et d'autres opérations. Même si certains chercheurs ont étudié les AG avec des chaînes de caractères ou avec d'autres structures variables, la plupart des recherches se focalisent sur les chaînes de caractères de longueur fixe. Quand un algorithme génétique est implémenté (voir algorithme 2), il est fait en général de la façon suivante : (1) Évaluer la valeur de la fonction objectif de tous les individus de la population. (2) Créer une nouvelle population se basant sur les opérations de croisement, les évaluations de correspondance et mutation, sur les individus que nous venons d'évaluer dans la première étape. (3) Éliminer l'ancienne population et itérer en utilisant la nouvelle. Une seule itération de cette boucle est appelée génération. Il faut noter qu'il n'y a pas de bases théoriques pour ce modèle d'implémentation. En fait, ce processus en entier, ou bien le modèle, n'est pas observé dans la nature, mais il semble être un modèle d'implémentation convenable. La première génération (génération 0) de ce processus opère sur une population d'individus engendrés aléatoirement. Et, à partir de cette dernière génération, les opérateurs génétiques opèrent en parallèle, avec l'évaluation de la fonction pour améliorer la population [Song et al. 2001], [Subrata et al. 2003].

TAB.3.2– Algorithme génétique.

Algorithme 2 : l'algorithme génétique classique

Début

$t=0$

Initialiser $P(t)$

Évaluer la population $P(t)$

Tant que condition de terminaison non satisfaite

faire

Début

$t = t+1$ Sélectionner $P(t)$ à partir de $P(t-1)$

Mutations

Croisements

Évaluations

Fin

Fin

3.2.3 Les algorithmes de colonies de fourmis

Description

Les algorithmes de colonies de fourmis sont des algorithmes inspirés du comportement des fourmis et constituent une famille de méta-heuristiques d'optimisation pour résoudre naturellement des problèmes complexes. Une telle aptitude s'avère possible en raison de la capacité des fourmis à communiquer entre elles indirectement, par le dépôt dans l'environnement de substances chimiques, appelées phéromones. Ce type de communication indirecte est appelé stigmergie. En anglais, le terme consacré à la principale classe d'algorithmes est en anglais Ant Colony Optimization (ACO). La principale illustration de ce phénomène est donnée par la figure 3.1 : un obstacle est placé sur le trajet des fourmis qui, après une étape d'exploration, finiront par emprunter le plus court chemin entre le nid et la source de nourriture [Goss et al. 1989]. Les fourmis qui sont retournées le plus rapidement au nid en passant par la source de nourriture sont celles qui ont emprunté le chemin le plus court. Il en découle que la quantité de phéromones déposées par unité de temps sur ce trajet est plus importante que sur les autres. Par ailleurs, une fourmi est d'autant plus attirée à un certain endroit que le taux de phéromones y est important. De ce fait, le plus court chemin a une probabilité plus importante d'être emprunté par les fourmis que les autres chemins et sera donc, à terme, emprunté par toutes les fourmis. Le premier algorithme d'optimisation s'inspirant de cette analogie a été proposé par Colorni, Dorigo et Maniezzo [Colorni et al. 1991, Dorigo et al. 1996], afin de résoudre le problème du voyageur de commerce.

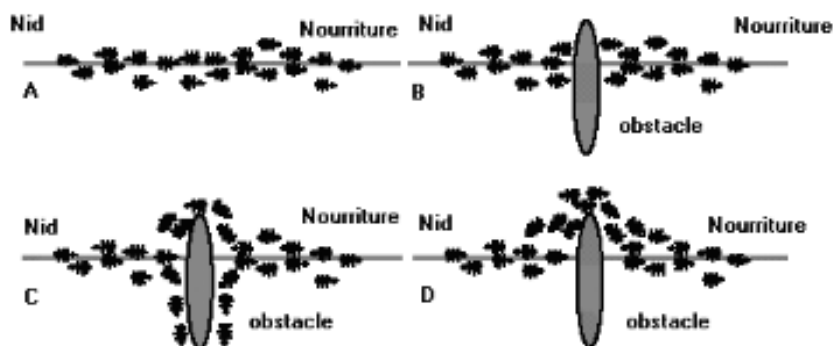


FIG.3.1– Les Fourmis réelles suivent un chemin entre le Nid et la Nourriture. (B) Un obstacle apparaît sur le chemin : Les Fourmis choisissent de tourner soit à gauche soit à droite avec

une probabilité égale. (C) La phéromone est déposée plus rapidement sur le chemin le plus court. (D) Toutes les fourmis ont choisi le chemin le plus court.

Les fourmis utilisent les pistes de phéromones pour marquer leur trajet, par exemple entre le nid et une source de nourriture. Une colonie est ainsi capable de choisir sous certaines conditions le plus court chemin vers une source à exploiter [Goss et al. 1989], sans que les individus aient une vision globale du trajet.

TAB.3.3– Algorithme de colonie de fourmis [Yang et al. 2009, Wang et al. 2008].

Algorithme 3 : pseudo code de l'algorithme de colonie de fourmis

Début

Initialiser la piste des phéromones

Répéter

Construire solution

Appliquer une recherche locale

Mettre à jour les pistes de phéromones

Jusqu'à la condition de terminaison

Fin

3.3 Personnalisation de l'apprentissage basée calcul évolutionnaire

Considérons deux répertoires d'informations. Le premier représente une grande population d'apprenants, atteignant plusieurs centaines et même des milliers de cours en ligne populaires. Ce répertoire comprend son profil, ses connaissances de fond, des évaluations, des activités d'apprentissage, des progrès et l'évaluation dans chacun d'eux. Le second correspond au large et grand nombre d'objets pédagogiques et leurs métadonnées connexes. Les cours peuvent contenir plusieurs centaines d'objets pédagogiques. Donc, étant donné n objets pédagogique dans le programme, nous sommes à la recherche de la séquence optimale qui s'adapte en premier à la structure pédagogique et en second aux besoins individuels des apprenants.

L'espace de la solution de la personnalisation de l'apprentissage comprend tous les $n!$ séquences possibles. En effet, l'espace de solutions est encore plus grand si toutes les séquences sont autorisées. Une solution possible est celle qui satisfait toutes les exigences mentionnées ci-dessus. En réalité, la personnalisation de l'apprentissage est un problème avec de nombreux paramètres qui doivent être optimisés simultanément dans un environnement d'apprentissage complexe.

Les Algorithmes approximatifs représentent la seule façon de s'attaquer aux grandes instances d'un tel problème. Parmi les différentes heuristiques et méta-heuristiques existantes, les approches communautaires ont attirées l'intérêt des chercheurs du e-Learning pour résoudre ce problème car ils peuvent s'adapter naturellement à l'environnement d'apprentissage en évolution rapide. En effet, L'efficacité et l'efficacité des parcours d'apprentissage spécifiques sont évalués par le temps nécessaire à l'achèvement d'un objectif pédagogique, ou le taux de réussite des apprenants, ou une combinaison de ces mesures [Tattersall et al. 2005, Janssen et al. 2005].

L'EIAH est parmi les environnements les plus difficiles à mettre en œuvre. Tout d'abord, une troisième partie prenante apparaît dans l'équation en plus du contenu de l'apprenant et de l'apprentissage; à savoir le tuteur. Deuxièmement, le contenu d'apprentissage est généralement présenté comme différents services (simulations, des forums de discussion, des évaluations ... etc.) plutôt qu'un simple texte ou un document [Gil et al. 2008]. La plupart des algorithmes qui sont appliqués sont les algorithmes de colonie de fourmis [Yang et al. 2009, Gil et al. 2008, Wang et al. 2008, Wong et al. 2009, Semet et al. 2003a, Semet et al. 2003b, Gutiérrez et al. 2007, Tattersall et al. 2005] et OEP [Wang 2009]. Les algorithmes génétiques combinés avec des techniques d'exploration de données [Guo 2009, Romero et al. 2002, Guo 2008]. Nous présentons différentes approches de personnalisation des objets pédagogiques.

3.3.1 La personnalisation basée ACO

Ant Colony Optimization (ACO) est la méta-heuristique la plus utilisée pour la personnalisation de l'apprentissage. ACO implique une colonie d'agents artificiels qui simulent le comportement des fourmis naturelles dans la résolution de problèmes d'optimisation. Dans cette perspective, les fourmis coopèrent pour trouver les chemins les plus courts comme les arbres couvrants de la source de nourritures à leur nid dans un délai raisonnable. Ces fourmis communiquent indirectement sous une forme connue sous le nom

« stigmergie » par la déposition de phéromones aromatiques sur les sentiers qu'ils suivent. Les chemins les plus fréquentés ont des niveaux élevés de phéromones tandis que les chemins les moins fréquentés oublient progressivement ce sens aromatique vu que les phéromones s'évaporent dans le temps [Dorigo et al. 1996, Yang et al. 2009, Wang, et al. 2008, Semet et al. 2003a, Semet et al. 2003b, Gutiérrez et al. 2007]. La combinaison des phéromones et l'évaporation produit des résultats robustes et dynamiques.

Principe

En effet, les apprenants sont modélisés comme des fourmis qui vagabondent dans un graphe où les objets pédagogiques représentent les nœuds du graphe, et l'éventuelle séquence des objets pédagogiques est modélisée par les arcs entre les nœuds. Les fourmis (les apprenants) cherchent le plus court chemin qui mène à leur nourriture ultime (objectif pédagogique). Les phéromones tout au long de la route (arcs) sont représentées comme des poids indiquant l'importance de chaque arc, contre les autres, selon le même nœud. Ces poids sont fondés sur divers facteurs en fonction de techniques spécifiques, par exemple les poids de l'arc peuvent augmenter / diminuer d'une manière proportionnelle aux score du succès / échec de l'apprenant sur les évaluations en suivant ces chemins.

Travaux

[Semet et al. 2003a] [Semet et al. 2003b, Valigiani et al. 2006] et [Gutiérrez et al. 2007] ont appliqué ACO à un programme d'études secondaires existant qui s'appel en français Paraschool Société. Dans leur approche, les cours sont décomposés en chapitres et un graphe est créé pour chaque chapitre. La valeur de la fonction fitness d'un arc s'appuie sur trois fonctions objectifs [Semet et al. 2003a, Semet et al. 2003b]: son poids pédagogique fixé par l'instructeur, l'historique personnel des destinations visitées par l'apprenant et son succès correspondant sur ce nœud. Et finalement, les phéromones de succès et d'échec émettent sur ce nœud par d'autres apprenants. Cette fonction de fitness détermine la probabilité de la sélection de l'arc qui est basée à la fois sur une recommandation pédagogique et la probabilité. Les auteurs ont décrit un graphe simple ne comportant que quatre nœuds est (voir figure 3.2). En supposant que l'apprenant a complété avec succès la ressource d'apprentissage 1 (LR1), les phéromones positives sont déposées derrière. La prochaine ressource d'apprentissage sera choisit d'une manière stochastique (LR3 avec une probabilité de 80%, LR4 avec une

probabilité de 20%). Si l'apprenant échoue dans LR1, puis les phéromones négatives sont déposées, alors la probabilité se réduit dans sa piste échouée. Les expérimentations ont révélées que les exercices les plus difficiles ne sont généralement pas choisis puisque l'accent était mis sur la maximisation du succès et non l'apprentissage. Pour s'attaquer à ce problème, [Valigiani et al. 2006] ont modifié la fonction de fitness de façon à favoriser les arcs menant à des points sur lesquels les apprenants moyens auront une probabilité de succès de 60%. En outre, les fonctions objectifs ont été légèrement modifiés afin de forcer en quelque sorte les apprenants à prendre un des exercices difficiles [Gutiérrez et al. 2007]. Cependant, les résultats montrent que les apprenants qui choisissent librement leur matériel d'apprentissage ont de meilleurs résultats que ceux guidés par le système.

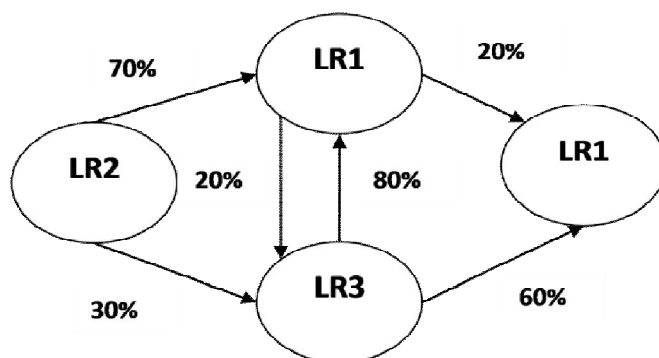


FIG 3.2 – Exemple de graphe pour la personnalisation de l'apprentissage [Al-Muhaideb et al. 2011].

3.3.2 La personnalisation basée EOP

Principe

[Chu et al. 2009] présentent une approche de composition de cours personnalisés basée sur l'optimisation par essais particulaires. Un codage binaire de particules est utilisé, où chaque dimension détermine si un objet pédagogique appartient à la séquence (1) ou non (0). Un essaim initial est généré de façon aléatoire. Les fonctions objectifs décident si *gbest* et *pbest* sont basés sur le degré de correspondance entre le but pédagogique de l'apprenant et les

concepts d'apprentissage: le matériel d'apprentissage, le degré de correspondance du niveau de difficulté, limite du temps d'apprentissage, et la supposition que l'apprenant a déjà étudié ce cours. La vitesse pour chaque dimension (*gbest* et *pbest*) est définie par une fonction de fitness proportionnelle et normalisée. Toute nouvelle position de dimension est stochastique. L'outil de composition des cours a été mis en place et testé dans un environnement réel. Les résultats des apprenants qui l'ont utilisé sont meilleurs que les résultats de leurs pairs qui ne l'ont pas utilisé. En outre, des questionnaires montrent que la plupart des apprenants sont satisfaits de la recommandation des cours en termes de niveau de difficulté et l'harmonie du but d'apprentissage.

Travaux

[De-Marcos et al. 2007] ont introduit le problème de satisfaction de contraintes (PSC) pour modéliser le problème de l'adaptation du contenu et de proposer un agent OEP pour le résoudre. Un essaim initial est généré de façon aléatoire avec une taille fixe. L'encodage entier est utilisé là où chaque dimension représente un nombre d'objet pédagogique. La fonction fitness est basée sur la minimisation d'une fonction de pénalité définie comme le nombre total de violations des contraintes d'ordre pédagogiques. Après, La vitesse est calculée et normalisée pour l'ensemble de particules. La nouvelle position de particules est choisie comme stochastique soit *gbest* ou *pbest*. La condition de terminaison est atteinte lorsque la valeur de la fonction fitness a atteint zéro ou un nombre spécifique d'itérations est écoulé. Le Profil de l'apprenant n'été pas inclus dans les critères de sélection. Des testes ont été effectués sur la personnalisation de cours pour un programme de Master en Ingénierie (Université de Alcalá, Espagne). Les résultats expérimentaux démontrent l'efficacité de cette technique. [De-Marcos et al. 2008b, De-Marcos et al. 2008a, De-Marcos et al. 2009a] ont amélioré la performance de leur système en limitant la croissance rapide de la vitesse observée dans de grands espaces de solutions.

3.3.3 La personnalisation basée AG

Représentation.

La première étape dans la conception d'un AG consiste à définir la représentation du chromosome. La représentation fait correspondre les objets réels (phénotypes) avec leur codage dans la population (génotype) [Eiben et al. 2007]. Dans le contexte de la

personnalisation de l'apprentissage, la représentation définit l'encodage d'une séquence d'apprentissage. La qualité de cette étape affecte fortement le succès de la recherche [Azough et al. 2007]. La représentation la plus utilisée est le codage entier de chromosome à longueur fixe [Lopes et al. 2009, Guo et al. 2009, Chen 2008] ; [De-Marcos et al. 2009b, Azough et al. 2007, Huang et al. 2007, Chen et al. 2008, Acampora et al. 2009]. Chaque gène représente le numéro de série attribué à un objet pédagogique.

Le nombre de gènes dans le chromosome est égal au nombre d'objets pédagogiques qui satisfont l'objet pédagogique but de l'apprenant. Certaines variantes existent, comme le codage binaire de chromosome avec une longueur fixe proposé par [Hovakimyan et al. 2004]. Dans ce codage, le nombre de gènes est égal au nombre total disponible des objets pédagogiques, et un gène est soit 0 si l'objet pédagogique correspondant n'est pas inclus dans la chaîne ou 1 si l'objet pédagogique est inclus dans la séquence d'apprentissage. Dans une valeur réelle encodage de chromosomes, un gène représente une information sélective des métadonnées [Seki et al. 2005].

Population et initialisation de paramètres.

Les paramètres de l'algorithme comme la taille de la population, les probabilités de reproduction et de mutation complètent la spécification de l'AG. Les valeurs de ces paramètres ont un impact considérable sur la qualité du résultat de recherche [Eiben et al. 2007]. La taille de la population varie généralement en fonction de la complexité du problème. Plus la taille de la population est grande, plus il y'a de la chance de trouver une meilleure qualité de solutions. Cependant, l'inconvénient, c'est la lenteur de la vitesse de recherche [Chen 2008]. La taille de la population, croisement et des probabilités de mutation sont généralement fixés à une certaine valeur [Seki et al. 2005, Chen 2008, Huang et al. 2007]; [Chen et al.2008, Acampora et al. 2009]. En effet, la population initiale est habituellement générée de façon aléatoire [Lopes et al. 2009, Seki et al. 2005, Guo et al. 2009, De-Marcos et al. 2009b, Azough et al. 2007].

La sélection des Parent.

Cette étape de sélection décide quels individus sont autorisés à devenir parents et subissent des opérateurs de recombinaison ou mutation. La sélection est probabiliste où les meilleurs individus obtiennent une chance plus élevée [Eiben et al. 2007].

La fonction de fitness.

La fonction fitness est une mesure de la qualité d'un parcours d'apprentissage. Les critères de fitness définissent ce qu'on entend par " l'amélioration " [Eiben et al. 2007]. Chaque auteur détermine une fonction de fitness différente. La plupart des auteurs se concentrent sur la satisfaction d'ordre des contraintes pédagogiques, la régularité du niveau de difficulté d'une manière qui maximise le degré de relation entre objets pédagogiques et l'équivalence des préférences des apprenants.

La recombinaison.

L'opérateur de recombinaison combine des informations de deux parents choisis au hasard dans un ou deux nouvelles progénitures. L'objectif est de produire de meilleurs individus en combinant les caractéristiques souhaitables des deux parents [Eiben et al. 2007]. Différents types de recombinaison ont été appliqués, y compris le croisement en un point [Lopes et al. 2009]; [Hovakimyan et al. 2004, Guo et al. 2009], croisement uniforme [Chen 2008, Huang et al. 2007], croisement en ordre [De-Marcos et al. 2009b] et croisement d'échange de sous tournée [Seki et al. 2005]. Le croisement et la mutation sont généralement effectués d'une manière qui préserve la séquence logique des objets pédagogiques [Seki et al. 2005, Guo et al. 2009, Chen 2008]. L'élimination des doublons pourrait être introduite pour éviter la convergence prématurée [De-Marcos et al. 2009b].

La Mutation.

La mutation est un opérateur aléatoire qui est appliqué à un génotype afin de produire un individu légèrement modifié (mutant) [Eiben et al. 2007]. La mutation est généralement un opérateur de fond qui est appliqué avec une faible probabilité. Le but est de créer de nouvelles voies d'apprentissage qui pourraient ne pas être obtenues en utilisant l'opérateur de recombinaison [Huang et al. 2007]. L'opération de mutation peut être appliquée de différentes manières, y compris la mutation binaire simple [Hovakimyan et al. 2004], la mutation et la mutation d'échange [Chen 2008, De-Marcos et al. 2009b, Huang et al. 2007]. La position de la mutation est basée sur la relation " faiblesse " des objets pédagogiques [Seki et al. 2005].

La sélection du survivant.

Le rôle de la sélection du survivant est de décider qui parmi les individus sont autorisés dans la prochaine génération. Cette décision est déterministe, elle repose habituellement sur la qualité et / ou l'âge d'un individu [Eiben et al. 2007]. Les stratégies fréquemment utilisées de la sélection du survivant sont le remplacement des générations [Hovakimyan et al. 2004] avec la préservation d'élite [Lopes et al. 2009, Seki et al. 2005, De-Marcos et al. 2009b].

La Condition de terminaison.

Les AGs sont des algorithmes stochastiques et il n'y a pas de garanties à atteindre la solution optimale. La condition de terminaison est habituellement une disjonction, atteindre une solution optimale ou satisfaire quelques condition [Eiben et al. 2007]. Cette condition peut être exprimée en termes de seuil de fitness [Hovakimyan et al. 2004, Seki et al. 2005, Azough et al. 2007], le nombre de générations [Chen 2008, Huang et al. 2007], ou une combinaison des deux [De-Marcos et al. 2009b].

3.4 Discussion

La caractéristique la plus intéressante de la technique ACO est son adaptation et sa robustesse où le matériel d'apprentissage est en constante évolution et les apprenants eux-mêmes sont en train de changer aussi. Cependant, d'autres voies d'apprenants suivants ne peuvent pas toujours se traduire par le meilleur chemin. Au lieu de cela, ce pourrait aboutir à tomber dans des optima locaux. Bien que cette approche donne plus d'importance aux besoins individuels des élèves, l'inconvénient est qu'elle aggrave le problème de démarrage à froid, surtout si le nombre d'utilisateurs n'est pas assez grand.

Une autre caractéristique de la technique de l'ACO est l'évaporation des phéromones prévues qui conduit à oublier progressivement la piste des fourmis antérieures qui n'ont pas été guidées et que la phéromone n'a pas encore accumulée. Cela renforce l'adaptation dynamique du système à divers scénarios, tels que la disparition / le changement d'un certain matériel d'apprentissage.

Un AG peut être utilisé pour générer une séquence d'objets pédagogiques appropriée pour un même comportement. Alors qu'un AG est généralement considéré comme un algorithme simple à mettre en œuvre, la difficulté est liée au réglage de ses différents paramètres. La performance de l'algorithme génétique est fortement influencée par ses paramètres qui

peuvent être généralement contrôlées par différentes stratégies. Un autre inconvénient de certaines implémentations de l'AG est le mauvais choix des opérateurs génétiques d'une manière qui produit des séquences illégales de programmes. Par exemple, en utilisant un point croisé et un croisement uniforme

Quant à OEP, c'est une technique qui partage de nombreux avantages et caractéristiques avec ACO. Contrairement à AG, EOP n'a pas à faire face à de nombreux paramètres, comme le souligne [Gil et al.2008], EOP pourrait être mis en œuvre dans des applications plus importantes à l'échelle où un grand nombre d'entrées appartient à différents concepts d'apprentissage dans différents niveaux. Une évaluation comparative montre que le EOP surpasse GA en termes de fitness moyenne obtenue et sa stabilité.

