

N° d'ordre : 2529

Thèse

présentée

en vue de l'obtention du

**DOCTORAT DE L'UNIVERSITE DE TOULOUSE
DELIVRE PAR L'INSTITUT NATIONAL POLYTECHNIQUE DE TOULOUSE**

École doctorale : **Informatique et Télécommunications**

Spécialité : **Informatique**

Par M : **Christophe Piombo**

Titre de la thèse **Modélisation probabiliste du style d'apprentissage et application à l'adaptation de contenus pédagogiques indexés par une ontologie**

Soutenue le 30 octobre 2007 devant le jury composé de :

M. Jean-Marc Labat

Président et
Rapporteur

MM.

Alain Ayache
Mohamed Quafafou
Hadj Batatia
Thierry Nodenot
Philippe Vidal

Directeur de thèse
Rapporteur
Membre (encadrant)
Membre
Membre

Epigraphe

« Quoique j'invente, quoique je fasse, je serai toujours au dessous de la vérité. Il viendra toujours un moment où les créations de la science dépasseront celles de l'imagination »

Jules Verne

Remerciements

J'aimerais en premier remercier deux hommes sans qui cette thèse n'aurait pu voir le jour et se terminer. Je remercie le Professeur Alain Ayache, mon Directeur de thèse et directeur de l'ENSEEIH, pour avoir soutenu mon dossier d'inscription en thèse et m'avoir ainsi permis de réaliser mon projet d'études. Je remercie le docteur Hadj Batatia, mon co-encadrant et maître de conférences à l'IPST-Cnam, de m'avoir pendant ces 6 années, soutenu, aidé et poussé à terminer cette thèse. Sa capacité à encadrer un étudiant, sa compétence dans le domaine scientifique et sa rigueur professionnelle m'ont servis d'exemple dans l'apprentissage du métier de chercheur.

Je suis sensible à l'honneur que me font les professeurs Jean-Marc Labat et Mohamed Quafafou en rapportant sur ma thèse. La qualité de leurs remarques m'ont permis d'améliorer le contenu de ce manuscrit.

Je remercie le professeur Philippe Vidal ainsi que Thierry Nodenot, maître de conférences, d'avoir acceptés d'examiner cette thèse et de participer au jury de cette dernière.

Durant cette thèse j'ai mené des études empiriques qui n'auraient pu se faire sans l'aide de certaines personnes. C'est pourquoi j'aimerais remercier les collègues de l'Université du Mirail, à savoir Christian Escribe, Nathalie Huet et Claudette Mariné qui ont fait passer à leurs élèves le questionnaire d'évaluation de style d'apprentissage et à Caroline Dupeyrat pour m'avoir aidé à le traduire en Français. J'ajoute à ces remerciements Claude Bégué principal du collège de Ramonville et Michelle Castaing professeur de mathématiques pour m'avoir permis de travailler avec leurs élèves de 5^{ème}.

Je remercie le professeur Michel Daydé de m'avoir accueilli dans le laboratoire LIMART où j'ai pu réaliser mes travaux de recherche ainsi que Sylvie Eichen, Josiane Gambino et Annie Adelantado les différentes secrétaires du laboratoire qui se sont occupées de mes différents dossiers administratifs et dossiers de missions.

Je remercie aussi l'ensemble des thésards du laboratoire pour leur aide à un moment ou un autre tout au long de ces années de thèse et plus particulièrement Zehour Ouksili, Clovis Tauber et Amhed Touhami pour leur amitié, leur gentillesse et leur proximité.

Enfin, je voudrais ajouter à ces remerciements mes collègues de travail à l'IPST-Cnam qui m'ont toujours soutenu et aidé dans cette activité par leur gentillesse et leur disponibilité. Un grand merci tout particulièrement à Karim Achaïbou, responsable du département informatique, et aux différents enseignants de ce même département. Je veux nommer Danielle Barbarou, Bertrand Raïff et Louis Randriamihamison. Que soit aussi remerciée Valérie Crouzil, secrétaire du département informatique, pour sa disponibilité et sa bonne humeur rayonnante.

Je voudrais ensuite terminer en remerciant toute ma famille d'avoir acceptée pendant ces 6 années, mon manque de disponibilité à cause de l'importante masse de travail à fournir dans le contexte de cette thèse.

A mon fils Florent et mon épouse Valérie...

Table des matières

EPIGRAPHE	3
REMERCIEMENTS.....	5
TABLE DES MATIERES	7
LISTE DES FIGURES.....	11
LISTE DES TABLEAUX	15
INTRODUCTION GENERALE.....	17
PROBLEMATIQUE.....	20
CONTRIBUTIONS	22
PLAN DE LA THESE	22
1. ETAT DE L'ART.....	25
1.1. INTRODUCTION	27
1.2. MODELISATION DE L'ELEVE.....	28
1.2.1. <i>Type du modèle.....</i>	28
1.2.2. <i>Données du modèle : connaissances et préférences.....</i>	33
1.2.3. <i>Acquisition du modèle : statique, dynamique ou mixte.....</i>	43
1.2.4. <i>Discussion.....</i>	45
1.3. MODELISATION DE CONTENUS	46
1.4. STRATEGIES D'ADAPTATION.....	50
1.4.1. <i>Adaptation de la présentation.....</i>	51
1.4.2. <i>Adaptation de la navigation.....</i>	51
1.4.3. <i>Modèle d'adaptation à plusieurs couches</i>	53
1.4.4. <i>Incertitude dans le choix de l'adaptation</i>	58
1.4.5. <i>Discussion.....</i>	62
1.5. STYLE D'APPRENTISSAGE.....	63
1.5.1. <i>Différentes approches et modèles.....</i>	63
1.5.2. <i>Style d'apprentissage et performances de l'élève.....</i>	64
1.5.3. <i>Style d'apprentissage et documents adaptatifs.....</i>	65
1.6. WEB2.0 : MODELES DE L'ELEVE ET DE CONTENUS.....	67
1.6.1. <i>Les ontologies.....</i>	67
1.6.2. <i>Ontologies et systèmes d'enseignement.....</i>	76
1.6.3. <i>Ontologies et « profiling ».....</i>	77
1.7. CONCLUSION.....	80
2. ACTIVITE DE PROFILAGE : ETUDE STATISTIQUE DU STYLE D'APPRENTISSAGE	81
2.1. INTRODUCTION	83
2.2. META-MODELE CONCEPTUEL DU PROCESSUS DE PROFILAGE.....	84
2.3. MODELE DE DEPENDANCE DU STYLE D'APPRENTISSAGE	86
2.3.1. <i>Indicateurs observables.....</i>	87
2.3.2. <i>Paramètres non observables.....</i>	88
2.3.3. <i>Dépendances entre indicateurs observables et paramètres non observables.....</i>	88
2.4. MESURE DU STYLE D'APPRENTISSAGE : QUESTIONNAIRE ILS.....	90
2.4.1. <i>Echantillon de la population pour l'étude.....</i>	90
2.4.2. <i>Mesures de l'étude.....</i>	90
2.4.3. <i>Conception et mise en œuvre de l'étude</i>	93
2.4.4. <i>Résultats de l'étude.....</i>	93
2.5. INDICATEURS OBSERVABLES POUR INFERER LE STYLE D'APPRENTISSAGE D'UN ELEVE.....	94
2.5.1. <i>Echantillon de la population pour l'étude.....</i>	94
2.5.2. <i>Mesures de l'étude.....</i>	94
2.5.3. <i>Conception de l'étude.....</i>	99
2.5.4. <i>Mise en œuvre de l'étude</i>	105
2.5.5. <i>Résultats de l'étude.....</i>	106
2.6. VARIABLES DE PROFILAGE DU STYLE D'APPRENTISSAGE	114