Cependant, la principale difficulté dans la mise en œuvre et l'évaluation de ces systèmes est le problème de démarrage à froid. Car, ces méthodes nécessitent une population suffisamment grande pour être efficace. En outre, elles peuvent être plus lentes à adapter de nouveaux contenus e-Learning, car il faut du temps pour que les apprenants découvrent et essayent de nouveaux contenus d'apprentissage et en conséquence les résultats de leur succès (ou l'échec) pour être pris en compte.

Des centaines de présentations de matériels d'apprentissage en forme d'objets pédagogiques apparaissent chaque jour avec diverses propriétés afin de faciliter la présentation et l'explication de ces concepts sous différents angles et en utilisant divers méthodes pour convenir à différents styles d'apprentissage et des niveaux de difficulté. Les méthodes basées sur les AGs et OEP sont parmi les méthodes d'optimisation les plus utilisées pour rechercher une bonne approximation au problème de personnalisation de l'apprentissage.

Les poids des différents paramètres qui constituent la fonction fitness déterminent leur importance relative dans la formation d'une séquence des objets pédagogiques. Ces pondérations peuvent être distribuées [Wang et al. 2009]. Alternativement, les poids peuvent être fixés par l'équipe pédagogique [Valigiani et al. 2006, Chu et al. 2009] ou par l'apprenant [Seki et al. 2005]. Les méthodes du calcul évolutionnaire peuvent jouer le rôle d'optimiser les poids de ces paramètres pour la personnalisation de l'apprentissage. Par exemple, les poids des paramètres qui influencent la fonction de fitness (Comme la correspondance du niveau de difficulté, le degré de pertinence des cours, les délais, etc.) peuvent être codés comme des gènes dans le chromosome qui le permette d'évoluer. Cela peut être une nouvelle direction de

recherche qui permet d'éviter la surcharge, biais potentiel et les écarts associés à la sélection de ces poids.

3.5 Conclusion

Dans ce chapitre, nous nous sommes concentré sur l'application des algorithmes du calcul évolutionnaires pour la personnalisation de l'apprentissage, bien entendu, celle-ci joue un rôle central dans les EIAH et représente un enjeu de recherche important. La personnalisation de l'apprentissage est considérée comme l'un des problèmes les plus difficiles, sachant que le succès d'un tel système dépend sur sa capacité à s'adapter à la connaissance préalable, les compétences et les besoins de l'apprenant. Nous avons fourni un aperçu du calcul évolutionnaire comme un domaine émergent et nous avons présenté un aperçu sur les méthodes utilisées pour rapprocher les solutions au problème de personnalisation. Il existe quelques approches principales qui ont été utilisées pour réaliser la personnalisation, comme ACO et OEP et AG. Cependant, l'absence d'une méthodologie standard pour évaluer la qualité des solutions produites est une difficulté majeure. Une façon d'approcher est à travers la comparaison des manuel / séquence d'instructeur générés [Chen et al. 2008]. Dans les systèmes d'apprentissage classiques, les objets pédagogiques sont fixes et pré-stocké. Cependant, dans des systèmes ouverts les objets pédagogiques peuvent être dynamiquement ajoutés, déplacés ou supprimés (par exemple, le système pourrait être en mesure pour récupérer un contenu pertinent sur le Web liés aux besoins de l'apprenant). La résolution de cette variante dynamique du problème de personnalisation de l'apprentissage apporterait une solution à l'évolution des besoins des apprenants et permet un certain degré de flexibilité du contenu liée à la dynamique de l'information sur le Web. Toutes les techniques ci-dessus sont des méthodes approximatives, mais certainement des outils précieux pour améliorer la qualité de l'apprentissage.

Chapitre 4

Le calcul de l'ADN

4.1 Introduction

Nous nous sommes intéressés à la méthode du calcul de l'ADN, donc dans ce chapitre nous allons détailler les mécanismes biologiques, à savoir les méthodes, les techniques et les étapes de l'algorithme du calcul de l'ADN ainsi que les différentes utilisations de la méthode pour résoudre des problèmes NP complets.

4.2 L'ADN composant génétique

L'ADN, l'acide désoxyribonucléique, a été progressivement découvert par des scientifiques. Les chercheurs comme Mendel, Morgan, Crick ou Watson ont mis des années avant de parvenir à leurs fins. La découverte de l'ADN ne commence qu'au XIX^e siècle. Cependant, des recherches ont débuté depuis longtemps mais leurs résultats n'ont débouché sur de faux résultats.

- **Structure de l'ADN**

L'**acide désoxyribonucléique (ADN)** est une biomolécule présente dans toutes les cellules ainsi que chez de nombreux virus et qui contient toute l'information génétique, appelée génotype, permettant le développement et le fonctionnement de ces organismes. L'ADN est un acide nucléique, au même titre que l'acide ribonucléique (ARN) [Cours de SVT 2de 2015]. Les acides nucléiques sont, avec les peptides et les glucides, l'une des trois grandes familles de biopolymères essentiels à toutes les formes de vie connues [Cours de SVT 2de 2015].

La plupart des molécules d'ADN sont formées de deux brins antiparallèles enroulés l'un autour de l'autre pour former une double hélice. Chaque **brin** résulte de l'**assemblage de nucléotides**, aux quatre bases azotées qui sont:

- Adénine (A)
- Cytosine (C)
- Guanine (G)
- Thymine (T)

Un nucléotide est formé de 3 molécules :

- Un sucre : le désoxyribose
- Un acide phosphorique
- Une base azotée

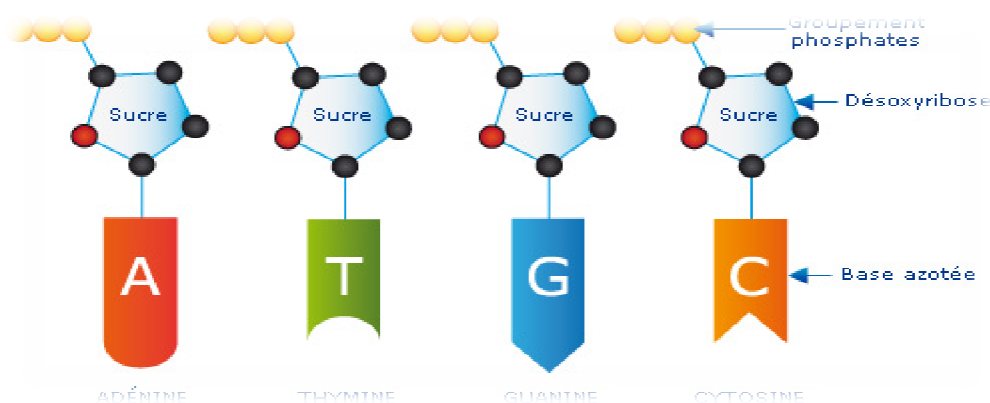


FIG.4.1 –Les bases de l'ADN [Cours de SVT 2de 2015].

Les bases azotées sont liées au désoxyribose qui lui-même est lié à un groupe phosphate. Les nucléotides polymérisés sont liés les uns aux autres par des liaisons covalentes entre le désoxyribose d'un nucléotide et le groupe phosphate du nucléotide suivant, formant ainsi une chaîne où alternent oses et phosphates, avec des bases azotées liées chacune à un ose. L'ordre dans lequel succèdent les nucléotides le long d'un brin d'ADN constitue la séquence de ce brin. C'est cette séquence qui porte l'information génétique. Celle-ci est structurée en gènes, qui sont exprimés à travers la transcription en ARN. Ces ARN peuvent être non codants (ARN de transfert et ARN ribosomique notamment), ou bien codants : dans ce cas, ils sont

traduits en protéines par des ribosomes, la succession des bases azotées sur l'ADN déterminant la succession des acides aminés constituant les protéines issues de ces gènes.

La correspondance entre bases azotées et acides aminés correspond au code génétique, Et l'ensemble des gènes d'un organisme constitue son génome.

Les bases azotées d'un brin d'ADN peuvent interagir avec les bases azotées d'un autre brin d'ADN à travers des liaisons hydrogène, qui déterminent des règles d'appariement entre paires de bases : l'adénine et la thymine s'apparient au moyen de deux liaisons hydrogène, tandis que la guanine et la cytosine s'apparient au moyen de trois liaisons hydrogène. Lorsque les séquences des deux brins sont complémentaires, ils peuvent s'apparier en formant une structure bicaténaire hélicoïdale caractéristique, la double hélice d'ADN.

Cette double hélice est bien adaptée au stockage de l'information génétique. La chaîne oses-phosphates est résistante aux réactions de clivage. De plus, l'information est dupliquée sur les deux brins du double hélice, ce qui accroît sa pérennité. Enfin, cette information peut être copiée à travers un mécanisme appelé réplication de l'ADN au cours duquel une double hélice d'ADN est recopiée fidèlement en une autre double hélice portant la même information. C'est en particulier ce qu'il se passe lors de la division cellulaire : chaque molécule d'ADN de la cellule mère est répliquée en deux molécules d'ADN, chacune des deux cellules filles recevant ainsi un jeu complet de molécules d'ADN identiques l'un à l'autre.

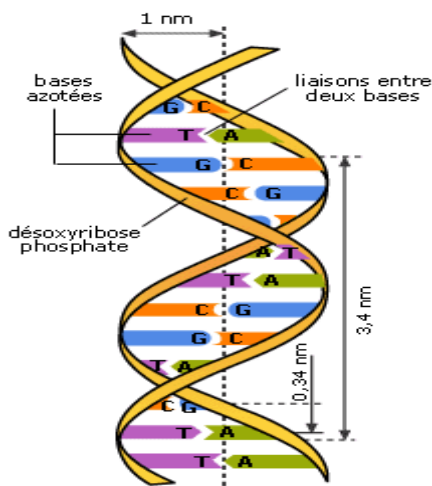


FIG.4.2 – La structure double hélice d'ADN [Cours de SVT 2de 2015].

Dans les cellules, l'ADN est organisé en structures appelées chromosomes. Ces chromosomes ont pour fonction de rendre l'ADN plus compact à l'aide de protéines, notamment d'histones, qui forment, avec les acides nucléiques, une substance

appelée chromatine. Les chromosomes participent également à la régulation de l'expression génétique en déterminant quelles parties de l'ADN doivent être transcrites. Chez les eucaryotes (animaux, plantes, mycètes et protistes), l'ADN est essentiellement stocké dans le noyau des cellules, avec une fraction d'ADN présent également dans les mitochondries ainsi que, chez les plantes, dans les chloroplastes. Chez les procaryotes (bactéries et archées), l'ADN est stocké dans le cytoplasme. Chez les virus qui en contiennent, l'ADN est stocké dans la capsid. Quel que soit l'organisme considéré, l'ADN est transmis au cours de la reproduction : c'est le support de l'hérédité.

La modification de la séquence des bases d'un gène peut conduire à une mutation génétique, laquelle peut être sans conséquence pour l'organisme ou, au contraire, être incompatible avec sa survie. À titre d'exemple, la modification d'une seule base (polymorphisme nucléotidique) d'un seul gène celui de la β -globine, une sous-unité protéique de l'hémoglobine A du génotype humain est responsable de la drépanocytose, une maladie génétique parmi les plus répandues.

En effet, l'étude de l'ADN de cellules d'un même organisme et d'organismes d'espèces différentes révèle quelques caractéristiques. Chaque ADN d'une espèce quelconque montre une quantité équivalente (mais différente d'une espèce à l'autre) de nucléotides A et T, ainsi que de nucléotides C et G. En revanche, ces proportions sont constantes pour deux molécules d'ADN issues de deux cellules différentes d'un même organisme. Ces données suggèrent que les nucléotides forment des « couples » au sein de la molécule d'ADN : l'adénine et la thymine sont complémentaires, ainsi que la guanine et la cytosine. La complémentarité des nucléotides se trouve vérifiée dans l'organisation en double hélice. Chaque brin est ainsi complémentaire de l'autre. Cette structure de l'ADN fut mise en évidence par Watson et Crick, en 1953.

4.3 Les outils pour le calcul de l'ADN

La majorité des expériences de la biologie moléculaire utilisent les quelques outils de base que nous verrons ci-dessous. Leur compréhension est nécessaire, et presque suffisante, pour travailler dans un laboratoire. Nous présentons d'abord ces outils qui sont les enzymes, les électrophorèse, et le PCR (Réaction en chaîne de Polymérase).

4.3.1 Les enzymes

La biologie moléculaire possède des instruments, ces instruments sont des enzymes extraits du monde vivant. Les enzymes sont de véritables magiciens nanométriques, et on ne comprend pas toujours exactement leurs fonctionnements, mais on sait s'en servir. C'est un peu comme conduire une voiture : la plupart d'entre nous ne savons pas exactement ce qu'il y a sous le capot, même si les principes générales nous sont connus [Houchmandzadeh 2004]. Cependant, nous savons parfaitement nous en servir pour aller d'un point à un autre. Notons qu'un enzyme est une protéine agissant comme catalyseur de nombreuses réactions biochimiques. Actuellement, près de 1 000 enzymes différents ont été identifiés. La science qui étudie les enzymes est l'enzymologie. Les enzymes sont des substances extrêmement efficaces.

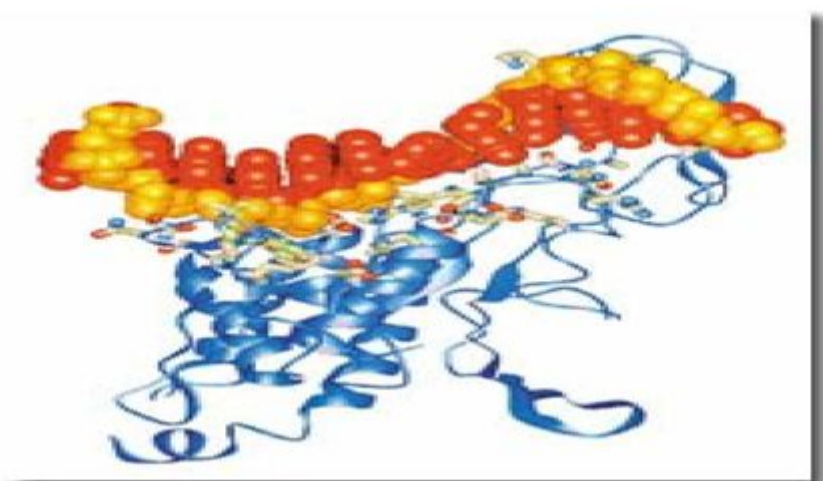


FIG.4.3– Enzymes autour de l'ADN durant la réplication et la synthèse protéique. Les zones rouges et jaunes dans l'image montrent les enzymes qui agissent avec l'ADN [Yahyia 2005].

4.3.2 Les polymérases

Comment fabriquer une nouvelle molécule d'ADN ? Watson et Crick [Watson et al.1953] ont mentionné dès le début, dans leur article de 1953 sur l'ADN, que la forme même de cette molécule peut être une indication pour sa synthèse : chaque brin devrait servir de moule pour fabriquer son complémentaire. Nous connaissons aujourd'hui, grâce au travail acharné de nombreux biochimistes, les enzymes qui dirigent ces actions de synthèse [Houchmandzadeh 2004].

Le premier enzyme de synthèse est découvert en 1955. Elle s'appelle PNPase, elle synthétise l'ARN en incluant de façon aléatoire les nucléotides présents dans la solution [Houchmandzadeh 2004]. Nirenberg et Matthaei, dans leur course pour briser le code génétique, l'ont utilisé pour synthétiser des ARN polyU en 1961 et établir ainsi la relation entre le triplet UUU et l'acide aminé phenylalanine. [Houchmandzadeh 2004].

Mais le maître absolu, l'enzyme de réplication du génome (nommé l'ADN polymérase) a été découvert en 1958 après le travail minutieux des biochimistes Arthur Kornberg et Robert Lehman de l'Université de Stanford. La réaction de base peut s'écrire $(ADN)_n + NTP \rightarrow (ADN)_{n+1} + P_i + P_i$. [Houchmandzadeh 2004] Une chaîne d'ADN de longueur n incorpore un nucléotide (présent en solution comme NTP un nucléotide triphosphate) et s'allonge d'une unité. Cette réaction est énergétiquement favorable. Le travail de l'ADN polymérase est de choisir le bon nucléotide parmi les 4 (ATP, CTP, TTP, GTP) présent dans la solution, en lisant le brin complémentaire. Nous verrons la biochimie de ce processus plus bas.

Ce qui est essentiel à retenir ici est que l'ADN polymérase ne fonctionne qu'en présence d'un brin d'ADN utilisé comme modèle. L'autre élément important est que l'ADN polymérase ne fonctionne que s'il y a déjà une séquence synthétisée, et correctement hybridée au brin modèle derrière lui ! La longueur minimale de cette séquence est de trois à quatre bases.

L'ADN polymérase est là pour dupliquer le génome et le transmettre à la génération d'après. Il doit donc être le plus précis possible. Son taux d'erreur est de l'ordre de 10^{-19} , c'est à dire que tous les milliards de base, il commet une erreur. Essayez donc de recopier 1000 livres en ne commettant qu'une seule faute pour avoir une idée de la précision de cette machine [Houchmandzadeh 2004]. Les considérations énergétiques - un mauvais base pairing coûte deux à trois liaisons hydrogènes - donne un taux d'erreur de l'ordre de 10^{-4} .

Il existe d'autres mécanismes de relecture qui permettent d'augmenter la précision. On peut constater que l'existence d'une séquence déjà hybridée derrière participe à ces mécanismes. En effet, supposons que l'ADN polymérase inclue un mauvais nucléotide et avance d'une séquence. La séquence qu'il a alors derrière lui n'est pas correctement hybridée et arrête le fonctionnement de l'ADN-polymérase. L'enzyme doit alors faire marche arrière, enlever la mauvaise base, et recommencer.

Aussi surprenant que cela puisse paraître, c'est vraiment comme cela que ça se passe. In vivo, pour dupliquer une molécule d'ADN, il faut séparer à un endroit les deux brins d'ADN et commencer à dupliquer chaque brin séparément. Le point de départ s'appelle l'origine de réplication. In vitro le problème ne se pose pas, puisqu'en élevant la température (autour de 60 à 90C) on sépare les deux brins.

L'ARN polymérase est un cousin très proche de l'ADN polymérase, à cette différence qu'il lit un brin d'ADN pour le copier en ARN. Il fut découvert en 1960 par plusieurs groupes de façon presque simultanée. Notons qu'ARN-polymérase est nettement moins précis que l'ADN-polymérase et ne possède pas plusieurs des mécanismes de correction d'erreurs de ce dernier. Enfin, le dernier enzyme de cette classe est appelé reverse transcriptase et qui polymérise l'ADN en lisant l'ARN ! Cet enzyme est propre aux rétrovirus dont le matériel génétique est l'ARN et non l'ADN.

4.3.3 Ligase

Cet enzyme, découvert en 1967, est utilisé pour coller deux brins d'ADN en créant une liaison covalente entre l'extrémité 3' de l'un et 5' de l'autre [Houchmandzadeh 2004]. Les deux morceaux doivent être immobilisés sur le brin complémentaire (voir Figure 4.4). In vivo cet enzyme est utilisé lors de la duplication d'ADN. Comme le chromosome est assez long, souvent sa duplication commence à plusieurs endroits en parallèle. La ligase intervient pour attacher ces copies partielles les unes aux autres. L'action de ligase est de former une liaison covalente phosphodiester entre deux morceaux d'ADN, immobilisés sur le brin complémentaire qui lui, est intact [Houchmandzadeh 2004].

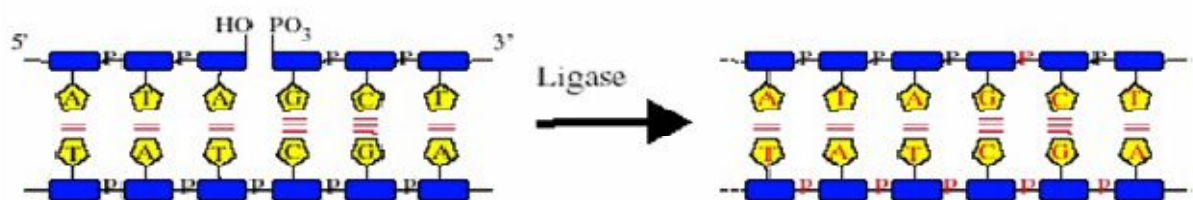


FIG.4.4– L'action de ligase. [Yahyia 2005].

4.3.4 Enzymes de restrictions

Les enzymes de restriction ont été découverts à partir de 1973. Ils sont capables de reconnaître spécifiquement une courte séquence, de 4 à 10pb, et de cliver l'ADN au site reconnu. Ils permettent de fragmenter l'ADN en segments de taille réduite, ou de le couper à tel ou tel site désiré (voir figure 4.5). Certains enzymes coupent le site en son milieu et produisent deux fragments dont les extrémités sont franches. Cependant, la plupart réalisent une coupure dissymétrique : on parle dans ce cas d'extrémités cohésives (chaque fragment possède une chaîne qui dépasse l'autre de quelques bases). Plusieurs centaines de ces enzymes ont été caractérisés : ils reconnaissent une grande variété de sites de coupure.

Les enzymes de restriction sont utilisés pour établir une carte de restriction de toute molécule d'ADN que l'on souhaite caractériser [Hofmeister et al. 2002]. Cela consiste à déterminer l'ordre des sites de restriction le long de cette molécule, qui vont produire, après "digestion enzymatique" de cette molécule, des fragments de tailles différentes dont la taille pourra être définie par électrophorèse.

Les enzymes de restriction appartiennent à la classe des endonucléases, c'est-à-dire des enzymes capables de cliver les liaisons phosphodiester entre deux nucléotides à l'intérieur d'un acide nucléique. Les endonucléases se différencient des exonucléases qui dégradent la molécule d'ADN à partir de l'une de ses extrémités (3' ou 5').

L'action des enzymes de restriction consiste à couper un double brin d'ADN à un endroit. Montrons l'exemple de EcoRI(une enzyme de restriction produite par la souche bactérienne Escherchia Coli) et son site de reconnaissance. Cette enzyme produit des bouts collants [Houchmandzadeh 2004].

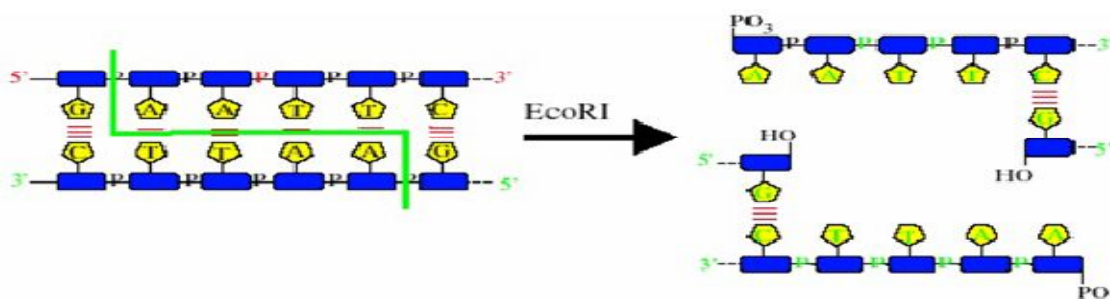


FIG.4.5– Actions des enzymes de restriction [Yahyia 2005].

4.3.5 Electrophorèse

Le principe de l'électrophorèse, qui consiste à faire migrer dans un certain milieu des molécules chargées en présence d'un champ électrique, date des années trente. Cependant, elle est devenue une technique très répandue d'analyse et de séparation d'ADN (on l'utilisait initialement pour la migration de protéines) avec l'utilisation de gel d'agarose (1973)[Lionnet et al.2005].

L'ADN est déposé dans un gel d'agarose en présence d'un marqueur qui devient fluorescent lorsqu'il s'intercale dans l'ADN. L'ADN est attiré vers l'électrode + (phosphates chargés négativement), mais les mailles du gel d'agarose permettent de séparer les molécules en fonction de leur taille [Lionnet et al. 2005]. Cette méthode permet de vérifier la taille de fragments d'ADN, et de les séparer ensuite. On peut aussi l'utiliser pour détecter la présence d'une séquence spécifique dans des fragments donnés. C'est la technique du Southern blot [Lionnet et al. 2005], après l'électrophorèse d'ADN double brin, le gel est transféré sur une membrane et l'ADN double brin dénaturé en simple brin. Cette étape permet de fixer l'ADN sur un support solide.

La réplique du gel sur la membrane est ensuite mélangée à une sonde, c'est-à-dire un simple brin d'ADN portant un marqueur (radioactif ou fluorescent). Après rinçage, seules les bandes du gel ayant intégré la séquence voulue retiennent la sonde, et on peut les révéler (par autoradiogramme ou simple photo dans le cas d'un marqueur fluorescent). Cette technique est utilisable pour l'ADN, ou les protéines (Western) [Lionnet et al. 2005]. (Voir figure 4.6)

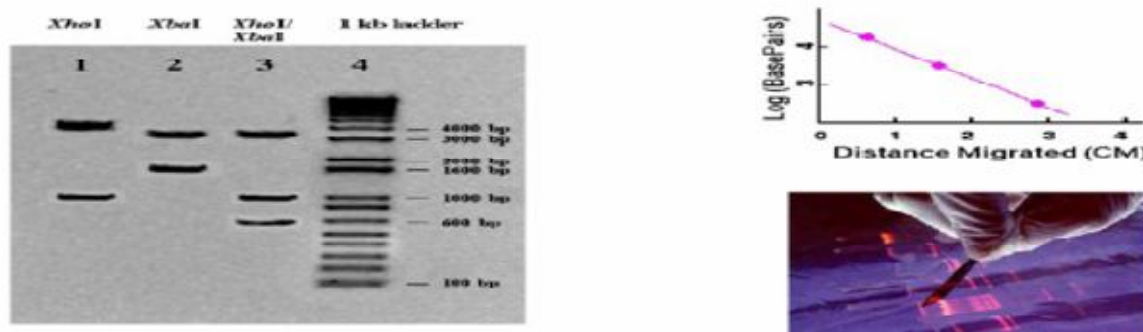


FIG.4.6– Electrophorèse [Lionnet et al. 2005].

Gauche : Photo du résultat d'électrophorèse d'ADN marqué par fluorescence sur gel d'agarose. Les trois premières colonnes montrent le même fragment coupé par les enzymes XhoI et/ou XbaI. La dernière colonne est une solution contenant des fragments de tailles connues. Ainsi, on peut mesurer la taille et séparer différentes molécules d'ADN ; droite haut : La distance de migration varie comme le logarithme de la taille de l'ADN ; droite bas : on peut extraire la bande voulue du gel et en récupérer l'ADN débarrassé des fragments parasites ; [Lionnet et al. 2005].

Pour résumer, l'électrophorèse en gel est un appareil qui nous permet de séparer un mélange de molécules d'ADN en fonction de leurs tailles. Si les pores du gel sont très petites, le pouvoir de résolution peut être d'une seule base : on peut séparer une séquence de 45 bases d'une séquence de 46, et c'est exactement cela qu'on utilise pour séquencer l'ADN. Le problème avec cela est qu'on ne peut pas faire cela sur des longues séquences d'ADN, disons plus que 300 bases. En général, on n'a pas besoin d'autant de résolution et on prépare des gels avec des pores plus grands. Savoir préparer ses gels est pour le biologiste moléculaire ce que savoir préparer ses couleurs est pour le peintre.

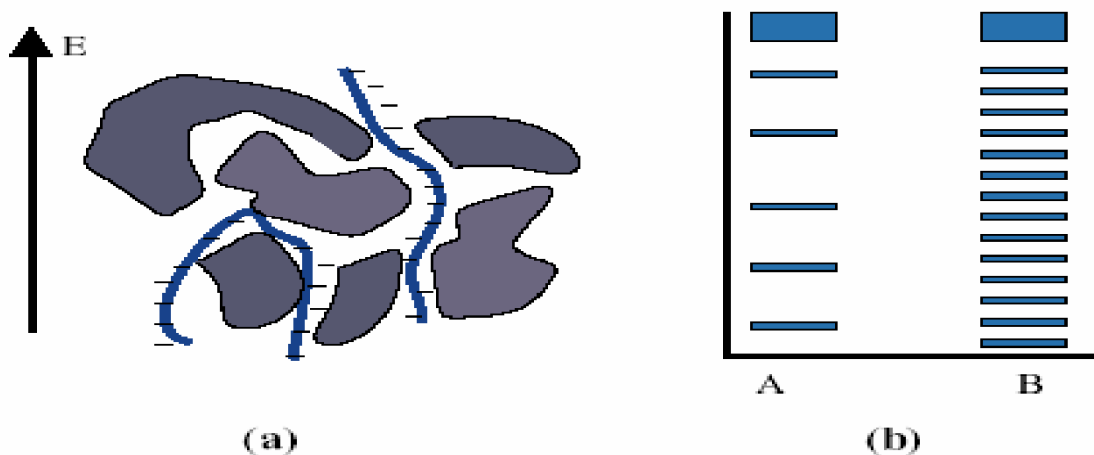


FIG.4.7– Le principe de l'électrophorèse en gel.

Pour le principe de l'électrophorèse en gel. (a) le gel est montré à l'échelle de quelques nanomètres, où les chaînes d'ADN (en bleu) migrent dans le champ. En bas à gauche, on montre le cas d'une molécule piégée : les forces s'exerçant sur ses deux bouts sont à peu près égales, et elle peut rester bloquée dans cette position. (b) l'appareil de l'électrophorèse

schématisé, d'une largeur d'environ 20 à 40 cm. Le haut et le bas sont branchés à des électrodes pour maintenir une différence de potentielle et créer un champ électrique à l'intérieur du gel. Au temps zéro, le mélange d'ADN (ou ARN, ou protéine) est chargé en haut (colonne A). Au bout d'un certain temps (de l'ordre de l'heure, toujours trop long de toute façon), les divers constituants migrent sous l'effet du champ, mais à des vitesses différentes, et se trouve donc à divers position du gel. Comme on ne connaît pas à priori leurs tailles, on charge toujours en parallèle (colonne B) un mélange de molécules d'ADN de longueurs connues. La colonne B constitue donc la règle qui nous permet de lire la taille des diverses bandes de la colonne A. Toutes ces molécules sont bien sûr transparentes, et il existe divers méthode pour les visualiser : soit on les a rendu radioactives par avance, et il suffit alors de poser une plaque photographique sur le gel, soit on doit, après la fin de la migration, utiliser des marqueurs visibles qui se lient spécifiquement à l'ADN (ou ARN, ou...). La taille des pores est un paramètre ajustable et dépend des besoins de l'utilisateur en pouvoir de résolution [Houchmandzadeh 2004].

4.3.6 PCR (Polymerase Chain Reaction)

L'amplification est l'art de copier des millions de fois un fragment d'ADN donné. La technique qui permet de faire une amplification exponentielle est appelée en anglais PCR (Polymerase Chain Reaction). Elle est d'une telle simplicité que l'on s'étonne qu'il ait fallu attendre jusqu'à 1985 pour qu'elle soit découverte. Cette technique a révolutionné la biologie moléculaire. Supposons que nous avons un fragment d'ADN double brin (ADNdb ou dsDNA en anglais) dont la séquence est connue, ou tout au moins la séquence de ses extrémités (quelques paire de base de chaque coté suffisent). Comment fabriquer une copie exacte in vitro ? Evidemment, nous devons utiliser l'ADN polymérase, mais l'ADN-Pol ne fonctionne que s'il y a déjà quelques séquences polymérisées derrière lui. Pour cela, on utilise des amorces (primers en anglais), des petites séquences de 4-5 pairs de base qui sont complémentaires aux extrémités 3' de chaque brin.

Voilà comment on procède [Houchmandzadeh 2004] : On démarre avec une solution contenant l'ADN-Pol, beaucoup de NTP (les quatre bases, évidemment) et beaucoup d'amorces, dans laquelle on balance notre fragment d'ADNdb. On chauffe alors notre solution autour de 90C, ce qui a pour effet de séparer les deux brins. On refroidit alors la

solution aux alentours de 60C. Les amorces peuvent alors s'hybrider avec les extrémités 3' et l'ADN-Pol peut commencer son travail en polymérisant un brin tout neuf complémentaire du brin qu'il lit. Nous sommes donc maintenant en possession de deux ADNdb, strictement similaire à l'original avec lequel on avait commencé (voir figure 4.7). Il suffit maintenant de chauffer à nouveau à 90C et recommencer le cycle pour avoir 4 copies et ainsi de suite. Le nombre de copie final sera 2^n , où n est le nombre de cycles. Le PCR met grandement à profit la nécessité pour l'ADN-Pol de démarrer avec des amorces déjà hybridées. Si nous avons plusieurs fragments d'ADN différent dans la solution, seul le fragment qui correspond aux bonnes amorces sera amplifié. Par exemple, pour déterminer si un patient a été infecté par un virus ou une bactérie donné (bien avant que les symptômes de la maladie apparaissent), on purifie l'ADN issue du patient, et on amplifie avec des amorces spécifiquement préparées pour le génome du pathogène [Houchmandzadeh 2004].

Des traces infimes de l'ADN du pathogène peuvent ainsi être détectées. Nous avons parlé des températures de 60 à 90C. La question qui se pose est comment la protéine ADN-Pol peut rester stable à ces températures, sans parler de son fonctionnement ? En fait, cela était la clef de la technique. Dans les versions primitives du PCR, ADN-Pol était ajouté à chaque cycle, pendant la phase basse température. Mais il existe des bactéries qui ne vivent que dans des conditions extrêmes, proches des sources chaudes au fond des océans et qui possèdent des versions thermostable de l'ADN polymérase. C'est la purification de ces polymérases (un des premiers et plus fameux s'appelle Taq-polymérase) qui a réellement lancé le PCR [Houchmandzadeh 2004]. Le PCR a bien sûr ses limites. On ne peut amplifier des fragments de plus que quelques kilobases pour deux raisons :

- in vitro, les polymérases manquent quelques mécanismes de correction d'erreur. Taq par exemple a un taux d'erreur de 10^{-4} (à comparer à 10^{-9}).
- La polymérase se détache de l'ADN au bout de quelques kilo bases. In vivo il existe des mécanismes qui le gardent accroché. Beaucoup d'efforts technologiques ont permis de pousser ces limites de plus en plus loin, mais on en est là de nos jours. Nous avons dit également plus haut que les séquences des extrémités doivent être connues. Cela est souvent le cas, puisque les fragments d'ADN ont été produits en « coupant » un long fragment par des enzymes de restriction. Les extrémités du fragment à amplifier sont donc ceux qui sont reconnu par l'enzyme de restriction utilisé. Enfin, notons que nous pouvons amplifier non pas tout le fragment d'ADN, mais seulement une fenêtre à l'intérieur qui nous intéresse. Il suffit pour cela que l'on utilise des amorces correspondant aux extrémités de la fenêtre en question.

Le principe du PCR. A basse température, les deux brins du fragment d'ADN sont liés par des liaisons hydrogènes. On les sépare en les portant à haute température. En refroidissant ensuite, on permet aux amorces de s'hybrider (former des liaisons hydrogènes) avec les extrémités 3' de chaque brin. L'ADN polymérase peut alors utiliser ces amorces pour polymériser un nouveau brin complémentaire de celui qu'il lit [Houchmandzadeh 2004].

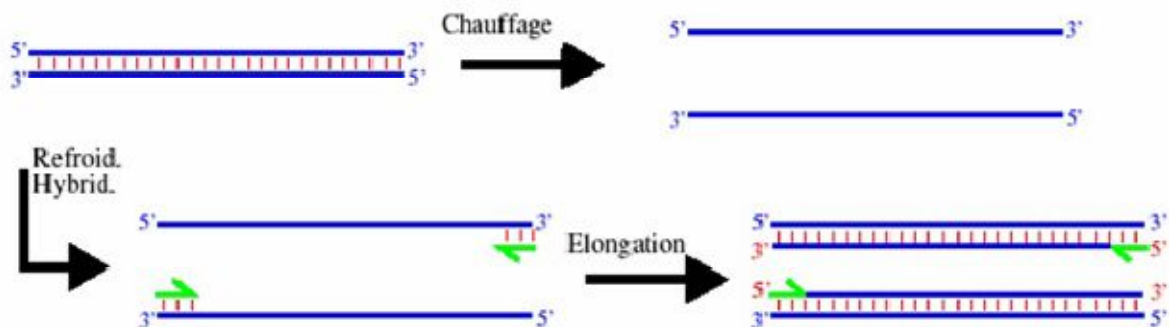


FIG.4.8– Le principe du PCR [Yahyia 2005].

4.4 De l'ADN biologique vers les ordinateurs à ADN

L'ordinateur marque seulement l'avancée la plus récente dans l'histoire des instruments de calculs : leur apparition n'en constitue ni la première ni la dernière étape. Ils ont, d'ailleurs eux aussi leurs limites : limites des quantités qu'ils peuvent stocker ; limites des vitesses d'exécution qui sont fixées par la loi de la physique et seront prochainement atteintes. La plus récente tentative propose de remplacer, une fois de plus, les outils de calcul : d'utiliser, en lieu et place des instruments électroniques, des instruments biologiques.

Exposés pour la première fois en 1994 par Leonard Adelman, les calculs(ou encore bio moléculaire) constituent un nouveau paradigme de calcul. S'appuyant sur la manipulation bio molécule pour résoudre des problèmes mathématiques, ils considèrent les processus naturels comme des modèles. En quoi consistent-ils ?

En effet, en codant des données sur des brins d'ADN et en utilisant les outils de la biologie moléculaire, nous allons pouvoir effectuer des opérations mathématiques. Cela requiert un milliard de fois moins d'énergie que les calculs électroniques traditionnels, tout en stockant

des données sur mille milliards de fois moins d'espace. Qui plus est, le calcul avec l'ADN est hautement parallèle : en théorie, des milliards de molécules sont soumises à des réactions chimiques c'est-à-dire effectuent des calculs simultanément.

Pour Leonard Adelman, ces deux processus ne sont pas seulement similaires : grâce aux progrès de la biologie moléculaire, la biologie peut être utilisée pour faire des mathématiques. C'est donc ce qu'il a fait lui-même pour résoudre un difficile problème de calcul. Plus précisément, le cas à sept points du problème hamiltonien.

Quant à l'ordinateur à ADN, est une des voies non électroniques actuellement explorées pour résoudre des problèmes combinatoires. Il ne prétend pas à la généralité et à la flexibilité d'un ordinateur général. Il s'agit plutôt d'un dispositif spécialisé comme peut l'être un processeur graphique, une carte son ou un convolveur. Son principe, énoncé par Leonard Adleman en 1994, « consiste à coder une instance du problème avec des brins d'ADN et à les manipuler par les outils classiques de la biologie moléculaire pour simuler les opérations qui isoleront la solution du problème, si celle-ci existe. »

En utilisant des fragments de brins d'ADN, on peut coder les contraintes d'une recherche sous forme d'enzymes. Dans un processus d'assemblage et de duplication de bases désoxyribonucléiques, les fragments ne répondant pas aux contraintes du problème sont éliminés par ces enzymes. En fin de processus, il ne reste plus que des chaînes ADN contenant la solution au problème cherché.

Un système de calcul utilisant de l'ADN s'appuie sur des mécanismes de codage fondamentalement différents de ceux de l'ordinateur conventionnel : dans nos machines classiques, c'est la manipulation de charges électriques portés par des électrons au sein de dispositifs de commutation électroniques (transistors) qui matérialise l'information codée sous une forme binaire. Avec les ordinateurs à base d'ADN, l'information est traduite en termes d'unités chimiques de l'ADN. Le principe du calcul avec un ordinateur à base d'ADN, consiste à synthétiser des séquences d'ADN particulières et de les laisser réagir dans un tube à essai.

– **Avantage**

- Réalisation des calculs massivement parallèles : plusieurs millions ou milliards de molécules qui interagissent entre elles.
- Les calculs de l'ADN peuvent utiliser un milliard de fois moins d'énergie qu'un ordinateur électronique.

– **Inconvénients**

- lenteur d'exécution

En effet, même les ordinateurs électroniques ont leurs limites: Il ya une limite à la quantité de données qu'ils peuvent stocker, et les lois physiques qui dictent les seuils de vitesse qu'ils vont bientôt atteindre. La tentative la plus récente à briser ces barrières est de remplacer, une fois de plus, les outils pour effectuer des calculs biologiques au lieu de ceux d'électrique.

TAB.4.1– Comparaison entre ordinateurs à ADN et ordinateurs conventionnels.

Caractéristique	Ordinateurs à ADN	Ordinateurs conventionnels
Supports de stockage	Acides nucléique	Semi-conducteurs
Capacité mémoire	Très élevée	Elevée
Opérateurs	Opérateurs bio moléculaires	Opérateurs logiques (AND, OR, NOT)
Opération	Simultanés (parallèle)	Bit (séquentiel)
Vitesse de chaque opération	Lent	Rapide
Processus	Stochastique	Déterministe

4.5 Exemples de résolutions de Problèmes combinatoires

Les problèmes connus comme des problèmes d'optimisation combinatoire, on été pratiquement résolus par la méthode du calcul de l'ADN. Parmi eux on peut citer le problème du chemin hamiltonien, le problème SAT et le problème du sac à dos.

4.5.1 Le problème du chemin hamiltonien

Le cas à sept points du problème du « chemin hamiltonien orienté », l'auteur dans [L.M. Adelman, 1998] a réussie à le résoudre. L'algorithme utilisé suit les étapes suivantes :

Dans l'étape de codage, chaque point du graphe est codé par un brin d'ADN constitué de vingt bases prises au hasard, puis pour chaque flèche, un oligonucléotide est généré dont les dix premières bases sont complémentaires des dix dernières de son point d'origine et les dix dernières correspondent aux dix premières du point suivant. Après, les séquences complémentaires vont s'apparier bout à bout par l'enzyme ligase avec les séquences d'ADN correspondantes.

Dans l'étape de l'amplification, le résultat obtenu est amplifié par PCR (Polymerase Chain Reaction) en utilisant comme amorces des oligonucléotides. La réaction amplifie et retient ainsi les molécules codant les chemins ayant le nœud de départ et le nœud d'arrivée.

Dans l'étape de séparation, une électrophorèse sur gel d'agarose permet de séparer et de recueillir les brins d'ADN de longueur désirée. Ainsi les produits de la PCR codant ce chemin vont comporter $7 \times 20 = 140$ paires de bases.

Dans l'étape suivante, une purification excessive pour chaque point de départ et d'arrivée est effectuée. Il s'agit, à chaque fois, de sélectionner des brins d'ADN codant pour un point donné en utilisant des billes magnétiques : seuls les brins d'ADN contenant la séquence recherchée sont retenus.

Enfin, dans la dernière étape, la vérification des molécules codant pour un chemin hamiltonien est effectuée, en procédant par PCR. La première PCR amplifie les résultats de l'étape précédente et confirme ou non l'existence d'une telle molécule. Si cette molécule existe, une seconde PCR confirme la présence des points internes du graphe en utilisant comme amorces les séquences d'oligonucléotides complémentaires de chacun de ces points.

Ainsi le problème est résolu par Adelman [L.M. Adelman, 1998], il ne pouvait être traité que sur un ordinateur non déterministe. C'est-à-dire un ordinateur capable de poursuivre un nombre illimité de calculs indépendants en parallèle.

4.5.2 Le problème du sac à dos

Dans [Sanchita et al. 2008] les auteurs ont proposé une stratégie de codage des items avec l'ADN. Cette stratégie est basée sur le codage des items et des profits. Le but principal est de maximiser le profit des éléments dans le sac à dos sans dépasser sa capacité.

Le principe du codage est le suivant : pour chaque item est utilisé deux brins d'ADN pour le coder, c'est-à-dire pour un item i tel que $i(1 \leq i \leq n)$: S_{i1} et S_{i2} sont utilisés pour le coder. S_{i1} est composé de trois parties : $S_{i1}' S_{i1}'' S_{i1}'''$. Où S_{i1}' représente le profit de cet item, et S_{i1}'' est calculé comme suit : $S_{i1}'' = (W_i - P_i)/2$ et $S_{i1}''' = W_i - P_i - (W_i - P_i)/2$. Quant à S_{i2} est représenté par le complément à l'inverse de la partie du centre de S_{i1} .

Une réplication des brins d'ADN dans la phase finale prend la combinaison de chaque ADN pour former le double brin d'ADN afin de trouver la solution optimale.

4.5.3 Le problème SAT

Dans [Wang et al. 2008] les auteurs ont proposés une méthode pour résoudre le problème SAT. L'idée de base est d'utiliser la technique de la réaction de chaîne de Ligase (Ligase Chain Reaction en anglais, en se basant sur quatre opérations : Ligature, amplification, fractionnement et fusion des brins d'ADN afin d'obtenir des solutions correctes. Ainsi, les auteurs ont éliminé le besoin de construire une solution complète de bibliothèque d'ADN afin de réduire le nombre de brins d'ADN requis dans le calcul de l'ADN, et rendre cette approche plus extensible pour résoudre de larges problèmes d'optimisation combinatoires. De plus, le graphe proposé pour la construction des chemins pour résoudre le problème SAT apporte quelques nouveaux concepts dans la théorie des graphes, Tout d'abord, chaque bord est intégré avec une porte logique qui détient une affectation de valeur pour le nœud correspondant et vérifie si un chemin est autorisé ou non, et les chemins sont générés étape par étape.