2.6.1.	<i>Variables non observables : Le style d'apprentissage de l'élève</i>	114
2.6.2.	<i>Variables observables</i>	115
2.7.	CONCLUSION	118
3.	APPRENTISSAGE SUPERVISE DU MODELE DE L'ELEVE	119
3.1.	INTRODUCTION	121
3.2.	SYSTEME ADAPTATIF PROPOSE	122
3.3.	MODELE DE L'ELEVE	124
3.4.	MODELE DE CONNAISSANCES	124
3.4.1.	<i>Les variables</i>	126
3.4.2.	<i>Relations causales : arcs et paramètres du réseau</i>	128
3.5.	MODELE DE PREFERENCES	131
3.5.1.	<i>Conception</i>	131
3.5.2.	<i>Apprentissage</i>	132
3.5.3.	<i>Inférence</i>	134
3.6.	UN RESEAU BAYESIEN POUR INFERER LES PREFERENCES D'UN ELEVE	136
3.6.1.	<i>Aperçu théorique</i>	136
3.6.2.	<i>Structure du réseau</i>	136
3.6.3.	<i>Entraînement du réseau</i>	142
3.6.4.	<i>Inférences et renforcement du réseau</i>	144
3.6.5.	<i>Validation du modèle</i>	146
3.7.	MACHINE A VECTEURS DE SUPPORT POUR INFERER LES PREFERENCES D'UN ELEVE.....	150
3.7.1.	<i>Aperçu théorique</i>	150
3.7.2.	<i>SVM multi-classes pour estimer les préférences d'un élève</i>	155
3.7.3.	<i>Validation du modèle</i>	157
3.8.	CONCLUSION	160
4.	CONTENU ADAPTATIF : MODELE, METHODE ET IMPLEMENTATION	161
4.1.	INTRODUCTION	163
4.2.	MODELE DE CONTENUS.....	164
4.2.1.	<i>Modèle conceptuel du domaine</i>	167
4.2.2.	<i>Modèle conceptuel pédagogique</i>	168
4.2.3.	<i>Modèle conceptuel d'objets d'apprentissage</i>	170
4.2.4.	<i>Décomposition Structurale du modèle de contenus</i>	171
4.3.	METHODE DE PRODUCTION DES CONTENUS	173
4.3.1.	<i>Recueil des besoins de formation du métier</i>	174
4.3.2.	<i>Analyse pédagogique</i>	175
4.3.3.	<i>Conception pédagogique</i>	176
4.3.4.	<i>Développement du matériel pédagogique</i>	177
4.4.	STRUCTURATION STANDARD DES CONTENUS	177
4.4.1.	<i>Nécessité de structurer les ressources</i>	178
4.4.2.	<i>Modèle SCORM d'agrégation de contenus</i>	178
4.4.3.	<i>Structuration proposée</i>	184
4.4.4.	<i>Données d'adaptation proposées</i>	187
4.5.	MISE EN ŒUVRE ET RESULTATS	209
4.6.	CONCLUSION	210
5.	ADAPTATION	211
5.1.	INTRODUCTION	213
5.2.	MESURES DE PERTINENCE EN IR.....	213
5.2.1.	<i>Mesures de similarité par calcul</i>	215
5.2.2.	<i>Mesures de similarité par comparaison</i>	216
5.3.	ADAPTER UNE SEQUENCE PEDAGOGIQUE AU STYLE D'APPRENTISSAGE	217
5.3.1.	<i>Architecture du système adaptatif</i>	217
5.3.2.	<i>Modèle d'adaptation</i>	219
5.3.3.	<i>Mesure de similarité sémantique</i>	222
5.4.	SIMULATIONS ET RESULTATS	224
5.5.	CONCLUSION	227