Le graphe G modélisant le problème SAT, a n nœuds. Dans la phase de calcul, l'ensemble de l'ADN restant dans le tube à essai correspond au graphe G dont Chaque nœud a deux bords à chaque voisin, et chaque arête comprend une porte logique, et toutes les portes logiques contiennent une valeur interne fixe de 0 ou 1, qui est conceptuellement nouvelle et différente des portes logiques classiques. Une Formule SAT est considérée comme une condition de

retenue pour générer des chemins à travers les nœuds. Tous les chemins partiels possibles qui commencent au début et à la fin, au plus tard nœud traité sont générés étape par étape pour chaque littéral et chaque clause. Un chemin est appelé un vrai chemin et réservé, si un littéral dans une clause est satisfait par la valeur du nœud de grille correspondant, sinon le chemin est considéré comme une fausse voie et enlevé.

Tous les vrais chemins pour tous les littéraux d'une clause sont fusionnés pour former un ensemble de véritables chemins pour que la clause; les vraies voies sont ensuite étendues et / ou tamisées pour la clause suivante.

Par conséquent, tous, et seulement, chemins partiels qui satisfont toutes les clauses qui ont été traitées sont générés.

4.6 Synthèse

L'idée principale du calcul de l'ADN était l'encodage des données dans les brins d'ADN et l'utilisation d'outils de la biologie moléculaire pour exécuter les opérations de calcul. Outre la nouveauté de cette approche, l'informatique moléculaire a le potentiel d'être plus performante que les ordinateurs électroniques. Par exemple, le calcul de l'ADN peut utiliser un milliard de fois moins d'énergie qu'un ordinateur électronique, en stockant des données dans un espace trillion de fois inférieure. De plus, le calcul de l'ADN est hautement parallèle: En effet, il pourrait y avoir des milliards et des milliards de molécules d'ADN subissant des réactions chimiques, et qui effectuent des calculs en même temps. Malgré la complexité de cette technologie, l'idée derrière l'ADN informatique suit d'une simple analogie entre les deux processus suivants, une biologique et une mathématique.

Les progrès de la biologie moléculaire permettent d'utiliser la biologie pour faire des calculs mathématiques. Plus précisément, les chaînes d'ADN peuvent coder l'information tandis que divers techniques de la biologie moléculaire effectuent de simples opérations. Ces possibilités pratiques de coder des informations dans une séquence d'ADN et la manipulation de brins d'ADN ont conduit Adleman pour résoudre le problème du chemin hamiltonien.

4.7 Conclusion

A travers ce chapitre, on peut conclure que : L'acide désoxyribonucléique (ADN) est la molécule qui contient toutes les informations génétiques. L'ADN est exactement le même

dans toutes les cellules de notre corps, mais unique à chaque individu (à l'exception des jumeaux identiques qui ont exactement le même profil génétique). Les multiples applications scientifiques et technologiques impliquant l'étude de l'ADN font appel à différentes techniques. Parmi ces techniques les plus répandues, la technique de l'électrophorèse, qui permet de séparer des molécules en fonction de leur taille et de leur charge en utilisant un courant électrique. On peut ainsi analyser et purifier dans un milieu gélifié (gel d'agarose, gel de polyacrylamide...) l'ADN, l'ARN, les protéines. La technique de PCR, l'amplification, ou l'art de copier des millions de fois un fragment d'ADN donné, est une technique extraordinaire, elle permet d'amplifier des séquences d'ADN de manière spécifique et d'augmenter de manière considérable la quantité d'ADN dont on dispose initialement. Elle nécessite de connaître la séquence des régions qui délimitent l'ADN à amplifier. Ces séquences serviront à synthétiser des amorces oligonucléotidiques complémentaires. Ces oligonucléotides serviront à délimiter la portion d'ADN à amplifier. L'ADN polymérase les utilisera comme amorces.

Chapitre 5

Modélisation de la personnalisation de l'apprentissage : les objets pédagogiques

5.1 Modèle global de la méthode proposée

Le but de cette thèse, comme il a été indiqué dans les sections précédentes, est d'élaborer un système de personnalisation des objets pédagogique pour différents apprenants. Afin d'atteindre cet objectif, nous avons commencé par mettre en avant une philosophie de notre système de personnalisation. Nous entendons par philosophie du système, le modèle général de la personnalisation proposée. L'architecture décrite dans la figure 5.1, représente l'idée générale de notre méthode.

Cette architecture descriptive de la méthode est basée essentiellement sur trois aspects, qui constituent notre contribution et qui sont liés à notre objectif qui est la personnalisation de l'apprentissage, la conception d'un profil apprenant (1), quels sont les éléments à prendre en considération dans un profil apprenant et quelles sont ces composants ?, et aussi la conception d'un graphe constituant les objets pédagogiques (2), c'est à dire quelles sont les relations pertinentes qui peuvent être élaborées dans un graphe basé objets pédagogiques ? Comment proposer à n'importe quel apprenant la séquence personnalisée d'objets pédagogiques afin de réaliser la tâche de personnalisation ? (3). Présentons ci-dessous le scénario qui décrit la méthode générale. A partir de ce scénario, nous allons discuter les fonctions principales sur lesquelles se base notre méthode. L'apprenant, qui est connecté au système choisit un objet pédagogique comme un objectif d'apprentissage, le module de personnalisation va ensuite s'occuper de lui délivrer les séquences d'objets qui correspondent à son profil. Nous allons décrire dans ce qui suit les différents composants de ce modèle générale de la méthode.

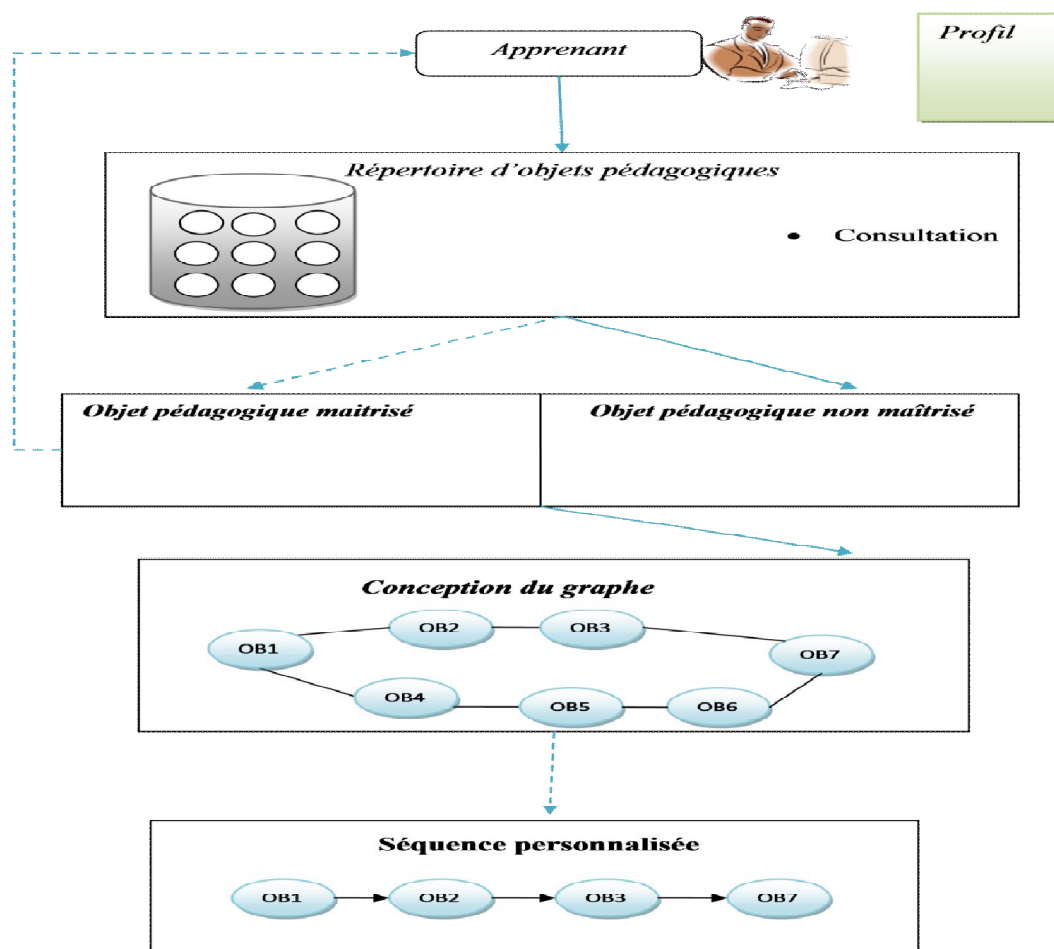


FIG.5.1– Modèle général de la méthode.

TAB.5.1– L'algorithme général de la méthode proposée.

Algorithme 4 : méthode générale de la personnalisation

Données : S: séquence personnalisée, g: graphe; Maîtrisé: Booléen, OP: Objet pédagogique ;

Résultat : Proposition d'une séquence personnalisée S;

Début

G ← Conception du graphe ;

Si (non maîtrisé OP) alors maîtrisé ← maîtrise (OP,g) ; Fin si ;

Sinon proposer (S) ;

Retourner (S) ;

Fin

5.2 Conception des objets pédagogiques

La question essentielle dans cette thèse est de développer une méthode de personnalisation des objets pédagogiques sélectionnés par un apprenant. Cette question, comment personnaliser l'apprentissage nous a poussés en premier lieu à poser la question suivante: comment modéliser l'ensemble des objets pédagogiques pour qu'on puisse chercher le chemin personnalisé.

Un objet pédagogique est définie, par l'IEEE (Institut des ingénieurs électriciens et électroniciens), comme « une entité, numérique ou non numérique, qui peut être utilisée pour l'apprentissage, l'éducation ou la formation ». L'idée principale derrière l'utilisation des objets pédagogiques est de diviser le contenu de l'enseignement en petits morceaux qui peuvent être réutilisés dans divers environnements d'apprentissage [Acampora et al. 2009]. Afin de permettre la réutilisation des objets pédagogiques, ils doivent être annotés avec des métadonnées, par exemple : (auteur, mots-clés, description, etc) [Debbah et al. 2014].

Pour cette fin, de nombreux schéma de métadonnées sont utilisés pour décrire ces objets pédagogiques. Nous citons : Sharable Content Operable le modèle objet de référence de contenu partageable(SCORM), Learning Object Metadata (LOM) développé par l'IEEE.

Dans notre approche, un objet pédagogique est un contenu d'apprentissage qui peut être délivré par un navigateur Web, qui est annoté avec une instance d'un schéma de métadonnées (interopérable avec la norme IEEE LOM) et qui est stocké et indexé dans un répertoire d'objets pédagogiques. Les objets pédagogiques contiennent des métadonnées qui les expliquent. Ces objets pédagogiques, qui constituent le contenu à apprendre (cours, leçons), sont assemblés à partir de grains pédagogiques puis associés pour constituer des parcours individuels de formation.

Ces objets pédagogiques sont aussi annotés avec un ensemble de propriétés qui les décrivent et un vocabulaire correspondant. La table 5.2 illustre les différentes propriétés et leur vocabulaire.

TAB.5.2– Valeurs des propriétés possible pour un objet pédagogique [Acampora et al. 2009].

Propriété	Vocabulaire
Type de ressource d'apprentissage	Exercice, Simulation, diagramme, figure, texte narrative, graphe, Index
Type d'interactivité	Active, expositif , mixte , indéfini
Niveau d'interactivité	Trop bas, Bas, moyen, élevé, très élevé
Difficulté	(ISO8061 format)
Temps d'apprentissage typique	Très facile, facile, moyen, difficile, très difficile
Langue	(anglais, fr, etc)
Contexte	Education primaire, éducation secondaire, premier cycle universitaire, second cycle universitaire.

5.3 Conception du profil de l'apprenant

« Personnaliser c'est adapter quelque chose à quelqu'un ». Afin de pouvoir adapter les objets pédagogiques aux besoins de chaque apprenant, il va falloir modéliser ces besoins dans une seule composante qui s'appelle : le profil de l'apprenant.

Un profil apprenant est Pour Self [Self 1994] destiné à fournir des informations sur l'objet modélisé, dans ce cas, l'apprenant qui utilise un système d'apprentissage. Quant à Höök [Höök et al. 1996], il définit le profil de l'apprenant comme une *connaissance* à propos de l'apprenant *explicitement* ou *implicitement* codée, utilisée par le système afin d'améliorer son interaction. Pour Koch [Koch 2000], le profil de l'apprenant est une *représentation abstraite* de quelque chose de réel qui est l'apprenant. C'est la croyance que le système a sur

l'apprenant représenté. Enfin, le profil de l'apprenant représente les propriétés de l'apprenant et la façon dont nous les représentons.

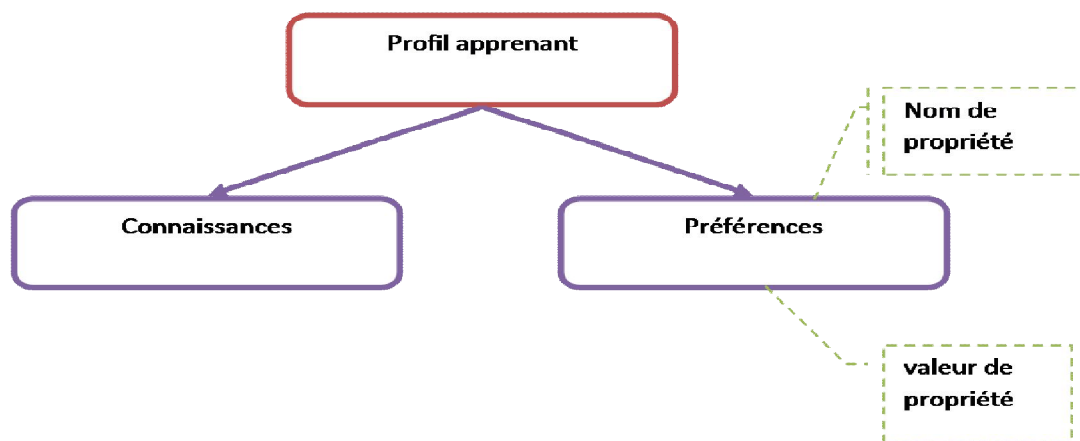


FIG.5.2.Composants du profil apprenant.

Tout profil apprenant est caractérisé par

- Un ensemble de connaissances
- Un ensemble de préférences

Une connaissance est caractérisée par :

- Un taux de maîtrise.

La valeur attribuée à ce taux de maîtrise est entre 0 et 1, où 0 indique l'absence totale de la connaissance sur cet objet, et 1 indique la connaissance totale de cet objet par rapport à cet apprenant. En effet le taux de maîtrise est calculé à travers des quizz ou bien des questions à choix multiples durant une phase d'évaluation.

Concernant la deuxième partie qui constitue le profil qui est les préférences est caractérisée par :

- Nom de la propriété
- Valeur de la propriété

Nous rappelons que ces propriétés existantes dans les préférences du profil de l'apprenant sont les mêmes qui caractérisent un objet pédagogique et cela afin de faire correspondre chaque objet avec l'état de préférence d'un apprenant donné.

La table 5.3 suivante résume les différentes propriétés qu'on a utilisées afin de définir les préférences de l'apprenant ainsi que l'ensemble de vocabulaire plausible afin de décrire ces propriétés.

TAB.5.3– Valeurs des propriétés des préférences.

Propriété	Vocabulaire
Type de ressource d'apprentissage	Exercice, Simulation, diagramme, figure, texte narrative, graphe, Index
Type d'interactivité	Active, expositif , mixte , indéfini
Niveau d'interactivité	Trop bas, Bas, moyen, élevé, très élevé
Difficulté	(ISO8061 format)
Temps d'apprentissage typique	Très facile, facile, moyen, difficile, très difficile
Langue	(anglais, fr, etc)
Contexte	Education primaire, éducation secondaire, premier cycle universitaire, second cycle universitaire.

5.4 Stratégies de conception d'un graphe

L'ensemble des objets pédagogiques est modélisé sous forme de graphe. Un graphe est un ensemble de nœuds reliés entre eux par des arcs. Les nœuds représentent les objets pédagogiques et les arcs représentent les relations entre ces objets. En effet, la recherche d'une séquence d'objets à partir d'un point de départ vers un point d'arrivée dans un graphe a une complexité polynomiale. Donc le problème de recherche tombe dans la catégorie des problèmes NP complet.

En plus, ce problème NP complet est un problème d'optimisation combinatoire. On qualifie généralement de « combinatoires » les problèmes dont la résolution se heurte à une explosion du nombre de combinaisons à explorer. Cette notion de problème combinatoire est formellement caractérisée par la théorie de la complexité qui propose une classification des problèmes en fonction de la complexité de leur résolution. Les Problèmes d'optimisation combinatoire ont pour but de maximiser (respectivement minimiser) une fonction objectif donnée. A chaque problème d'optimisation on peut associer un problème de décision dont le but est de déterminer s'il existe une solution pour laquelle la fonction objectif soit supérieure (respectivement, inférieure) ou égale à une valeur donnée. La complexité d'un problème d'optimisation est liée à celle du problème de décision qui lui est associé.

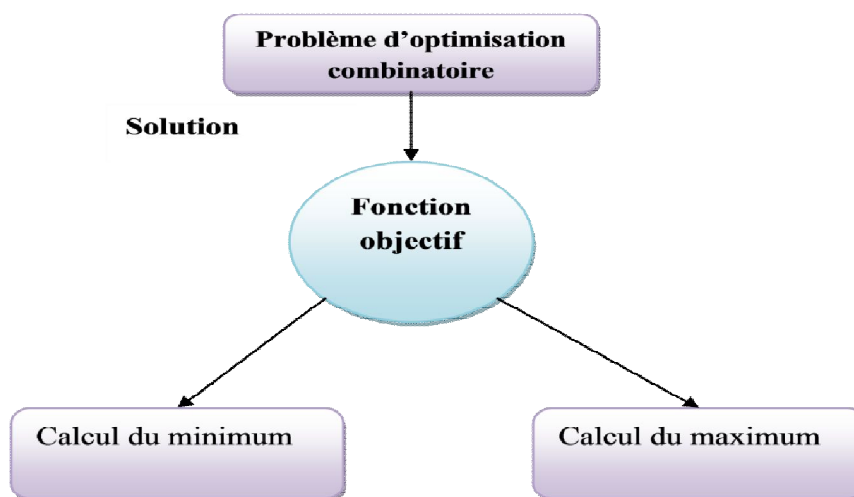


FIG.5.3– Solutions d'un problème d'optimisation combinatoire.

– Relations entre objets pédagogique

Les objets pédagogiques sont reliés entre eux par deux types de relations à savoir une relation de *pré requis* et une relation de *sortie*. L'utilisation de ces deux relations va pouvoir :

- organiser le graphe des objets pédagogiques
- mettre en priorités les objets pédagogiques

– Définition 1

La relation de pré requis peut être définie de différentes manières : elle définit la nécessité d'acquisition d'un objet pédagogique (i) avant d'accéder à l'objet pédagogique (j). Par exemple si un objet pédagogique (i) est un pré-requis de l'objet pédagogique(j), l'acquisition de l'objet pédagogique (j) implique l'acquisition de l'objet pédagogique (i). Elle indique aussi par exemple que l'apprentissage d'un objet pédagogique A est assujéti à la maîtrise de l'objet pédagogique B qui le précède.

– Définition 2

La relation de sortie définit l'acquisition d'un objet pédagogique (j) engendre comme conséquence l'acquisition de l'objet pédagogique (i). La figure 5.4 montre le schéma d'un objet pédagogique OP1 qui a comme pré requis un objet pédagogique OP2 et comme sortie un objet pédagogique OP3.

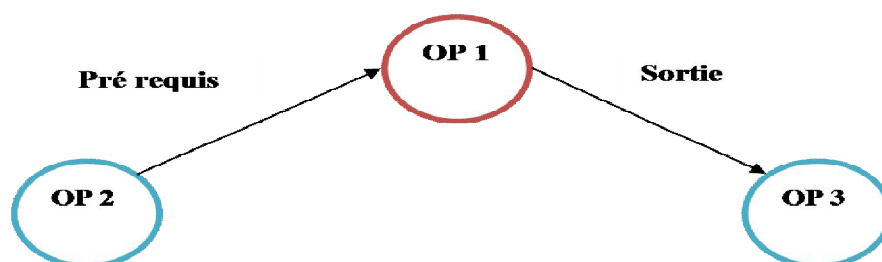


FIG.5.4. Les relations entre objets pédagogiques.

Cette façon de concevoir le graphe permet aux enseignants experts de mieux organiser leur graphe des objets pédagogiques.

5.5 Calcul du poids proposé

Vu les attentes particulières que nous avons par rapport à la personnalisation des objets pédagogiques, notamment la prise en compte du taux de maîtrise de chaque objet pédagogique ainsi que la prise en compte des préférences d'apprentissage pour chaque apprenant (les propriétés décrites sur la table et leurs valeurs) nous avons opté pour le calcul du poids pour chaque relation dans le graphe. Le calcul du poids se fait dans le but de pondérer le graphe, et ainsi de donner une valeur pour chaque relation dans le graphe.

Nous avons opté pour une formule qui permet de le faire. Cette formule se base tout d'abord sur le calcul du taux de maîtrise entre chaque deux objets pédagogiques. De plus, on a proposé une autre formule pour le calcul des préférences entre deux objets pédagogiques.

$$TM_{(OP_i, OP_j)} = \frac{Taux_{maîtrise} OP1 + Taux_{maîtrise} OP2}{2} \quad (5.1)$$

Où: TM : représente le taux de maîtrise de chaque objet pédagogique. Donc le taux de maîtrise entre deux objets OP_i et OP_j est égal à la somme des deux taux de maîtrise divisée par deux.

$$PA(OP_i, OP_j) = \frac{Préférence_{d'apprentissage} OPi + préférence_{d'apprentissage} OPj}{2} \quad (5.2)$$

Où PA : représente la préférence d'apprentissage représentée sur la table (5.3). Ce qui fait que la préférence d'apprentissage entre deux objets OP_i et OP_j est égal à la somme des deux préférences d'apprentissage divisée sur deux.

Le poids des arcs entre les objets pédagogiques est en effet calculé comme suit :

$$P(OP_i, OP_j) = \frac{TM(OP_i, OP_j) + PA(OP_i, OP_j)}{2} \quad (5.3)$$

Où : P : représente le poids entre deux objets OP_i OP_j , qui représente la somme des taux de maîtrise et les préférences d'apprentissage divisée sur deux.

La figure (5.5) montre un exemple d'un graphe composé d'un ensemble d'objets pédagogiques avec des poids calculés à l'aide de la formule.

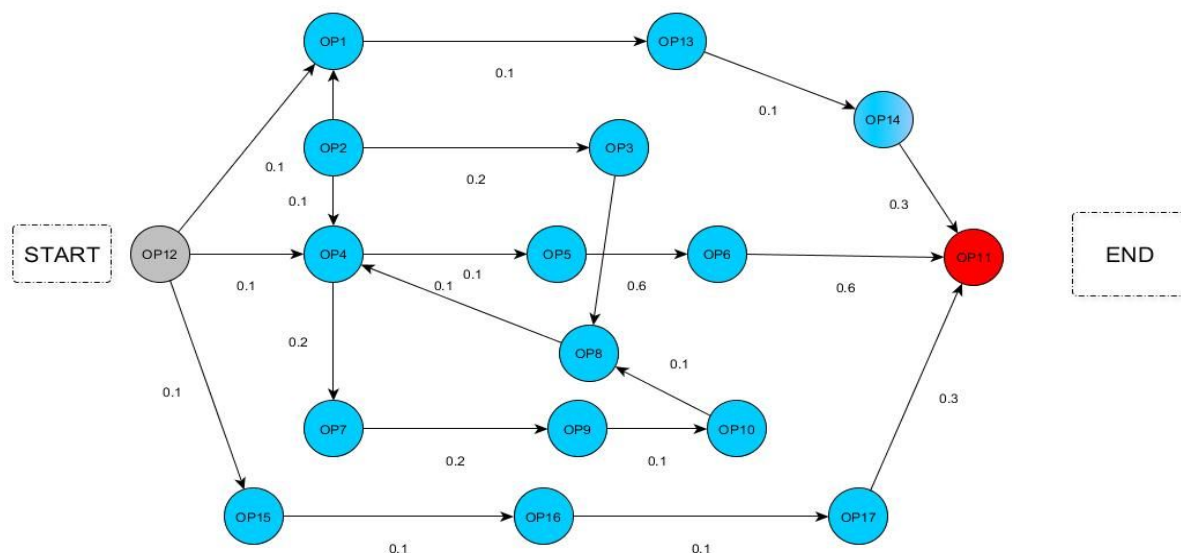


FIG.5.5– Exemple d'un graphe basé poids.

5.6 Conclusion

Dans ce chapitre, nous avons proposé le modèle de la méthode de personnalisation de l'apprentissage avec tous ses éléments de base : la conception du profil apprenant, les étapes de construction du graphe et la composition de la séquence personnalisée comme résultat.

En résumé, Les problèmes d'optimisation combinatoires, ont été résolus à l'aide de plusieurs méthodes bio-inspirées, on note parmi ces méthodes la méthode du calcul de l'ADN, qui a été abordé dans un chapitre précédant. Nous allons utiliser cette méthode pour faire de la personnalisation au sein de notre système. Nous allons détailler cette approche dans le chapitre suivant.

Chapitre 6

Personnalisation de l'apprentissage et calcul ADN

6.1 Introduction

Nous avons défini les grandes lignes de notre approche de personnalisation des objets pédagogiques, à savoir le profil de l'apprenant, la conception du graphe, et la recherche de la séquence personnalisée. Nous allons maintenant nous baser sur la tâche de personnalisation afin de chercher la séquence personnalisée où une solution est proposée en utilisant la méthode du calcul de l'ADN.

Nous allons tout d'abord commencer le chapitre par la conception d'une séquence personnalisée, suivie par la description des différentes étapes appliquées et puis nous terminons par exposer les résultats expérimentaux obtenus par notre algorithme.

6.2 Conception d'une séquence personnalisée

Dans le contexte des environnements informatiques d'apprentissage humains, la notion de séquence signifie une suite d'objets pédagogiques reliés entre eux par des relations. Le système va extraire les informations à partir du profil de l'apprenant qui choisit ses préférences par rapport aux objets disponibles. Dans notre contexte, on définit une séquence personnalisée de la manière suivante :

Définition

"Une séquence personnalisée est une séquence adaptée aux besoins d'un apprenant donné".

Le problème consiste donc à chercher la séquence personnalisée (optimale) par rapport à un objectif d'apprentissage dans un graphe. Ce problème est un problème d'optimisation combinatoire qui nécessite dans notre approche l'application de la méthode du calcul de l'ADN afin de le résoudre.

6.2.1 La recherche d'une séquence dans un graphe

La séquence d'apprentissage comprend un ensemble spécifique d'objets pédagogiques à partir d'un point de départ vers un point d'arrivée dans un graphe. En effet, le but est de chercher la séquence qui conduit l'apprenant vers son objectif d'apprentissage, et cela entre un objet pédagogique de départ et un objet pédagogique d'arrivée. Pour résoudre ce problème, nous allons proposer de calculer une fonction objectif.

6.2.2 Le calcul de la fonction objectif

La fonction objectif sert à évaluer des individus séquences en brins d'ADN représentant un chemin dans un graphe. le problème d'optimisation consiste à minimiser ou à maximiser cette fonction jusqu'à l'obtention de l'optimum. Dans notre cas, le système va chercher la séquence optimale en minimisant la somme des poids totaux (P) (voir équation 6.1) calculés pour un chemin dans le graphe afin de proposer la séquence optimale qui répond au mieux à l'état des préférences et des connaissances de chaque apprenant.

$$\text{Min } f(W) = \sum_{i,j} P(OP_i, OP_j) \quad (6.1)$$

Où $P(OP_i, OP_j)$ représente le poids calculé entre l'objet pédagogique OP_i et l'objet pédagogique OP_j .

6.3 Scénario de la personnalisation de l'apprentissage

Après la sélection de certains objets pédagogiques basés sur les connaissances et les préférences, nous allons donc passer maintenant à la description du scénario de la personnalisation de l'apprentissage proposée.

Le scénario passe par trois étapes :

- Une étape de sélection des objets reliés au but pédagogique.
- Une étape de pré-évaluation des connaissances et calcul du poids dans le graphe.
- Une étape d'optimisation de la séquence d'apprentissage basée ADN.

– **Etape 1 : sélection du but d'apprentissage**

A cette étape, on propose à l'apprenant un ensemble d'objets pédagogiques, l'apprenant doit choisir parmi l'ensemble de ces objets, l'objet but pour déclencher le processus de recherche de la séquence personnalisée dans le graphe.

– **Etape 2 : pré-évaluation des connaissances et calcul du poids dans le graphe**

Une fois l'objectif déterminé, une pré-évaluation des connaissances de l'apprenant est effectuée, et l'ensemble d'objets pédagogiques reliés entre eux par l'objectif va permettre de construire le graphe.

– **Etape 3 : optimisation de la séquence d'apprentissage basée ADN**

Après avoir effectué le calcul du poids dans le graphe, le processus d'optimisation de la séquence d'apprentissage basée ADN est déclenché, afin de délivrer une séquence personnalisée à la fin.

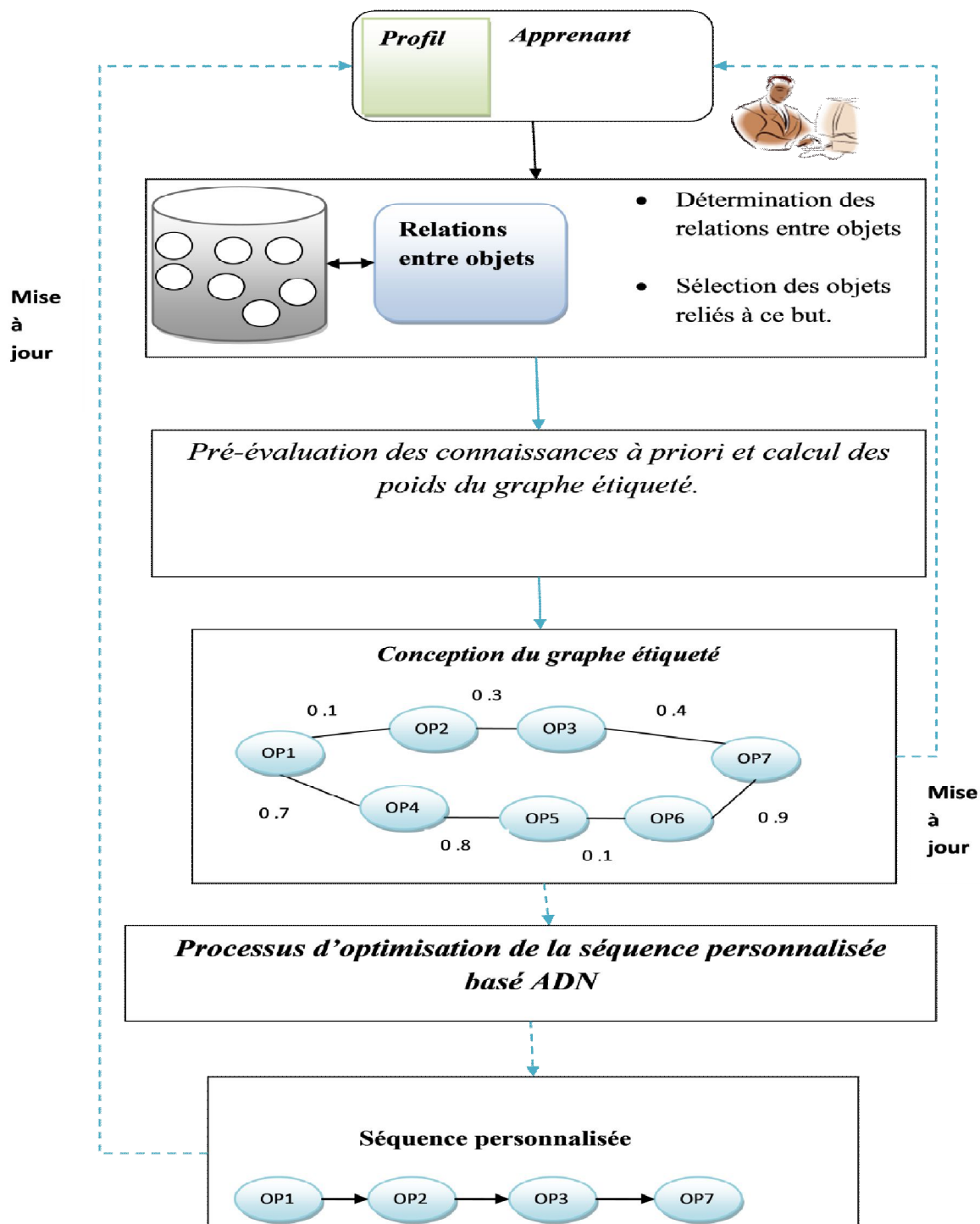


FIG.6.1– Scénario de la personnalisation.

6.3.1 Construction du profil de l'apprenant

Le profil de l'apprenant est caractérisé par l'ensemble des connaissances et l'état des préférences de l'apprenant. L'état des préférences dans le profil de l'apprenant est le même utilisé pour décrire les propriétés des objets pédagogiques.

Connaissances A priori (Taux de maîtrise des objets pédagogiques)		Etat des préférences d'apprentissage
	OP9(0.1)	Langue: An
OP1(0.1)	OP10(0.1)	Niveau d'interactivité: élevé
OP2(0.1)	OP11(0.1)	Type d'interactivité: Active
OP3(0.3)	OP12(0.1)	Type de ressource
OP4(0.1)	OP13(0.1)	d'apprentissage: texte narratif
OP5 (0.1)	OP14(0.1)	Difficulté: Moyenne
OP6 (0.3)	OP15(0.1)	Contexte: éducation supérieur
OP7 (0.3)	OP16(0.1)	Temps d'apprentissage typique :
OP8(0.1)	OP17(0.1)	facile

FIG.6.2– Représentation du profil apprenant après détermination du but.

6.3.2 Caractéristiques d'un objet pédagogique

Chaque objet pédagogique est caractérisé par un ensemble de métadonnées descriptives (voir section 5.2). Cet ensemble de métadonnées est enregistré et répertorié dans le répertoire des objets pédagogiques. La figure 6.3 montre la représentation des métadonnées d'objets pédagogiques.

6.4 Application de la méthode du calcul de l'ADN

La méthode du calcul de l'ADN à été appliquée sur l'exemple du graphe proposé sur la figure 5.6 de la section précédente. Commençons d'abord par décrire les étapes de l'algorithme du calcul de l'ADN.

OP1 (An, élevé, Actif, Simulation, Moyenne, éducation supérieur, facile)
OP2 (An, élevé, Actif, texte, Moyenne, éducation supérieur, très facile)
OP3 (An, bas, Actif, texte narratif, Moyenne, éducation supérieur, facile)
OP4 (An, élevé, Actif, texte narratif, élevé, éducation supérieur, facile)
OP5 (An, élevé, Actif, texte narratif, Moyenne, éducation supérieur, facile)
OP6 (An, élevé, Actif, texte narratif, Moyenne, éducation supérieur, facile)
OP7 (An, élevé, Expositif, texte narratif, Moyenne, éducation supérieur, facile)
OP8 (An, élevé, Actif, Simulation, bas, éducation supérieur, facile)
OP9 (An, élevé, Expositif, texte narratif, Moyenne, éducation supérieur, facile)
OP10 (An, élevé, Actif, texte narratif, élevé, éducation élevé, facile)
OP11 (An, élevé, Actif, texte narratif, Moyenne, éducation supérieur, facile)
OP12 (An, élevé, Actif, Simulation, bas, éducation supérieur, facile)
OP13 (An, élevé, Expositif, texte narratif, Moyenne, éducation supérieur, très facile)
OP14 (An, élevé, Actif, Simulation, Moyenne, éducation supérieur, facile)
OP15 (An, élevé, Expositif, texte narratif, Moyenne, éducation supérieur, facile)
OP16 (An, élevé, Actif, texte narratif, élevé, éducation supérieur, Medium)
OP17 (An, élevé, Actif, texte narratif, Moyenne, éducation supérieur, Difficile)

FIG.6.3– La représentation des métadonnées d’objets pédagogiques.

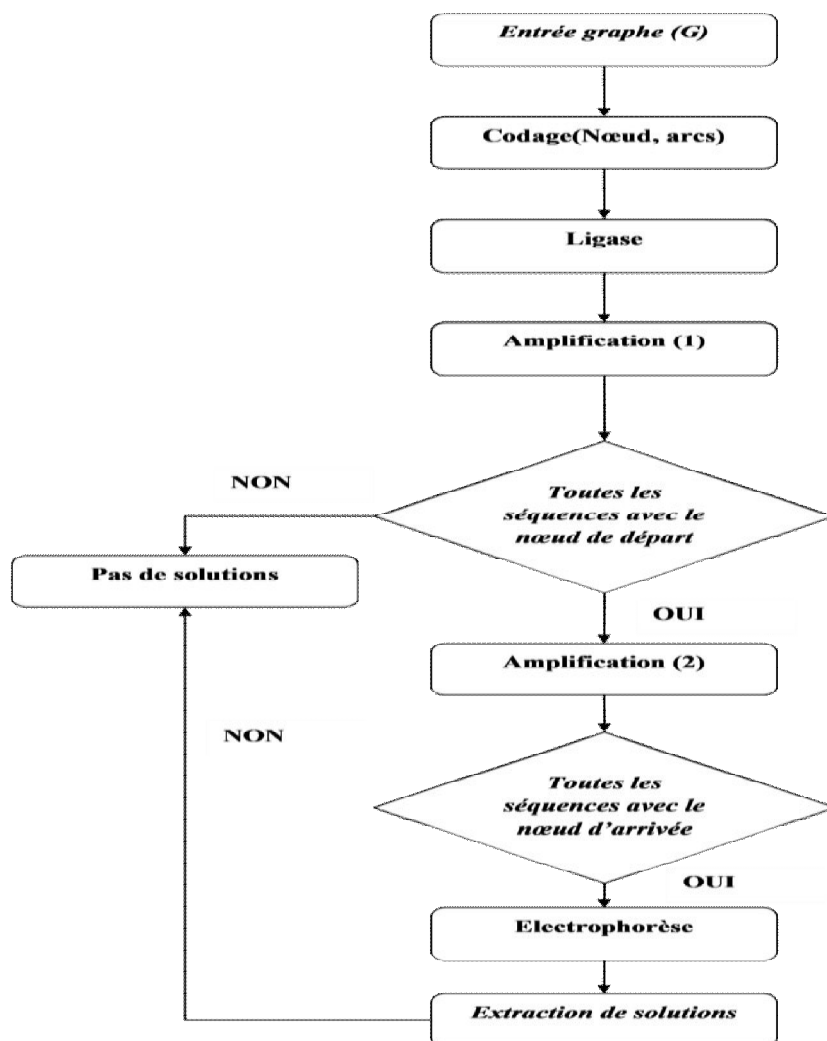


FIG.6.4– Organigramme représentant les étapes de l’algorithme du calcul de l’ADN.

6.4.1. Étape 1 : codage

Le codage permet de coder les nœuds ainsi que les arcs du graphe en leurs attribuant des codes contenant les quatre bases « A » « C » « G » « T ». La génération des codes de cette étape se fait d’une manière aléatoire. Soit b l’objet pédagogique de départ de la séquence dans le graphe, et P le poids et e l’objet pédagogique de l’arrivée de la séquence. [Debbah et al. 2014]. La création des brins représentant les séquences se base sur les caractéristiques suivantes :

- (a) $E_i, i = \{1, \dots, I\}$: avec I séquences nucléotides générées aléatoirement, chaque séquence est composée d'un ensemble d'objets décrits dans le graphe.
- (b) $\bar{E}_i, i = \{1, \dots, I\}$: avec I séquences nucléotides complémentaires de E_i décrits dans le graphe.
- (c) $E_p, p = \{0, 1, \dots, P\}$: avec P séquences nucléotides aléatoires générées pour tous les poids P dans le graphe.
- (d) $\bar{E}_p, p = \{0, 1, \dots, P\}$: avec P séquences complémentaires de E_p générées pour tous les poids P dans le graphe.

Afin de construire les brins d'ADN, chaque nœud OP_i de E_i du graphe est codé aléatoirement avec les 4 bases {A, C, G, T}. Ensuite, pour chaque poids P_i des arcs entre les nœuds nous générons au hasard 4 bases d'ADN qui augmente avec 2^2 de bases d'ADN pour chaque poids P_{i+1} . Les tables 6.1 et 6.2 montrent respectivement les codes des brins d'ADN pour les nœuds et les poids du graphe pondéré [Debbah et al. 2014].

TAB.6.1– Codes ADN des objets pédagogiques.

Objets pédagogiques	Codes ADN	Complément ADN
OP1	TTGA	AACT
OP2	CTTC	GAAG
OP3	TATA	ATAT
OP4	CCCG	GGGC
OP5	GCAT	CGTA
OP6	TAGC	ATCG
OP7	CATC	GTAG
OP8	ATTC	TAAG
OP9	AAGG	TTCC
OP10	CCGT	GGCA
OP11	TTCC	AAGG
OP12	CAAG	GTTC
OP13	CGAT	GCTA
OP14	AATA	TTAT
OP15	CGCT	GCGA
OP16	AAAA	TTTT

OP17	CACT	GTGA
------	------	------

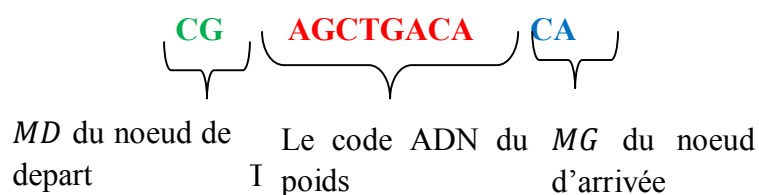
TAB.6.2– Codes ADN des poids.

Poids	Code ADN
P1= 0.1	ATAG
P2= 0.2	AGCTGACA
P3= 0.3	AACCGGAGGTCC
P5= 0.6	GCGGTCCAACCGCAA

6.4.2 Étape 2 : construction de brins d'ADN d'arcs

Pour former les brins d'ADN d'arcs entre deux objets pédagogiques (b) et (e), nous joignons à la moitié droite qu'on appelle *MD* (pour Moitié Droite) du premier objet pédagogique (b) le code ADN correspondant au poids entre les deux objets pédagogiques et la moitié gauche *MG* (pour Moitié Gauche) de l'objet pédagogique (e). L'arc est représenté par : $MD_b + E_p + MG_e$. Juste après on met les brins résultats dans un tube à essai T1.

Par exemple l'arc: OP4->OP7 est représenté comme suit:



TAB.6.3– Codes ADN des arcs.

Arcs	Brin ADN
OP15->OP16	CTATAGAA
OP2->OP4	TCATAGCC
OP10->OP8	GTATAGAT
OP4->OP7	CGAGCTGACACA
OP12->OP15	AGATAGCG
OP12->OP4	AGATAGCC
OP3->OP8	TAAGCTGACAAT
OP6->OP11	GCGCGGTCCAACCGCAAATT
OP5->OP6	ATGCGGTCCAACCGCAAATA
OP2->OP1	TCATAGTT
OP17->OP11	CTAACCGGAGGTCCTT
OP9->OP10	GGATAGCC
OP2->OP3	TCAGCTGACATA
OP8->OP4	TCATAGCC
OP1->OP13	GAATAGCG

OP16->OP17	AAATAGCA
OP7->OP9	TCAGCTGACAAA
OP13->OP14	ATATAGAA
OP14->OP11	TAAACCGGAGGTCCTT
OP4->OP5	CGATAGGC

6.4.3 Étape 3 : Ligature

Une opération de ligature est réalisée en utilisant des enzymes ligase dans un tube à essai T1, en fait, tous les brins d'ADN qui correspondent à la moitié droite de leurs nœuds qui rencontrent la moitié droite de leurs brins d'ADN vont se réunir dans un seul brin d'ADN. Cette opération est répétée jusqu'à l'obtention d'un chemin aléatoire dans le graphe. En conséquence, l'opération de ligature générera des brins d'ADN codant toutes les séquences aléatoires dans le graphe. Nous mettons tous les brins d'ADN de séquences dans un tube à essai T2. La table 6.4 montre l'ensemble des brins d'ADN des séquences aléatoires dans le graphe.

TAB.6.4 –Résultats de l'opération de ligase.