CONCLUSION GENERALE.....	229
LIMITATIONS.....	232
PERSPECTIVES.....	232
BIBLIOGRAPHIE	235
ANNEXES.....	257

Liste des figures

Figure 1-1.1 Architecture de notre système adaptatif.....	21
Figure 1-1 : Architecture générale des AIWBES	27
Figure 1-2 : Modèle expertise-partielle	29
Figure 1-3 : Modèle différentiel	30
Figure 1-4 : Modèle perturbation.....	31
Figure 1-5 : Réponse incorrecte dans un problème de soustraction de deux nombres.....	31
Figure 1-6 : Modèle social.....	32
Figure 1-7 : Modèle de connaissances et modèle d'attitudes (préférences) de l'élève	33
Figure 1-8 : Exemple de relations AND-OR dans le modèle GM.....	36
Figure 1-9 : le modèle de connaissances de LAOS	37
Figure 1-10 : Représentation XML d'un Modèle de connaissances de l'élève.....	38
Figure 1-11 : couche domaine/couche activité	39
Figure 1-12 : exemple de ressource basée sur les principes de l'intelligence visuel/spatial. 42	
Figure 1-13 : Représentation XML d'un modèle de contenus	47
Figure 1-14 : modèle de contenu du système INSPIRE	48
Figure 1-15 : Modèles d'adaptation (AM) et (PM) du framework LAOS	54
Figure 1-16 : Les trois couches de AM	55
Figure 1-17 : Concepts du modèle GM de LAOS	56
Figure 1-18 : exemple de règles d'adaptation directes	56
Figure 1-19 : Structure FOR- DO.....	57
Figure 1-20 : Le modèle d'adaptation de LAOS dans MOT.....	57
Figure 1-21 : Modèle de contenu du système GIAS	59
Figure 1-22 : Etape de prédiction	60
Figure 1-23 : Attributs et leurs valeurs	61
Figure 1-24 : processus de génération des sujets.....	62
Figure 1-25 : différents niveaux sémantiques dans les ontologies	68
Figure 1-26 : pile du Web sémantique.....	69
Figure 1-27 : classification des ontologies dans le domaine de la modélisation des utilisateurs.....	75
Figure 1-28 : architecture de OntobUM	78
Figure 2-1 : diagramme de classes du processus de profilage.....	86
Figure 2-2 : dépendances entre \mathcal{D} , \mathcal{T} , et \mathcal{M}	88
Figure 2-3 : relations entre \mathcal{D} , \mathcal{M} , \mathcal{T} et \mathcal{P}	89
Figure 2-4 : dimension du style	91
Figure 2-5 : dimension sensorielle (\mathcal{D}_1).....	91
Figure 2-6 : dimension de progression (\mathcal{D}_2)	92
Figure 2-7 : dimension de réflexion (\mathcal{D}_3)	92
Figure 2-8 : dimension du raisonnement (\mathcal{D}_4)	92
Figure 2-9 : dimensions adaptées du modèle.....	95
Figure 2-10 : variation du coefficient TAL	98
Figure 2-11 : exemple d'objectifs d'apprentissage.....	101
Figure 2-12 : matériel pédagogique graphique - contraint	102
Figure 2-13 : matériel pédagogique graphique - libre	102
Figure 2-14 : matériel pédagogique textuel - contraint	103
Figure 2-15 : matériel pédagogique textuel - libre	103
Figure 2-16 : pré-test en début de séquence	104
Figure 2-17 : évaluation sommative	104

Figure 2-18 : Relation entre U et TAL	109
Figure 2-19 : Variation de la note du post-test par groupe.....	110
Figure 2-20 : Variation de la note du post-test par groupe après filtrage.....	111
Figure 2-21 : Variation de la note du post-test par groupe et niveaux de Bloom	112
Figure 2-22 : Variation du TAL par groupe	113
Figure 2-23 : Variation de l'effort d'un élève par groupe.....	114
Figure 3-1 : architecture du système proposé.....	123
Figure 3-2 : variation du niveau d'acquisition	125
Figure 3-3 : Inférence du prochain objectif.....	126
Figure 3-4 : modèle de connaissance	128
Figure 3-5 : approche top-down.....	129
Figure 3-6 : inférence des métadonnées M_i en fonction de $\{T_i, D_i\}$	134
Figure 3-7 : Inférence du style de l'élève.....	135
Figure 3-8 : dépendances entre M_i, P_i et D_i	139
Figure 3-9 : dépendances entre M_i, P_i et T_i	139
Figure 3-10 : schéma réseau bayésien.....	140
Figure 3-11 : structure du réseau bayésien.....	141
Figure 3-12 : réseau bayésien.....	146
Figure 3-13 : processus de validation du réseau bayésien (RB)	147
Figure 3-14 : pourcentage de style correctement estimé par le RB	148
Figure 3-15 : processus de validation du réseau bayésien (RB)	148
Figure 3-16 : pourcentage de M_i correctement estimées par le RB.....	149
Figure 3-17 : la transformation non linéaire des données offre une séparation linéaire des exemples dans un nouvel espace.....	151
Figure 3-18 : L'hyperplan séparateur optimal est celui qui maximise la marge dans l'espace de redescription	152
Figure 3-19 : changement d'espace pour transformer une séparation non linéaire en linéaire	152
Figure 3-20 : structure de l'estimateur de préférences	156
Figure 3-21 : pourcentage de style correctement estimé par SVM.....	158
Figure 3-22 : pourcentage de \mathcal{M} correctement estimé par SVM	159
Figure 4-1 : structure des modèles de contenus	165
Figure 4-2 : Approche conceptuelle du modèle de contenus	166
Figure 4-3 : exemple de décomposition du domaine	167
Figure 4-4 : exemple de décomposition pédagogique.....	169
Figure 4-5 : exemple d'objets d'apprentissage	170
Figure 4-6 : représentation UML de la décomposition structurelle du modèle de contenus	171
Figure 4-7 : vision parcelaire de la méta-ontologie du modèle de contenu	172
Figure 4-8 : représentation UML des dépendances du modèle de contenus.....	173
Figure 4-9 : cycle de vie du modèle de contenus	174
Figure 4-10 : exemples d'assets	179
Figure 4-11 : exemple de SCO.....	180
Figure 4-12 : Organisation des contenus.....	180
Figure 4-13 : métadonnées LOM	181
Figure 4-14 : les 9 catégories de LOM.....	182
Figure 4-15 : séquence et navigation (Seq. & Nav.) pour les activités	183
Figure 4-16 : Deux niveaux d'adaptation.....	184
Figure 4-17 : Structure du paquet SCORM.....	186