Séquence	Brin ADN
OP17->OP11	CTAACCGGAGGTCCTT
OP15->OP16->OP17->OP11	CTATAGAAAAATAGCACTAACCGGAGGTCCTT
OP3->OP8->OP4->OP7->OP9->OP10	TAAGCTGACAATTCATAGCCCGAGCTGACACATCAG CTGACAAAGGATAGCC

OP8->OP4->OP7->OP9->OP10	TCATAGCCCGAGCTGACACATCAGCTGACAAAGGATAGCC
OP6->OP11	GCGCGGTCCAACCGCAAATT
OP2->OP3->OP8->OP4->OP5->OP6->OP11	TCAGCTGACATATAAGCTGACAATTCATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAA CCGCAAATT
OP9->OP10->OP8->OP4->OP7	GGATAGCCGTATAGATTCATAGCCCGAGCTGACACA
OP13->OP14->OP11	ATATAGAATAAACCGGAGGTCCTT
OP14->OP11	TAAACCGGAGGTCCTT
OP2->OP3->OP8->OP4->OP7->OP9->OP10	TCAGCTGACATATAAGCTGACAATTCATAGCCCGAGCTGACACATCAGCTGACAAAGGATAGCC
OP2->OP1->OP13->OP14->OP11	TCATAGTTGAATAGCGATATAGAATAAACCGGAGGTCCTT
OP12->OP1->OP13->OP14->OP11	AGATAGTTGAATAGCGATATAGAATAAACCGGAGGTCCTT
OP7->OP9->OP10->OP8->OP4->OP5->OP6->OP11	TCAGCTGACAAAGGATAGCCGTATAGATTCATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGT CCAACCGCAAATT
OP10->OP8->OP4->OP7->OP9	GTATAGATTCATAGCCCGAGCTGACACATCAGCTGACAAA
OP16->OP17->OP11	AAATAGCACTAACCGGAGGTCCTT

OP2->OP4->OP7->OP9->OP10->OP8	TCATAGCCCGAGCTGACACATCAGCTGACAAAGGATAGCCGTATAGAT
OP8->OP4->OP5->OP6->OP11	TCATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP1->OP13->OP14->OP11	GAATAGCGATATAGAATAAACCGGAGGTCCTT
OP12->OP4->OP5->OP6->OP11	AGATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP4->OP5->OP6->OP11	CGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP2->OP4->OP5->OP6->OP11	TCATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP10->OP8->OP4->OP5->OP6->OP11	GTATAGATTCATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP12->OP15->OP16->OP17->OP11	AGATAGCGCTATAGAAAAATAGCACTAACCGGAGGTCCTT
OP4->OP7->OP9->OP10->OP8	CGAGCTGACACATCAGCTGACAAAGGATAGCCGTATAGAT
OP5->OP6->OP11	ATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP9->OP10->OP8->OP4->OP5->OP6->OP11	GGATAGCCGTATAGATTCATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP3->OP8->OP4->OP5->OP6-	TAAGCTGACAATTCATAGCCCGATAGGCATGCGGTC

>OP11	CAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP12->OP4->OP7->OP9->OP10->OP8	AGATAGCCCGAGCTGACACATCAGCTGACAAAGGATAGCCGTATAGAT

6.4.4 Étape 4 : Amplification (T2, T3)

Tous les brins d'ADN existants en éprouvette T2 et qui commencent par l'objet pédagogique de départ sont amplifiés par la Réaction en Chaîne par Polymérase (Polymerase Chain Reaction en anglais) (PCR) en utilisant des séquences complémentaires comme amorces et enzymes (des polymérases de l'ADN). Dans ce cas, l'amorce est complémentaire de la séquence de l'objet pédagogique de départ. Après cela, ils sont roulés dans un tube à essai T3. Ainsi T3 contient tous les brins d'ADN de pistes avec le nœud de départ similaire [Debbah et al. 2014].

TAB.6.5– Résultats du Réaction en chaîne par Polymérase (PCR) (1).

Séquence	Brin ADN
OP12->OP1->OP13->OP14->OP11	AGATAGTTGAATAGCGATATAGAATAAACCGGAGGTC CTT
OP12->OP4->OP5->OP6->OP11	AGATAGCCCGATAGGCATGCGGTCCAACCGCAAATAG CGCGGTCCAACCGCAAATT
OP12->OP15->OP16->OP17->OP11	AGATAGCGCTATAGAAAAATAGCACTAACCGGAGGTC CTT
OP12->OP4->OP7->OP9->OP8->OP10	AGATAGCCCGAGCTGACACATCAGCTGACAAAGGATA GCCGTATAGAT

6.4.5 Étape 5 : Amplification (T3, T4)

Tous les brins d'ADN avec l'objet pédagogique de départ obtenus et restaurés dans T3 de l'étape précédente, sont amplifiés encore par la Réaction en chaîne par Polymérase (PCR) en utilisant l'objet pédagogique de fin d'amorce. Nous amplifions ceux qui se terminent avec le bon objet pédagogique et nous mettons les résultats dans le tube à essai T4. Lorsque l'amplification se termine, le tube à essais T4 ne contient que les brins ayant l'objet pédagogique de début et de fin. Les résultats de cette étape sont présentés dans la table 6.6.

TAB.6.6– Résultats du Polymerase Chain Reaction (PCR)(2).

Séquence	Brin ADN
OP12->OP1->OP13->OP14-> OP11	AGATAGTTGAATAGCGATATAGAATAAACCGGAGGTCCTT
OP12->OP4->OP5->OP6-> OP11	AGATAGCCCGATAGGCATGCGGTCCAACCGCAAATAGCGCGGTCCAACCGCAAATT
OP12 ->OP15->OP16->OP17->OP11	AGATAGCGCTATAGAAAATAGCACTAACCGGAGGTCCTT

6.4.6 Étape 6 : la séparation (T4, T5)

Nous séparons les brins d'ADN avec la bonne longueur des brins d'ADN à partir de tube à essai T4 et on les dispose dans le tube à essai T5 par l'électrophorèse en gel d'agarose. L'électrophorèse en gel d'agarose permet la séparation et la récupération des brins d'ADN de la bonne longueur. Le résultat de la séparation et la récupération des brins d'ADN est le plus petit brin d'ADN qui contient la solution au problème.

TAB.6.7– Résultats de la séparation.

Séquence	Brin ADN
OP12->OP15->OP16->OP17->OP11	AGATAGCGCTATAGAAAAATAGCACTAACCGGAGGT CCTT

6.5 PERSO : système pour la personnalisation de l'apprentissage PERSO

6.5.1 Démarche pédagogique de PERSO

PERSO repose sur une démarche pédagogique destinée à la personnalisation des objets pédagogiques dans un EIAH. L'ensemble des objets pédagogiques est présenté sous forme de graphe. L'objectif étant de proposer la séquence personnalisée pour un apprenant donné.

L'approche déjà proposée dans le chapitre précédent part de l'idée que pour accéder à un objet pédagogique, il faut comparer les valeurs de ses métadonnées avec les valeurs des préférences de l'apprenant afin de calculer le poids des relations dans le graphe.

6.5.2 Architecture de PERSO

6.5.2.1 Architecture logicielle

Cet outil est basé sur une architecture trois tiers. Cette architecture appelée aussi architecture client/serveur deuxième génération, ou client/serveur distribué, est caractérisée principalement par l'invariance à l'échelle (elle peut supporter un nombre important de clients). Une telle architecture applique les principes suivants :

- La représentation est toujours prise en charge par le poste client.
- Le traitement est pris en charge par un serveur intermédiaire.

La figure 6.5 ci-dessous montre l'architecture client/serveur du système.

Pour l'architecture le premier tiers est l'interface qui représente le point d'accès des apprenants à travers les applets. Le deuxième tiers s'occupe de l'exécution des applets et de la représentation des pages HTML. Le dernier tiers (les données) doté d'une application serveur qui traite toutes les requêtes transmises par les clients à travers RMI (Remote method Invocation).

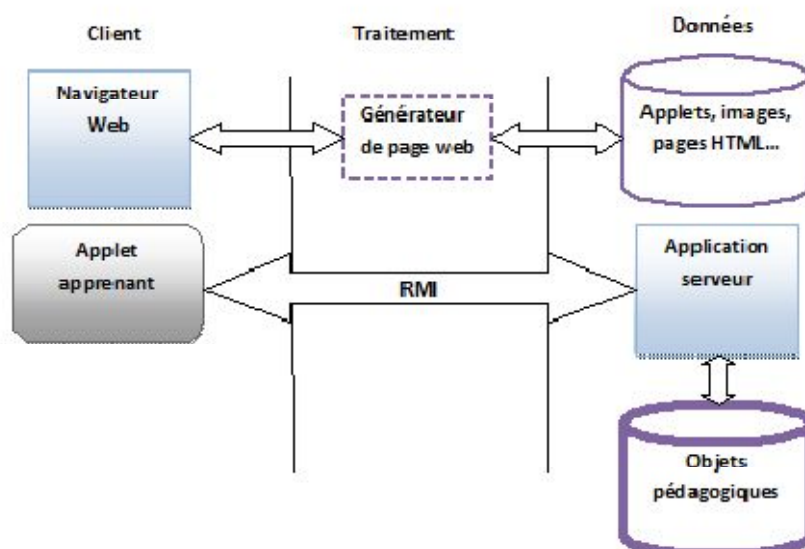


FIG.6.5– Architecture logiciel de PERSO.

6.5.2.2 Architecture fonctionnelle

PERSO est un outil de personnalisation de l'apprentissage des objets pédagogiques permettant à l'apprenant de choisir un but d'apprentissage et de délivrer le meilleur cours adaptatif au profil de l'apprenant sous forme d'une séquence personnalisée.

Il existe deux sessions. Une session d'enregistrement de profil de l'apprenant, d'une base d'objets pédagogiques. Et une deuxième session de personnalisation qui permet d'atteindre notre objectif.

Pour cela, l'outil est constitué principalement de deux composants : interface utilisateur (apprenant), et le mécanisme d'optimisation de cours qui est dans notre cas l'algorithmique.

Le système est doté de deux principales interfaces : la première est une page HTML ou figure l'ensemble des objets pédagogiques (choix éventuel des cours à apprendre) relatifs au domaine de l'algorithmique, et la deuxième est gérée par une applet java qui constitue l'interface apprenant qui contient l'ensemble des objets pédagogiques dans la séquence proposée.

Les objets pédagogiques sont générés à l'aide du logiciel [Articulate StoryLine 2015] qui permet de construire des objets pédagogiques en utilisant le standard LOM (Learning Object Metadata). En effet les objets pédagogiques résultant sont représentés graphiquement sous une forme de pages HTML où l'apprenant peut accéder à leurs contenus en cliquant dessus.

Que se passe-t-il dans une session d'enregistrement ?

Durant la session d'enregistrement, un formulaire sous forme de QCM (Questions à Choix Multiples) sur les préférences est proposé à l'apprenant sur lequel il va répondre (voir annexe). Et une pré-évaluation des connaissances sur le domaine à étudier est effectuée aussi à l'aide d'un QCM afin de construire un profil initial.

Que se passe-t-il dans une session de personnalisation ?

Après la session d'enregistrement, la session de personnalisation consiste à proposer Et une pré-évaluation des connaissances sur le domaine à étudier est effectuée aussi à l'aide d'un QCM (Questions à Choix Multiples) afin de construire un profil initial.

En second lieu, l'apprenant devrait choisir un objet pédagogique parmi ceux proposés afin de l'apprendre durant sa session d'apprentissage.

Après, dès que l'apprenant effectue son choix, le système va s'occuper de lui délivrer la séquence personnalisée correspondante à son profil.

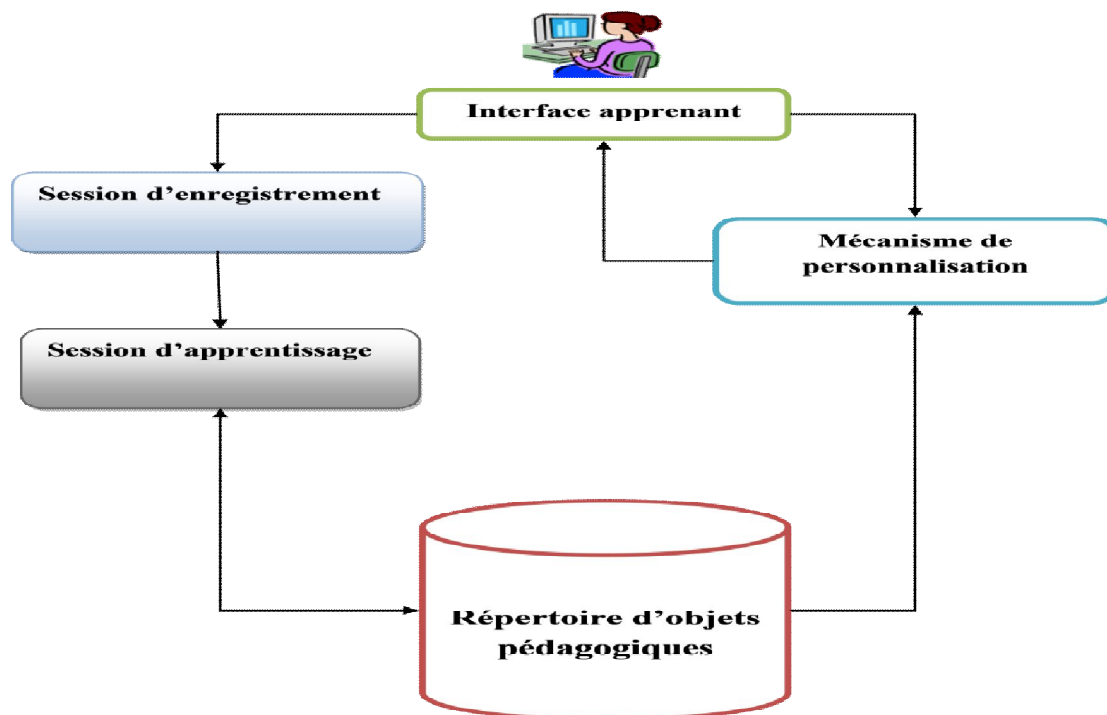


FIG.6.6–Architecture fonctionnelle de PERSO.

6.5.3 Expérimentation PERSO

Afin de valider la méthode proposée qui consiste à générer automatiquement un cours sous forme de séquence d'objets pédagogiques personnalisée, nous allons passer à une évaluation réelle qui se déroule avec une population de 25 étudiants de l'Ecole Préparatoire Aux Sciences et Techniques Annaba durant le premier semestre 2015.

Les étudiants ont été invités à utiliser le système PERSO pour accéder aux objets pédagogiques construits sur la base du module de l'algorithmique.

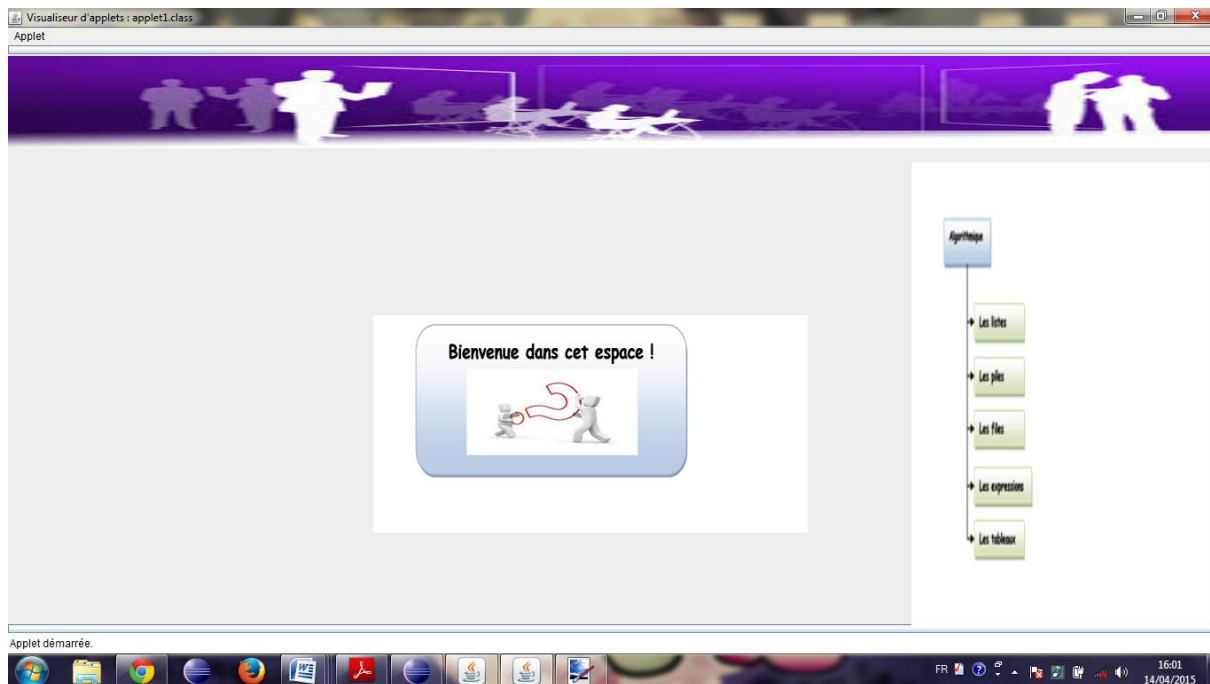


FIG 6.7 – Interface principale du système PERSO.

6.5.3.1 Résultats

Le cours qui a été choisi de l'algorithmique comporte 16 objets pédagogiques selon la figure (6.8) proposés aux étudiants, seulement 5 objets pédagogiques ont été maîtrisés par le système PERSO, 11 objets pédagogiques n'ont pas été maîtrisés par le système.

Le tableau suivant présente les différents taux de maîtrise des objets par le système PERSO.

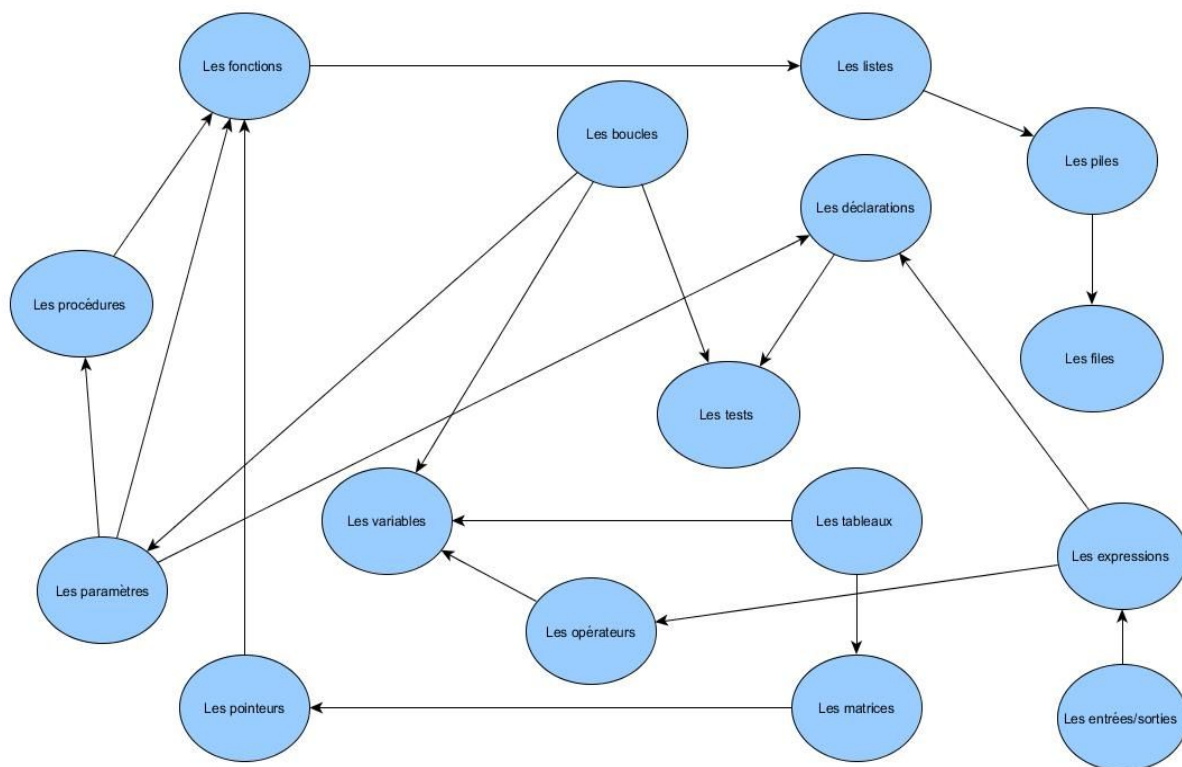


FIG.6.8– Graphe des objets pédagogique du module algorithmique.

TAB.6.8– Taux de maîtrise des objets présenté par PERSO.

Maitrisé	Non maîtrisé
3,125%	96,875%

TAB.6.9– Préférences des objets présentés par PERSO.

Préféré	Non préféré
10,75%	89,25%

La pré-évaluation des apprenants lors de la sélection du but d'apprentissage nous a permis d'observer que parmi les 16 objets présents dans le cours seulement 5 objets en moyenne ont été maîtrisés, et 11 objets en moyenne n'ont pas été maîtrisés. Un objet est maîtrisé si pour chaque apprenant sa note est égale à 10.

Le poids calculé change en fonction de deux paramètres qui sont le degré des préférences d'apprentissage et le taux de maîtrise. Afin de pouvoir observer l'effet des deux paramètres sur le poids des relations dans le graphe, nous allons varier des valeurs différentes. La figure (6.9).

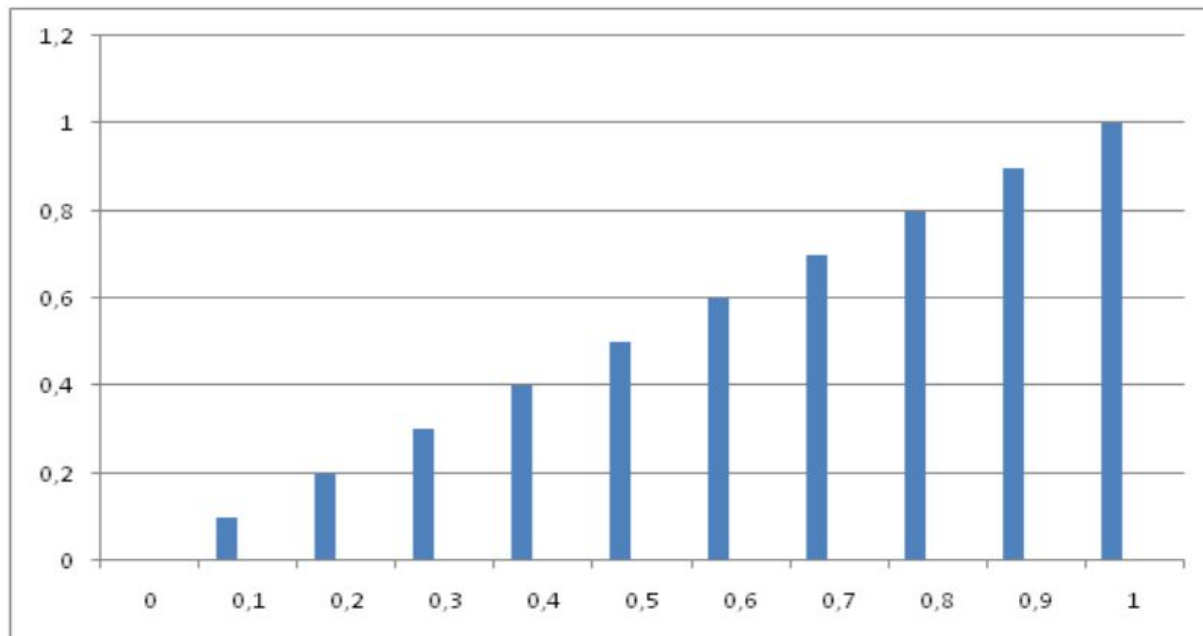


FIG.6.9–Variation des poids en fonction des paramètres PA, et taux de maîtrise TM.