Figure 4-18 : le concept de <Tache>	187
Figure 4-19 : relation d'influence entre concepts <Tâche>.....	187
Figure 4-20 : exemple de concept <Tâche>	189
Figure 4-21 : exemple de description LOM d'un concept <Tâche>	190
Figure 4-22 : le concept <ObjectifGlobal>	191
Figure 4-23 : relation d'influence entre concepts <ObjectifGlobal>	191
Figure 4-24 : exemple de concept <ObjectifGlobal>	193
Figure 4-25 : exemple de description LOM de concept <ObjectifGlobal>.....	193
Figure 4-26 : le concept <ObjectifFinal>	194
Figure 4-27 : relation d'influence entre concepts <ObjectifFinal>	194
Figure 4-28 : exemple de concept <ObjectifFinal>.....	196
Figure 4-29 : exemple de description LOM de concept <ObjectifFinal>	197
Figure 4-30 : exemple de description SCORM du concept <Sequence> pour un objectif final.....	198
Figure 4-31 : le concept <Sequence>	198
Figure 4-32 : exemple de concept <Sequence>.....	200
Figure 4-33 : exemple de description LOM de concept <Sequence>	200
Figure 4-34 : exemple de description SCORM du concept <Sequence> pour un objectif final.....	201
Figure 4-35 : le concept <Activite>.....	202
Figure 4-36 : exemple de concept <Activite>	204
Figure 4-37 : exemple de description LOM de concept <Activite>.....	205
Figure 4-38 : le concept de <Ressource>	206
Figure 4-39 : exemple de concept <Ressource>.....	207
Figure 4-40 : exemple de description LOM de concept <Ressource>	208
Figure 4-41 : exemple de description SCORM de concept <Ressource>.....	208
Figure 5-1 : architecture d'un système adaptatif (\bar{D}_i est le style réel inconnu de l'élève).....	218
Figure 5-2 : Deux niveaux d'adaptation	219
Figure 5-3 : Structure de contenus adaptatifs	221
Figure 5-4 : algorithme du sélecteur de séquence	223
Figure 5-5 : algorithme du compositeur	224
Figure 5-6 : relations entre Acteurs, Tâches et Artefacts	279
Figure 5-7 : décomposition structurelle de la tâche « requirements ».....	279
Figure 5-8 : dépendance entre les tâches métiers	280
Figure 5-9 : décomposition structurelle des compétences métier.....	280
Figure 5-10 : dépendances des compétences métier.....	281
Figure 5-11 : décomposition structurelle des objectifs pédagogiques.....	282
Figure 5-12 : dépendances des objectifs pédagogiques.....	283
Figure 5-13 : maquette écran du verbe « Identify »	283
Figure 5-14: Results of Learning Process.....	323
Figure 5-15: Results of Content	324
Figure 5-16: Results of Exercises and Tests.....	326
Figure 5-17: Results of Navigation & Orientation	327

Liste des tableaux

Tableau 1-1 : Modèle de préférences de l'élève.....	40
Tableau 1-2 : correspondance entre éléments OWL et entités concept map.....	69
Tableau 2-1 : Question ILS par dimension du style.....	90
Tableau 2-2 : degré de confiance de la mesure du style.....	91
Tableau 2-3 : Facteurs expliqués par la première version du questionnaire ILS	93
Tableau 2-4 : variables pour l'étude.....	100
Tableau 2-5 : matériel pédagogique utilisé par groupe	100
Tableau 2-6 : Répartition des Styles dans la population étudiée.....	106
Tableau 2-7 : Facteurs expliqués par le questionnaire RILS.....	107
Tableau 2-8 : Variables observées lors de la séquence pédagogique.....	107
Tableau 2-9 : Variation des variables mesurées lors de l'expérimentation.....	108
Tableau 2-10 : Pourcentage de résultats éliminés par groupe	111
Tableau 3-1 : méthodes d'apprentissage	133
Tableau 3-2 : matrice de données après filtrage.....	143
Tableau 3-3 : Pourcentage de résultats éliminés par groupe	143
Tableau 4-1 : représentation du concept de <Tache>.....	188
Tableau 4-2 : représentation du concept <ObjectifGlobal>	192
Tableau 4-3 : représentation du concept <ObjectifFinal>.....	195
Tableau 4-4 : représentation du concept <Sequence>.....	199
Tableau 4-5 : représentation SCORM du concept <Sequence>.....	200
Tableau 4-6 : représentation du concept <Activite>	203
Tableau 4-7 : représentation du concept <Ressource>.....	206
Tableau 5-1 : dimensions d'adaptation sémantique d'une séquence pédagogique.....	220
Tableau 5-2 : correspondances adaptation / structure de contenus	221
Tableau 5-3 : 16 modèles de séquences pédagogiques.....	225
Tableau 5-4 : Précision et Rappel du système.....	226
Tableau 5-5 : type de ressources en fonction du type d'activités.....	284

Introduction générale

Automatiser l'enseignement a été la solution rêvée pour résoudre le problème d'éducation né de l'exode rural du début du vingtième siècle. Individualiser l'instruction à l'aide de machines a été la motivation initiale de beaucoup de travaux. Divers outils, incluant le phonographe, tourne-disques, radio, puis télévision ont été utilisés pour enseigner. Mais la première individualisation prendra la forme de l'*enseignement programmé*. Cette approche, née du courant psychologique béhavioriste, prône la décomposition d'une matière en fragments élémentaires administrés à l'élève d'une manière linéaire par questions et réponses. Commencant sous forme de programmes manuels, cette méthode a vite fait sentir le besoin de disposer de machines à enseigner. Pour répondre aux besoins des militaires et des industriels, la machine à enseigner de Crowder a été conçue. Elle utilise le principe de programmation à branchement pour distiller à l'élève des connaissances de plus en plus complexes en fonction de ses réponses. La difficulté d'utiliser ce genre de machine a poussé les chercheurs à considérer d'autres modes d'interaction entre la machine et l'élève.

Dès 1959, Gordon Pask propose le concept de machine qui s'adapte au style d'apprentissage de l'élève. Il est le premier à s'intéresser à l'ajustement de l'enseignement en fonction des réponses correctes et erronées de l'élève. L'enseignement programmé devient alors une discipline à part entière réunissant chercheurs, producteurs et formateurs avec des publications et des manuels. Intervint alors la naissance de l'*enseignement assisté par ordinateur* (EAO). Plusieurs systèmes d'enseignement et logiciels auteurs voient le jour dans plusieurs universités, notamment américaines. Des tutoriaux et des simulations avec des interactions graphiques se développent et se répandent. De nombreux laboratoires de recherche en EAO sont créés à travers le monde. Le développement de l'intelligence artificielle a apporté des améliorations aux méthodes et modes d'interaction des logiciels d'apprentissage. Cela a donné naissance aux *systèmes tuteurs intelligents*. Plusieurs de ces systèmes sont créés dans des domaines aussi variés que la géographie, les mathématiques, le dépannage de circuits électroniques, et la médecine.

Sous l'influence des travaux de Jean Piaget, le courant de l'apprentissage *constructiviste* s'est développé. La naissance des sciences cognitives, conjuguée aux efforts dans l'intelligence artificielle, apporte alors une rupture avec le béhaviorisme. Seymour Papert crée le langage *LOGO* spécialement conçu pour l'éducation. L'élève programme la machine pour explorer et construire sa connaissance au lieu de l'inverse. Ce langage entraîne de nombreuses recherches et développements, souvent concluants. Cependant, l'usage de LOGO reste dans un cadre d'enseignement classique, à cause de ces limitations et des besoins de formation des maîtres. L'apparition de la programmation objet avec Smalltalk est un autre fait marquant l'évolution des technologies éducatives. Il apportera une nouvelle approche de conception des contenus et environnements d'apprentissage.