La première question à laquelle nous cherchons à répondre en terme statistique concerne la corrélation entre les taux de maîtrise avant et après apprentissage via le système PERSO.

Le calcul du coefficient de corrélation a engendré la valeur $-0,01354157$ entre le taux de maîtrise avant et après utilisation du système PERSO. Cette valeur montre qu'il y'a une faible corrélation entre les deux taux de maîtrise.

Les résultats obtenues par le système avant et après utilisation du système ont été comparé aux ceux obtenus après. Les différents paramètres des distributions des 400 taux de maîtrise attribués par PERSO sont mentionnés dans le tableau 6.10 ci-après.

TAB.6.10– Paramètres de distributions des taux de maîtrise par PERSO.

Taux de maîtrise	Moyenne	Ecart type	Médiane	Min	Max	Etendu
Avant proposition	14%	0,32826072	10	0	10	10
Après proposition	53,25%	2,87054002	0	0	10	10

Un deuxième constat concerne les deux moyennes (moy_avant=2,9375 , moy_après=8,52). On peut clairement remarquer qu'il y a une différence considérable entre la moyenne des taux de maîtrise avant et après proposition des séquences personnalisées.

Les écarts types fournissent une mesure statistique de la dispersion moyenne des taux de maîtrise, qui présentent également des valeurs élevées.

6.5.3.2 Discussion des résultats

Nous discutons dans cette partie quelques questions par rapport aux résultats obtenus durant l'expérimentation. De ce fait, nous comptons souligner les différents avantages et les limites observés de cette approche, pour ouvrir la voie sur les perspectives que nous voyons pour améliorer le système.

La comparaison des deux groupes de taux de maîtrise obtenus a révélée une différence importantes entre la moyenne obtenue avant utilisation du système et après.

D'une autre part, les résultats de l'analyse de corrélation en utilisant le coefficient de pearson (-0,01354157) indique une faible association linéaire entre les deux groupes de taux de maîtrise.

En général, les résultats de cette méthode montrent l'efficacité de la méthode de personnalisation. Cette efficacité revient à l'automatisation du calcul du poids entre les objets pédagogiques dans le graphe en terme de :

- Paramètre PA qui représente les préférences d'apprentissage pour calculer le poids

- Le taux de maîtrise pour calculer le poids.
- Le poids résultat pour calculer la fonction objectif.

En ce qui concerne le mécanisme de personnalisation, l'expérimentation de cette approche a révélé que pour un seul objet pédagogique but, il se peut qu'il y ait plusieurs séquences personnalisées.

6.6 Conclusion

Parmi les objectifs de cet outil est l'amélioration des connaissances des apprenants en leur proposant des séquences d'objets pédagogiques personnalisés et ainsi augmenter leur motivations.

Avec cet outil, on peut personnaliser l'apprentissage des apprenants et ainsi assurer la proposition des séquences d'objets pédagogiques pertinentes même pour un grand groupe d'apprenants.

Notons aussi que l'enrichissement de la base des objets pédagogiques sera important pour assurer une personnalisation plus performante.

Conclusion générale et perspectives

1. Conclusion

Jusqu'à présent, la plupart des recherches dans les EIAH ont mis l'accent sur l'architecture, la présentation, la collaboration et la distribution. Avec le grand nombre hétérogène d'utilisateurs et des ressources d'une part, et l'avancement rapide de la technologie et la vitesse de communication d'autre part, l'attention s'est dirigée aujourd'hui beaucoup plus à un service personnalisé. En même temps, la richesse du matériel d'apprentissage disponibles et les différences des apprenants et de leurs besoins, rend la satisfaction des apprenants d'aujourd'hui difficile de se limiter à un seul fournisseur d'apprentissage. Plutôt, ils préféreraient avoir l'ensemble de matériaux d'apprentissage disponibles à leurs services. Pour répondre à cette exigence, les EIAH ont besoin d'outils intelligents flexibles et dynamiques plus avancés qui peuvent faciliter la découverte, l'importation, la composition et la livraison du contenu d'apprentissage de différentes ressources. Beaucoup de travaux de recherches ont été investis dans des systèmes personnalisés. Il ya eu une croissance d'intérêt (bien que toujours relativement petite) pour enquêter sur les méthodes du calcul évolutionnaire afin d'améliorer les services des systèmes d'apprentissage et leurs capacités.

1.1 Synthèse

Dans cette thèse nous avons d'abord présenté un état de l'art sur la personnalisation de l'apprentissage dans les EIAH, et les différentes méthodes du calcul évolutionnaires utilisées pour faire de la personnalisation de l'apprentissage dans un EIAH. Cependant, nous nous sommes intéressés à la méthode du calcul de l'ADN sur laquelle se reposent nos contributions. En se basant sur les limites des techniques proposées dans la littérature, nous avons proposé nos contributions comme solutions à ces limites.

L'approche proposée a été conçue dans le but de délivrer à n'importe quel apprenant avec un profil donné contenant ses connaissances et ses préférences, une séquence personnalisée indépendamment du domaine d'apprentissage. Cette prise en compte de la variabilité des apprenants et leurs différents profils permet non seulement d'offrir un service personnalisé mais aussi de guider l'apprenant dans son parcours d'apprentissage.

L'approche de personnalisation proposée assure une modélisation des connaissances et des préférences des apprenants réellement possédées, et aide les apprenants à progresser.

- ***Contribution #1 : Une approche de personnalisation des objets pédagogiques***

L'approche a été conçue dont le but de modéliser n'importe quel domaine d'apprentissage constitué d'un ensemble d'objets pédagogiques. Elle est basée essentiellement sur la stratégie de conception de graphe constituant les objets pédagogiques reliés entre eux. La modélisation, consiste à identifier les relations entre les objets pédagogiques, et calculer le poids des relations entre ces objets dans le graphe en utilisant le degré du taux de maîtrise de chaque objet dans le graphe, ainsi que les préférences des apprenants par rapport aux objets pédagogiques proposés. En effet, l'avantage du calcul du poids dans le graphe est de favoriser un chemin dans le graphe par rapport à un autre afin de pouvoir réaliser des calculs de la fonction objectif plus tard.

- ***Contribution #2 : Une méthode de personnalisation de l'apprentissage basée sur le calcul de l'ADN***

En plus du calcul du poids calculé précédemment, un calcul de la fonction objectif est effectué automatiquement en utilisant la méthode du calcul de l'ADN qui joue un rôle dans la recherche de la séquence personnalisée. Cette méthode a montrée son efficacité à résoudre les problèmes combinatoires et obtenir ainsi des résultats fiables et satisfaisants.

- ***Contribution #3 : un système de personnalisation PERSO***

Le système de personnalisation est un système qui se base essentiellement sur deux formules : une formule de calcul du poids des relations dans le graphe, et une formule de minimisation les poids.

2. Perspectives

Les travaux effectués dans cette thèse dégagent de nombreuses perspectives de recherche qui méritent d'être explorés que ce soit pour l'approche proposée pour la personnalisation des objets pédagogiques ou pour la méthode basée sur le calcul de l'ADN .

- vu la taille du graphe et la quantité d'objets pédagogiques qui le compose, il serait judicieux d'utiliser une ontologie pour modéliser les objets pédagogiques au lieu de les décrire avec des métadonnées et les indexer dans une base de données. Et cela, afin d'ajouter de la sémantique aux différents objets et définir encore plus de hiérarchies et de relations entre eux.
- La méthode du calcul de l'ADN présentée a été validée seulement par un ordinateur, nous envisageons dans nos travaux futurs de faire une application réelle de la méthode dans un laboratoire biologique avec tous les moyens qu'il faut.
- Étendre le système PERSO, par l'intégration d'une interface enseignant expert pour pouvoir lui donner la possibilité de rajouter l'ensemble des objets pédagogiques et effectuer des modifications par rapport au contenu.

Nous envisageons de poursuivre notre travail vers une méthode générique de modélisation des objets pédagogiques. Nous faisons l'hypothèse que l'utilisation du formalisme de conception du graphe associé à une bonne fonction de fitness, peut nous créer une personnalisation des objets pédagogiques.

Ceci constitue une liste bien évidemment non exhaustive des perspectives, l'objectif étant d'améliorer et de faire évoluer la méthode tout en gardant les avantages décrits précédemment.

Dans les systèmes d'apprentissage classiques, les objets pédagogiques sont fixes et pré-stockés. Cependant, dans des systèmes ouverts les objets pédagogiques peuvent être dynamiquement ajoutés, déplacés ou supprimés (par exemple, le système pourrait être en mesure de récupérer un contenu pertinent sur le Web liés aux besoins de l'apprenant). La résolution

de cette variante dynamique du problème de personnalisation de l'apprentissage apporterait une solution à l'évolution des besoins des apprenants et permet un certain degré de flexibilité du contenu lié à la dynamique du Web information. Les approches communautaires représenteraient de bons candidats pour lutter contre un tel problème, car ils ont fait leurs preuves pour être en mesure de s'adapter à des environnements complexes et changeants. De plus, il est maintenant bien connu que leurs hybridation avec d'autres techniques, telles que techniques de recherche ou immunitaires locales peuvent réellement améliorer la capacité du processus de recherche à exploiter des solutions sous-optimales. Les solutions proposées doivent donner plus de considération pour les normes des systèmes d'apprentissage afin d'améliorer l'interopérabilité [Rey-lópez et al. 2008]. En outre, les normes doivent être revues à servir les besoins des systèmes plus généraux d'apprentissage adaptatif [Rey-lópez et al. 2008]. Comme la décentralisation et l'universalité sont les principales propriétés du Web, nous ne nous attendons pas à avoir " parcours d'apprentissage «parfaits». " La décentralisation exige des compromis de [Berners-Lee et al. 2001].

Bibliographie

[Aamodt 2005] Aamodt, A. “Case-based reasoning and Intelligent tutoring”. SAIS-SSLS, Sweden, pp. 8-22, 20-22 avril 2005.

[Acampora et al. 2009] Acampora,G., Gaeta,M., Loia,V. “Hierarchical optimization of personalized experiences for e-Learning systems through evolutionary models”. International journal of Neural Computing Application. doi:[10.1007/s00521-009-0273-z](https://doi.org/10.1007/s00521-009-0273-z), 2009.

[Adelman 1998] Adleman, L.M. “Computing with DNA the manipulation of DNA to solve mathematical problems in redefining what is meant by computation”, international journal of Science America, Vol 279, pp. 34–40. 1998.

[Ainsworth 2000] Ainsworth, S. “Redeem : Its authoring tools and human teaching”. Workshop on Modelling Human Teaching Tactics - ITS2000, 2000.

[Al-Muhaideb et al. 2011] Al-Muhaideb, S., Menai, M.E.B. “Evolutionary computation approaches to the Curriculum Sequencing problem”, International Journal of Natural Computing, Vol 10 No 2, pp. 891-920. 2011.

[Anderson et al. 1985] Anderson, J.R., Reiser, B.J. “The LISP tutor . Byte”, vol 10 N°4, pp. 159-175,1985.

[Anjaneyulu 1997] Anjaneyulu, K. S. R. “Concept Level Modelling on the WWW Workshop Intelligent Educational Systems on the World Wide Web”. The 8th World Conference on Artificial Intelligence in Education (AIED, 1997), pp.18-22,1997.

[Aouag 2008] Aouag,S. “ Individualisation de l’apprentissage dans un Système Tuteur Intelligent : cas de l’apprentissage de la lecture dans le système AMICAL” . Thèse doctorat, UNIVERSITE BLAISE PASCAL CLERMONT-FERRAND II, 11 Janvier 2012.

[Arnaud 2004] Arnaud, M. “La gestion des ressources avec les métadonnées”. Journée: Normes et standards éducatifs, 26 mars 2004, Lyon, France. 2004.

[Azough et al. 2007] Azough, S., Bellakih, M. “Re-use of resources for adapted formation to the learner”. The international symposium on computational intelligence and intelligent informatics (ISCIII 2007), pp. 213–217, 2007.

[Bentley et al. 2004] Bentley, T., Miller, R. “Personalisation; Creating the Ingredients for Systematic and Society-wide Change”. Personalised Learning Conference. 2004.

[Berners-Lee et al. 2001] Berners-Lee, T., Hendler, J., Lassila, O. “The semantic web”. *International Journal of Science American*, vol 284 No 5, pp. 35–43. 2001.

[Bessot 2002] Bessot, A. “Cours : Introduction à l'analyse des situations didactiques équipe DDM, Laboratoire Leibniz”. 2002.

[Bouhineau et al. 2005a] Bouhineau, D., Bronner, A., Chaachoua, H., Nicaud, J.-F. “Patrons d'exercices pour APLUSIX : Une étape du développement de l'EIAH occasion d'un travail entre didacticiens et informaticiens”. *Environnements Informatiques pour l'Apprentissage Humain (EIAH'2005)*, Montpellier, France, pp. 377-382, 25-27 mai 2005.

[Bouhineau et al. 2005b] Bouhineau, D., Bronner, A., Chaachoua, H., Nicaud, J.-F. “Informaticiens et didacticiens peuvent-ils travailler ensemble ? Une étape du développement de l'EIAH APLUSIX : la mise en place de patrons d'exercices et d'une carte de tests”, *Les cahiers Leibniz*, Grenoble, France, mai 2005.

[Bourda 2002] Bourda, Y. « *Des objets pédagogiques aux dossiers pédagogiques (via l'indexation)* ». Document numérique, Vol 6, pp. 115-128. DOI : 10.3166/dn.6.1-2.115-128. 2002.

[Brost 1997] Brost, WN. "Construction of Engineering Ontologies for Knowledge Sharing and Reuse". Ph.D. Dissertation, University of Twente. 5 september 1997.

[Brousseau 1986] Brousseau, G. "La théorisation des phénomènes d'enseignement des mathématiques". Thèse d'Etat. 1986.

[Brusilovsky et al.1996] Brusilovsky, P.,Schwarz, E.,Weber,G. "ELM-ART: an intelligent tutoring system on World Wide Web". Third international conference on intelligent tutoring systems, (ITS 1996), pp. 261–269. 1996.

[Brusilovsky et al. 1998] Brusilovsky, P., Eklund,J., Schwarz,E. " Web-based Education for All: A Tool for Development Adaptive Courseware" . Seventh International World Wide Web Conference, Computer Networks and ISDN Systems, pp. 291-300, 14-18 avril 1998.

[Brusilovsky et al. 1999] Brusilovsky,P. "Adaptive and Intelligent Technologies for Web-based Education" . Künstliche Intelligenz, international Journal of Artificial Intelligence in Education, Vol 13, pp 159-169, 1999.

[Bull et al. 2003] Bull, S., McEvoy, A. T., Reid, E. "Learner Models to Promote Reflection in Combined Desktop PC / Mobile Intelligent Learning Environments". Workshop on Learner Modelling for Reflection, International Conference on Artificial Intelligence in Education 2003 (AIED'2003), Sydney, Australia, pp. 199-208, 2003.

[Burgos et al. 2005] Burgos, D., Michel Arnaud, M., Patrick Neuhauser, P. Koper, R. "IMS Learning Design : la flexibilité pédagogique au service des besoins de l'e-formation". Association EPI Décembre 2005. Disponible à <http://www.epi.asso.fr/revue/articles/a0512c.htm>. 2005.

[Burton 1982] Burton, R. R. "Diagnosing bugs in a simple procedural skill". Intelligent Tutoring Systems. London, Academic Press, 1982.

[Campbell et al. 2007] Campbell, R.J., Robinson, W. “Personalized learning: Ambiguities in theory and practice”. *British Journal of Educational Studies*, vol 55 No 2, pp. 135-154, 2007. ISSN 0007-1005. 2007.

[Champaign et al. 2010] Champaign, J., Cohen, R. “A model for content sequencing in intelligent tutoring systems based on the ecological approach and its validation through simulated students”. *The twenty-third international Florida artificial intelligence research society conference (FLAIRS 2010)*, pp. 486–491. 2010.

[Chandrasekaran et al. 1999] Chandrasekaran, B., Josephson, J.R. “The Ontology of Tasks and Methods”, *AAAI Technical Report SS-97-06*. 1999.

[Chen et al. 2006] Chen, C., Liu, C., Chang, M. “Personalized curriculum sequencing utilizing modified item response theory for web-based instruction”. *International Journal of Expert System Application*, vol 30 No2, pp. 378–396, 2006.

[Chen et al. 2008] Chen, C., Peng, C., Shiue, J. “Ontology-based concept map for planning personalized learning path”. *The IEEE conference on cybernetics and intelligent systems (CCIS 2008)*, pp. 1337–1342, 2008.

[Chen 2008] Chen, C. “Intelligent web-based learning system with personalized learning path guidance”. *International Journal of Computers and Education*, vol 51, pp. 787–814, 2008.

[Chu et al. 2009] Chu, C., Chang, Y., Tsai, C., “PC2PSO: Personalized e-Course Composition based on Particle Swarm Optimization”. *International Journal of Applied Intelligence*, vol 34 No 1, pp. 141-154, 2009.

[Colorni et al. 1991] Colorni,A., Dorigo, M., Maniezzo,V. “ Distributed Optimization by Ant Colonies”. The First European Conference on Artificial Life, pp. 134–142,1991.

[Cueilliez 2006] Cueilliez, O. “ *Composition de documents électroniques personnalisés basée sur des ontologies : Application au e-Learning*”. Mémoire présenté en vue d’obtenir le diplôme d’ingénieur en Informatique, conservatoire national des arts et métiers, 2006.

[Debbah et al. 2014] Debbah,A., Mohamed Ben Ali.Y. “Solving the Curriculum Sequencing Problem With DNA Computing”, Journal of Distance Education Technologies,Vol 12 No 4, pp. 1-18, ISSN: 1539-3100,Octobre Novembre 2014.

[De Bra et al. 1998] De Bra, P., Calvi,L. “ AHA! An open Adaptive Hypermedia Architecture”. The New Review of Hypermedia and Multimedia, vol 4, pp. 233-254, 1998.

[De La Passardière et al. 2004] De La Passardière,B., Jarraud,B., ManUeL,P. “ un profil d'application du LOM pour C@mpuSciences ”. *international journal of Sciences et Technologies de l'Information et de la Communication pour l'Education et la Formation*, Vol 11, pp.11-57. 2004.

[De-Marcos et al. 2007] De-Marcos, L., Pages,C., Martinez, JJ., Gutiérrez,JJ. “Competency-based learning object sequencing using particle swarms”. The 19th IEEE international conference on tools with artificial intelligence, vol 2, pp. 111–116.2007.

[De-Marcos et al. 2008a] De-Marcos, L., Martinez, JJ., Gutiérrez, JA. “Swarm intelligence in e-learning: a learning object sequencing agent based on competencies”. The 10th annual conference on genetic and evolutionary computation. ACM Special Interest Group on Genetic and Evolutionary Computation, pp. 17–24. 2008.

[De-Marcos et al. 2008b] De-Marcos, L., Barchino,R., Martinez,J.,Gutiérrez,J., Hilera,J. “Competency-based intelligent curriculum sequencing: comparing two evolutionary

approaches”. The IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, vol 3, pp. 339–342. 2008.

[De-Marcos et al. 2009a] De-Marcos, L., Barchino, R., Martinez, J., Gutiérrez, J. “A new method for domain independent curriculum sequencing: a case study in a web engineering master program”. International Journal of Engineering Education, vol 25 No 4, pp. 632–645, 2009.

[De-Marcos et al. 2009b] De-Marcos, L., Martinez, J., Gutiérrez, JA., Barchino, R., Gutiérrez, JM. “A new sequencing method in web-based education”. The IEEE congress on evolutionary education (CEC, 2009), pp. 3219–3225. 2009.

[Dorigo 1992] Dorigo, M. “Optimization, learning and natural algorithms”. PhD thesis, Politecnico di Milano, Italie. 1992.

[Dorigo et al. 1996] Dorigo, M., Maniezzo, V., Colorni, A. “Ant System : optimization by a colony of cooperating agents”. IEEE Transactions on Systems, Man, and Cybernetics, Part B : Cybernetics, Vol. 26. pp. 29–4, 1996.

[Duclosson et al. 2005] Duclosson, N., Jean-Daubias, S., Riot, S. “AMBRE-enseignant : un module partenaire de l’enseignant pour faciliter l’intégration de AMBRE en classe”, Rapport de recherche, LIRIS, Villeurbanne, France, 2005.

[Eberhart et al. 1995] Eberhart, R., Kennedy, J. “A new optimizer using particle swarm theory”. The sixth international symposium on micro machine and human science, pp. 39–43. 1995.

[Eberhart et al. 1998] Eberhart, R.C., Shi, Y.V.W., Porto, N.S., Eiben, A.E. “Comparison between genetic algorithms and particle swarm optimization”. The 7th International Conference on Evolutionary Programming VII, pp. 611-616. 1998.

[Eberhart et al. 2007] Eberhart,R., Shi,Y. “Computational intelligence ”. Morgan Kaufmann Publishers, Madison. 2007.

[El Dor 2012] El Dor,A. “Perfectionnement des algorithmes d'optimisation par essaim particulaire : applications en segmentation d'images et en _electronique”, Thèse de Doctorat, UNIVERSITÉ PARIS-EST, 5 décembre 2012.

[El Mezouary et al. 2011] El Mezouary, A., Battou, A., Cherkaoui, C., Mammass, D. (2011). “A semantic model for improving the retrieval of Learning Objects in AHS”. *Acte de la conférence internationale Multimedia Computing and Systems (ICMCS)*, du 7 au 9 avril à Ouarzazate. Disponible à http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5945632. 2011.

[Eiben et al. 2007] Eiben,G., Smith,J. “Introduction to evolutionary computing”. Natural computing series. ISBN: 978-3-642-07285-7. 2007.

[Fogel et al. 1965] Fogel, L., Owens, A., Walsh, M. “Artificial intelligence through a simulation of evolution”. In Callahan A, Maxfield M, Fogel LJ (eds) *Biophysics and cybernetic systems*. Pp. 131–155. 1965.

[Gil et al. 2008] Gil,A., Penalvo.FJ. “Learner course recommendation in e-Learning based on swarm intelligence”. *International Journal of Universal Computer Science*, vol 14 No 16, pp. 2737–2755, 2008.

[Girard et al. 2007] Girard, S., Johnson, H. “DividingQuest: opening the learner model to teachers”. *Environnements Informatiques pour l'Apprentissage Humain (EIAH'2007)*, Lausanne, Suisse, pp. 329-334, 27-29 juin 2007.

[Goldberg 1989] Goldberg,D.E. “Genetic Algorithms”. Adison-Wesley, 1989. ISBN 0201157675. 1989.