La démocratisation de la micro-informatique dans les années 1980 permit d'accélérer le développement d'applications éducatives par ordinateur. Le multimédia est devenu une industrie dans laquelle le secteur du *CD-ROM* d'enseignement par ordinateur a pris un essor considérable. Le ralentissement de ce mouvement dans les années quatre vingt dix a coïncidé avec l'arrivée du *Web*. Ce média facile d'accès et d'utilisation a accéléré le développement de contenus d'apprentissage hypertextes. Publier son cours sur Internet est devenu le loisir de tout enseignant moderne ! Les entreprises et autres organismes publics y ont vu le moyen de délivrer « la connaissance qu'il faut au moment où il faut » à leurs employés. *L'enseignement à distance* par le réseau est devenu un domaine d'intérêt pour les universités aussi. Des systèmes de gestion de l'apprentissage (*learning management systèmes*) se sont répandus. Ils permettent de stocker, délivrer, et administrer des cours ou programmes entiers par le *Web*. Afin de valoriser économiquement le secteur du *e-learning*, des normes et standards (e.g. SCORM, LOM, IMS LD) se sont créés. Ils visent pour leur majorité la normalisation de la description des contenus pédagogiques en vue de faciliter leur interopérabilité et réutilisation.

La convergence entre les tuteurs intelligents et l'enseignement interactif par le web a engendré une nouvelle forme de systèmes dits, *environnements intelligents pour l'apprentissage humain (EIAH)*. C'est la naissance du nouveau paradigme de séparation du contenu et du contrôle. L'évolution de la technologie a ainsi permis de reconsidérer la personnalisation de l'instruction à l'élève. Les EIAH gèrent des *objets d'apprentissages* granulaires et réutilisables. Ils les exploitent en fonction de *modèles de connaissances et de préférences* de l'élève afin de lui fournir un enseignement adapté. Des efforts considérables de recherche ont été déployés pour créer des méthodes, modèles et outils pour cette approche. *L'adaptation* d'un ensemble de contenus pédagogiques a été vue comme la personnalisation de la navigation et de la présentation aux besoins de l'élève. Les critères à utiliser pour cette activité et la manière de les recueillir reste de l'ordre de la recherche, malgré un début de normalisation comme IMS LIP.

Problématique

Dans ce contexte, cette thèse est une contribution aux efforts de recherche sur l'adaptation de l'enseignement à l'élève. La problématique posée est comment proposer à l'élève une expérience pédagogique qui soit appropriée non seulement à ses connaissances mais aussi à sa manière d'apprendre. Nous avons reformulé ce problème en fixant les objectifs scientifiques suivants :

- établir un modèle formel prédictif des caractéristiques psychologiques de l'élève définissant sa manière d'apprendre
- concevoir un modèle de contenus pédagogiques incluant les besoins d'adaptation aux caractéristiques de l'élève
- proposer une stratégie d'adaptation permettant de mettre en correspondance le modèle de l'élève et le modèle de contenus pour recommander les activités pédagogiques qui maximisent les performances.

Pour atteindre le premier objectif, nous avons opté pour la caractérisation du style d'apprentissage de l'élève. L'index de learning style de Felder a été choisi pour sa portée générale. Nous avons mené une étude empirique pour établir un modèle de dépendance statistique pour prédire l'effet du style sur le comportement de l'élève. L'analyse des résultats nous a permis de valider trois nouvelles mesures du comportement de l'élève : l'énergie d'apprentissage, le degré d'attention, et le niveau d'adaptation de la séquence à son style. Les résultats de cette étude nous ont permis de proposer une approche probabiliste générique pour modéliser le style de l'élève. Nous avons implanté ce modèle à l'aide de deux formalismes mathématiques : les réseaux bayésiens et les machines à vecteurs de support.

Le travail sur le deuxième objectif a consisté à créer une méta-ontologie décrivant le contenu pédagogique. Elle se compose de trois modèles conceptuels : métier, pédagogique, et de ressources. Nous avons développé une méthode d'ingénierie pédagogique en cinq phases pour instancier cette méta-ontologie pour créer un cours. Le livrable final de la méthode est un ensemble d'objets d'apprentissage indexés par une ontologie. Nous avons montré comment un tel ensemble peut être structuré et décrit à l'aide de SCORM CAM et LOM. Cela a nécessité l'adaptation de SCORM et l'amélioration de LOM pour répondre aux besoins de l'adaptation.

Le dernier objectif a été traité en formulant l'adaptation comme un problème de recherche de la séquence optimale parmi celles décrites dans le contenu. L'optimalité est mesurée à l'aide d'une distance de similarité sémantique des séquences. Quatre dimensions, constituées chacune de plusieurs métadonnées sont à la base de notre modèle d'adaptation.