[Goss et al. 1989] Goss, S., Aron, S., Deneubourg, J. L., Pasteels, J. M. “Self-organized shortcuts in the Argentine ant”. *International Journal of Naturwissenschaften*, Vol 76 No 12, pp. 579–581, 1989.

[Gruber 1993] Gruber, T. “A translation approach to portable ontology specifications”, *International journal of Knowledge Acquisition*, Vol. 5, No. 2.1993

[Guéraud et al. 2007] Guéraud, V., Michelet, S., Adam, J.-M. “Suivi de classe à distance : propositions génériques et expérimentation en électricité”. *Environnements Informatiques pour l'Apprentissage Humain (EIAH'2007)*, Lausanne, Suisse, pp. 167-178, 27-29 juin 2007.

[Guo 2008] Guo,Q. “Implement individual web-based learning environment”. *The 19th international conference on database and expert systems application*, pp. 639–643. IEEE. doi:[10.1109/ DEXA.2008.17](https://doi.org/10.1109/DEXA.2008.17). 2008.

[Guo 2009] Guo,Q. “Implement web learning environment based on data mining”. *International Journal of Knowledge Based Systems*, vol 22, pp. 439–442. 2009.

[Gutiérrez et al. 2007] Gutiérrez,S., Pardo,A. “Sequencing in web-based education: approaches, standards and future trends”. *Studies in Computational Intelligence (SCI)*, vol 62, pp. 83–117, 2007.

[Gutiérrez et al. 2007] Gutiérrez, S., Valigiani, G., Collet,P., Kloos,CD. “Adaptation of the ACO heuristic for sequencing learning activities”. *The EC-TEL 2007 poster session*, pp. 17–20. doi:[10.1.1.142.8362](https://doi.org/10.1.1.142.8362). 2007.

[Hauger et al. 2007] Hauger,D., Kock,M. “State of the art of adaptivity in e-Learning platforms”. *Workshop at adaptivity and user modeling in interactive systems (ABIS 2007)*, Halle/Salle, Germany. 2007. doi:[10.1.1.91.6409](https://doi.org/10.1.1.91.6409).2007.

[Hofmeister et al. 2002] Hofmeister,T., Schöning,U., Schuler,R., Watanabe,O. “Annual Symposium on Theoretical Aspects of Computer Science” (STACS), pp.192-202, 2002.

[Hofmann 1990] Hofmann, Y. "Introduction de l'informatique dans la formation en soins infirmiers". *Informatique et Santé*, vol. 3, pp. 17-24, 1990.

[Höök et al. 1996] Höök, K., KARLGREN, K, J., WAERN, A. "A glass box approach to adaptive hypermedia". *International Journal of User models and User adaptive interaction*, vol 6, pp. 157-184. 1996.

[Holland 1975] Holland, J. H. "Adaptation in natural and artificial systems". MIT Press, Cambridge, 1992. 1st edition:1975, The University of Michigan Press, Ann Arbor. 1975.

[Holland 1992] Holland, J. H. "Adaptation in natural and artificial systems". MIT Press, 1992.

[Houchmandzadeh 2004] B.Houchmandzadeh . "Ballade autour de quelques thèmes de biologie". 2004.

[Hovakimyan et al. 2004] Hovakimyan,A., Sargsyan,S. Barkhoudaryan,S. "Genetic algorithm and the problem of getting knowledge in e-Learning systems". The IEEE international conference on advanced learning technologies (ICALT 2004), pp. 336–339, 2004.

[Huang et al. 2007] Huang,M., Huang,H., Chen,M. " Constructing a personalized e-Learning system based on genetic algorithm and case-based reasoning approach". *International Journal of Expert System with Application*, vol 33, pp. 551–564. 2007.

[Hû 2001] Hû,O. "Contribution à l'évaluation des logiciels multimédias pédagogiques", Thèse de doctorat, UNIVERSITE DE TECHNOLOGIE DE COMPIEGNE, 30 août 2001.

[Hwang et al. 2008] Hwang,G., Yin,P., Wang,T., Tsengand,J., Hwang,G. "An enhanced genetic approach to optimizing autoreply accuracy of an e-Learning system". *International Journal of Computers & Education* vol 51No 1, pp. 337–353, 2008.

[Janssen et al. 2005] Janssen, J., Tattersall, C., Waterink, W., Den Berg, B., van Es, R., Bolman, C., Koper, R. “Self-organising navigational support in lifelong learning: how predecessors can lead the way”. *International Journal of Computers & Education*, vol 49 No 3, pp. 781–793. 2005.

[Jean-Daubias 2004] Jean-Daubias, S. “De l'intégration de chercheurs, d'experts, d'enseignants et d'apprenants à la conception d'un EIAH”. *Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie (TICE'2004)*, Compiègne, France, pp. 290-297, 20-22 octobre 2004.

[Jean-Daubias et al. 2009] Jean-Daubias, S., Guin, N. “AMBRE-teacher: a module helping teachers to generate problems”. *2nd Workshop on Question Generation, International Conference on Artificial Intelligence in Education (AIED 2009)*, Brighton, Grande-Bretagne, AIED 2009 Workshops Proceedings 1, pp. 38-42, 6-7 juillet 2009.

[Johnson 1986] Johnson, W.L. “Intention-based diagnosis of errors in novice programs”. Palo Alto, CA, Morgan Kaufmann Series In Research Notes In Artificial Intelligence, 1986.

[Kay et al. 1997] Kay, J., Kummerfeld, B. “User models for customized hypertext”. *Intelligent hypertext: Advanced techniques for the World Wide Web. Nicholas*, vol 13 No 26, pp. 47-69, 1997.

[Kennedy et al. 1995] Kennedy, J., Eberhart, R.C. “Particle swarm optimization”. *The IEEE international conference on neural networks, (ICNN 1995)*, pp. 1942–1948. 1995.

[Koper 2001] Koper, R. “Modeling units of study from a pedagogical perspective. The pedagogical meta-model behind EML”. Open University of the Netherlands 2001, <http://eml.ou.nl/introduction/articles.htm> (consulté Mai 2009).

[Koza 1992] Koza, J.R. “Genetic programming”. MIT Press, Cambridge. 1992.

[Lefevre 2009] Lefevre,M. “Processus unifié pour la personnalisation des activités pédagogiques : méta-modèle, modèles et outils”. Thèse de doctorat, UNIVERSITÉ CLAUDE BERNARD LYON 1, 1er Décembre 2009.

[Leroux 2002] Leroux, P. “ Machines partenaires des apprenants et des enseignants - Étude dans le cadre d'environnements supports de projets pédagogiques”. Habilitation à Diriger des Recherches en Informatique, Université du Maine, 20 décembre 2002.

[Leroux 2006] Leroux, P. “Micromondes et robotique pédagogique”. Environnements informatiques pour l'apprentissage humain (Traité IC2, série Cognition et traitement de l'information). M. Grandbastien, J.-M. Labat, pp. 311-332, 2006.

[Lionnet et al. 2005] Lionnet,T., Croquette,V . “Introduction à la Biologie Moléculaire ”, 25 mars.URL : <http://www.phys.ens.fr/~biolps/>. 2005.

[López et al. 1998] López, J., Millán, M.E., Pérez-De-La-Cruz, J.L., Triguero, F. “ ILESA: a Web-based Intelligent Learning Environment for the Simplex Algorithm ”. The 4th International conference on Computer Aided Learning and Instruction in Science and Engineering (CALISCE, 1998), Göteborg, Sweden, pp. 399-406, 1998.

[Lopes 2009] Lopes, RS., Fernandes, MA. “Adaptative instructional planning using workflow and genetic algorithms”. The eighth IEEE international conference on computer and information science (ACIS 2009), pp. 87–92. 2009.

[Malaxa 2005] Malaxa,V. “A Framework for Metadata Creation Tools”. *Journal of Knowledge and Learning Objects*,Vol 1, pp. August 2005.

[Mitrovic 1998] Mitrovic, A. “A Knowledge-Based Teaching System for SQL”. Association for the Advancement of Computing in Education (AACE), 1998.

[Mufti-Alchawafa et al. 2004] Mufti-Alchawafa, D., Luengo, V., Vadcard, L. “ Architecture d’un environnement d’aide à l’apprentissage de la chirurgie orthopédique”. Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie (TICE'2004), Compiègne, France, 20-22 octobre 2004.

[Murray et al. 1998] Murray,T., Condit,C., Haugsjaa,E. “MetaLinks: “A preliminary framework for concept-based adaptive hypermedia” . Workshop "WWW-Based Tutoring" at "4th International Conference on Intelligent Tutoring Systems (ITS 1998), San Antonio, TX, 1998.

[Murray 2003] Murray,T. “MetaLinks: Authoring and Affordances for Conceptual and Narrative Flow in Adaptive Hyperbooks”. International Journal of Artificial Intelligence in Education (IJAIED), vol. 13, pp. 199-233, 2003.

[Murray 2003a] Murray, T. “Eon: Authoring Tools for Content, Instructional Strategy, Student Model, and Interface Design”. Authoring Tools for Advanced Technology Learning Environments. T. Murray, S. B., and S. Ainsworth, editors, Kluwer Academic Publisher, 2003.

[Nicaud et al. 2003] Nicaud, J.-F., Bouhineau, D., Chaachoua, H., Huguet, T., Bronner, A. “A computer program for the learning of algebra: description and first experiment”. Eleventh International PEG Conference, St. Petersburg, Russie, juin 2003.

[Paquette et al. 2002] Paquette,g., Rosca, l. “Organic Aggregation of Knowledge Objects in Educational Systems Center for Inter-university Research on Telelearning Applications” CIRTA-LICEF, Télé-université, Université du Québec. 2002.

[Pernin 2003] Pernin, J.P. “Objets pédagogiques : unités d’apprentissage, activités ou ressources” Revue "Sciences et Techniques Educatives", Hors série 2003 " Ressources numériques, XML et éducation", pp 179-210, avril 2003, éditions Hermès.

[Rechenberg 1965] Rechenberg,I. “Cybernetic solution path of an experimental problem”. Royal Aircraft Establishment, UK.1965.

[Rey-lópez et al. 2008] Rey-lópez,M., Brusilovsky,P., Meccawy, M., Daz-Redondo, R., Fernandez Vilas, A., Ashman, H. “Resolving the problem of intelligent learning content in learning management systems”. International Journal of e-Learning, vol 7 No 3, pp. 363–381. 2008.

[Ríos et al. 1999] Ríos,A., Millán,E., Trella,M., Pérez-De-La-Cruz,J.R., Conejo,R. “Internet Based Evaluation System” . The *9th World Conference of Artificial Intelligence and Education (AIED, 1999)*, pp. 387-394, 1999.

[Romero et al. 2002] Romero,C., Bra,PD., Ventura,S., De Castro,C. “Using knowledge levels with AHA! for discovering interesting relationships”. World conference on e-Learning in corporate, government, healthcare, and higher education, pp. 2721–2722, 2002.

[Sanchita et al. 2008] Sanchita,P., De Sarkar,A. “Solution to the 0/1 knapsack problem based on DNA computing” , International journal of theoretical and applied information technology ,vol 4 No 6, pp. 526-529. 2008.

[Seki et al. 2005] Seki,K., Matsui, T.,Okamoto,T. “An adaptive sequencing method of the learning objects for the elearning environment”. International Journal of Electronics and Communications in Japan, Vol 88, No. 3, pp. 54–71, 2005.

[Self 1994] Self,J. “The role of student models in learning environments”, Transactions of the Institute of Electronics, Information and Communication Engineers, vol 77 No1, pp. 3-8. 1994.

[Semet et al. 2003a] Semet,Y., Jamont,Y., Biojout,R., Lutton,E., Collet,P. “Ant Colony Optimization for e-Learning: observing the emergence of pedagogic suggestions”. The IEEE swarm intelligence symposium, pp. 46–52, 2003.

[Semet et al. 2003b] Semet,Y., Jamont,Y., Biojout,R., Lutton,E., Collet,P . “Artificial ant colonies and e-Learning: an optimization of pedagogical paths”. The 10th international conference on humancomputer interaction (HCII, 2003), 2003.

[Shwefel 1965] Shwefel,H. “Kybernetische Evolution als Strategie der Experimentellen Forschung in der Stromungstechnik”. Diploma thesis, Technical University of Berlin, Germany,1965.

[Song et al. 2001] Song,Y.H., Irving,M.R. “Optimization techniques for electrical power systems Part 2 Heuristic optimization methods”. International Journal of Power Engineering vol 15 No 3, pp. 151–160, 2001.

[Sørmo et al. 2002] Sørmo, F., Aamodt, A. “Knowledge communication and CBR”. *6th European Conference on Case-Based Reasoning (ECCBR 2002)*, pp. 47-59, 2002.

[Subrata et al. 2003] Subrata,R., Zomaya, A. Y. “A comparison of three artificial life techniques for reporting cell planning in mobile computing”. IEEE Transactions on Parallel and Distributed Systems, vol 14 No 2, pp. 142–153,2003.

[Tattersall et al. 2005] Tattersall,C., Manderveld,J., den Berg,B., van Es,R., Janssen,J., Koper,R. “Self organising wayfinding support for lifelong learners”. *International journal of Education and Information Technologies*, Vol 10 No 1-2, pp. 111-123, 2005.

[Tchounikine 2009] Tchounikine,P. “Précis de recherche en ingénierie des EIAH ”, en ligne sur le Web, 2009 (<http://membres-liglab.imag.fr/tchounikine/Precis.html>).2009.

[Tsang 1993] Tsang,E. “Foundations of constraint satisfaction”. Academic Press, London, ISBN 0-12-701610-4. 1993.

[Valigiani et al. 2006] Valigiani,G., Lutton,E., Jamont,Y., Biojout,R., Collet,P. “Automatic rating process to audit a man-hill”. WSEAS Transaction on Advanced Engineering Education ,vol 3 No 1, pp. 1–7, 2006.

[Wang et al. 2008] Wang,X. , Bao,Z., Hu,J., Wang,S., Zhan,A. “Solving the SAT problem using a DNA computing algorithm based on ligase chain reaction”. International journal of bio systems BioSystems Vol 91, pp. 117–125, 2008.

[Wang et al. 2008] Wang,T., Wang,K., Huang,Y. “Using a style-based ant colony system for adaptive learning”. Interantional Journal Expert System with Application, vol 34 No 4, pp. 2449–2464. 2008.

[Wang et al. 2009] Wang, TJ., Tsai, KH. “Interactive and dynamic review course composition system utilizing contextual semantic expansion and discrete particle swarm optimization”. International journal of Expert System with Application. Vol 36 No 6. pp. 9663–9673. 2009.

[Watson et al. 1953] Watson, J.D.,Crick, F.H.C. “Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid”, Nature (London), vol 171,pp. 737-738.1953.

[Weber 1996] G,Weber. “Individual selection of examples in an intelligent learning environment”. International *Journal of Artificial Intelligence in Education*, volume. 7 N°1, pp.3-31,1996.

[Wiley 2002] Wiley, D. A. “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy”, in David A. Wiley, *The Instructional Use of Learning Objects*». 2002.

[Wong et al. 2009] Wong, L., Looi, C. “Adaptable learning pathway generation with ant colony optimization”. *International Journal of Education Technology and Society*, Vol 12 No 3, pp. 309–326. 2009.

[Yahyia 2005] Yahyia, H. “Le Miracle de la Création dans l'ADN, Harun Yahya International”. ISBN: 2-914949-35-9. 2005.

[Yang et al. 2009] Yang, YJ., Wu, C. “An attribute-based ant colony system for adaptive learning object recommendation”. *International Journal of Expert System with Application*, vol 36 No 2, pp. 3034–3047, 2009.

[Yang et al. 2005] Yang, JT., Yu, PT., Chen, NS., Tsai, CY., Lee, CC., Shih, TK. “Using ontology as scaffolding for authoring teaching materials”. *International Journal of Distance Education Technologies*, vol 3 No1, pp. 81–96. 2005.

Webographie

[Meirieu 1994] Meirieu,P. “individualisation, différenciation, personnalisation : de l'exploration d'un champ Sémantique aux paradoxes de la, formation”, <http://www.meirieu.com/articles/individualisation.pdf>

[Rodet 2009] Rodet,J. “*Individualisation*”. Disponible à <http://www.novantura.com/wiki/wakka.php?wiki=individualisation>. Date d'accès : 2009.

[LiliMath 2000] “LiliMath”. <http://lilimath.free.fr/> (dernière visite : septembre 2009).

[Teleos 2003] “Teleos”.<http://www-clips.imag.fr/arcade/projets/teleos/TELEOS.html> (dernière visite : septembre 2009).

[Cours de SVT 2de 2015] “universalité de l'ADN”. Valable au <http://www.maxicours.com/se/fiche/3/3/356033.html/2e> (dernière visite: 2015).

[Niso2004]Niso.“UnderstandMetadata”,
<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.

[W3C 2004] “*OWL Web Ontology Language Guide*”. W3C Recommendation 10 February 2004 . <http://www.w3.org/TR/owl-guide/>. (dernière visite 2015)

[Articulate StoryLine 2015] “articulate storyline”. <https://fr.articulate.com/>, (dernière visite 2015).

Annexe A

(1). Quelques interfaces du système PERSO

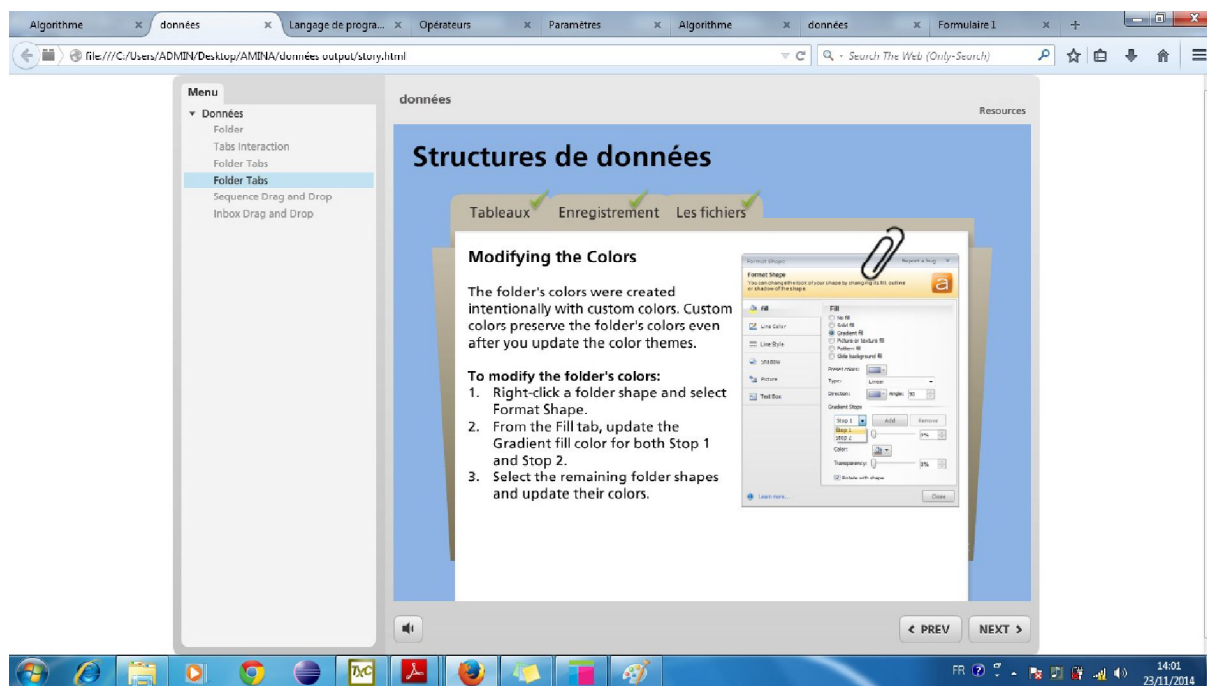


FIG.A.1– Exemple d'un objet pédagogique

The screenshot shows a web browser window with the URL `127.0.0.1:32767/start.html#?pre=formulaire_2`. The browser tabs include 'Algorithme', 'données', 'Langage de progra...', 'Opérateurs', 'Paramètres', 'Algorithme', 'données', 'Formulaire 2', and 'Formulaire 2'. The left sidebar shows a 'Sitemap' with 'Home', 'Formulaire 1', and 'Formulaire 2'. The main content area contains a form with the following questions and options:

- Quel est votre type de ressource préférée ?
 - Exercices
 - Simulation
 - Texte narrative
 - Figure
 - Graphe
- Quel est votre type d'interactivité ?
 - Active
 - expositif
 - Mixte
- Quel est votre niveau d'interactivité ?
 - trop bas
 - Bas
 - Moyen
 - Supérieur
- Quel est votre langue préférée ?
 - Anglais
 - Français
- Quel est votre temps d'apprentissage ?
 - Facile
 - Moyen
 - Difficile
- Quel est le contexte dans vous vous y trouver ?
 - Education primaire
 - Education secondaire
 - Education universitaire

Buttons at the bottom: '<<Précédent' and 'Valider'.

FIG.A.2– Formulaire des préférences

The screenshot shows a web browser window with the URL `127.0.0.1:32767/start.html#?pre=formulaire_1`. The browser tabs include 'Algorithme', 'données', 'Langage de progra...', 'Opérateurs', 'Paramètres', 'Algorithme', 'données', 'Formulaire 1', and 'Formulaire 1'. The left sidebar shows a 'Sitemap' with 'Home', 'Formulaire 1', and 'Formulaire 2'. The main content area contains a form with the following questions and options:

Veuillez remplir ce formulaire :

- c'est quoi les éléments de bases d'un algorithme ?
 - Déclaration, types, variables
 - Déclaration, variable
 - Déclaration
- Quelles sont les différents types de données ?
 - Caractère, entier, double
 - Double, entier
 - Réel, double, entier, caractère
- Quels sont les éléments nécessaire pour la déclaration ?
 - Type de données, variable
 - variables seulement
 - Type de données
- citer les langages de programmation les plus connus ?
 - C++, C#, Java, XML
 - C++, java
 - HTML, C++, java

Button at the bottom: '>>Suivant'.

FIG.A.3– Exemple d'un QCM

Publications réalisées au cours de la thèse

Publications et conférences internationales

- Amina.Debbah, Bensebaa.Tahar, Hafidi.Mohamed, Yamina Mohamed Ben Ali “Adaptation of the links in an educational hypermedia”, In proceeding of [Education and e-Learning Innovations \(ICEELI\), 2012.](#) DOI: [10.1109/ICEELI.2012.6360676](#)
Publication Year: 2012 , Page(s): 1 - 6 2012.
- Amina.Debbah, Yamina Mohamed Ben Ali: “Solving the Curriculum Sequencing Problem with DNA computing”, International journal of Distance Education Technologies, Vol 12, (4),pp. 1-18, 2014.