L'intégration de ces travaux s'est faite à travers la conception d'un système d'enseignement adaptatif (Figure 1.1). L'architecture de ce système a été la trame qui a guidé le développement des différentes parties. Deux parties de ce système ont été mises en œuvre dans deux projets européens. En effet, la méthode et le modèle de contenus ont été à la base du projet européen UP2UML. La partie modélisation de l'élève a été réutilisée dans le projet KPLAB dans le cadre de l'activité de profilage des utilisateurs d'un système de travail coopératif.

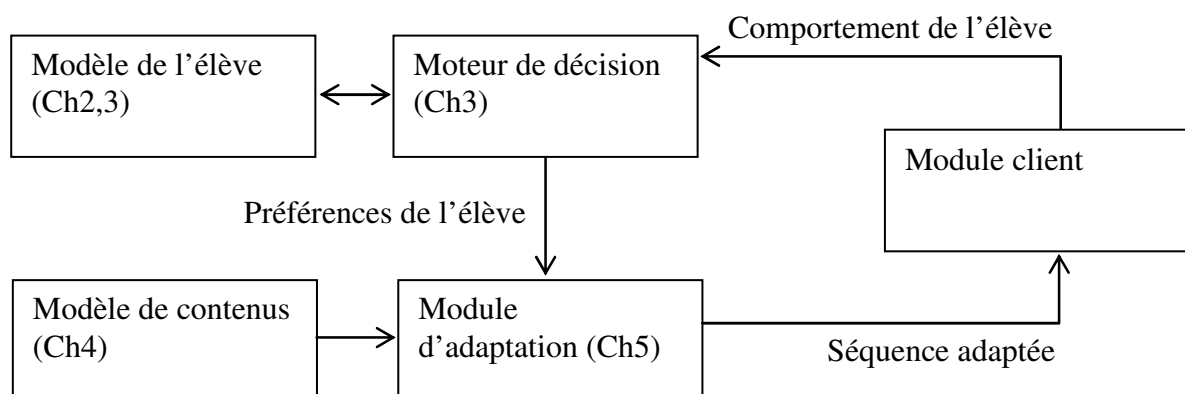


Figure 1-1.1 Architecture de notre système adaptatif

Contributions

Les travaux de cette thèse ont abouti aux contributions scientifiques suivantes :

- une étude empirique montrant la dépendance entre le style d'apprentissage, l'activité pédagogique et le comportement de l'élève
- trois mesures originales pour caractériser le comportement de l'élève : l'énergie d'apprentissage, le degré d'attention, et le niveau d'adaptation de la séquence
- un réseau bayésien pour estimer par renforcement le style d'apprentissage et recommander les métadonnées de la séquence pédagogique adaptée
- un deuxième classificateur probabiliste fondé sur le formalisme de machine à vecteurs de support pour estimer le style et recommander la séquence adaptée
- un modèle de contenus capable de capturer la variation des méthodes et approches pédagogiques, composé de trois parties : les connaissances métier, les aspects pédagogiques, et les caractéristiques structurelles et de forme des ressources
- une méthode d'ingénierie permettant la création d'un cours selon notre modèle de contenus, produisant une ontologie et des objets d'apprentissage granulaires structurés selon la norme SCORM
- une stratégie d'adaptation de l'enseignement au style de l'élève qui se fonde sur le niveau de connaissance, le type de navigation, la nature des activités, le genre de média, et le type de ressource
- une mesure de similarité sémantique pour évaluer la distance entre deux séquences pédagogiques, permettant de sélectionner la séquence optimale, au sens de notre stratégie d'adaptation, pour un style donné

Plan de la thèse

En plus de cette introduction générale, la thèse est structurée en cinq chapitres. Nous conseillons une lecture linéaire des chapitres. Une conclusion générale revient sur les limitations de nos travaux et les perspectives que nous en déduisons pour nos futures recherches. Un aperçu du contenu de chaque chapitre est donné ci-après.

Chapitre 1 : Etat de l'art

Ce chapitre présente une revue de la littérature des modèles et techniques liés à la conception des systèmes d'enseignement adaptatifs. Nous y couvrirons les modèles de l'élève, les modèles de contenus et les stratégies d'adaptation. Nous finirons en montrant comment les nouveaux systèmes d'enseignement tentent d'utiliser les ontologies pour la conception de leurs modèles de connaissances. Les méthodes numériques telles que les réseaux bayésien et les machines à vecteurs de support utilisées pour implémenter le modèle de l'élève, seront présentées dans le chapitre 3.

Chapitre 2 : Activité de profilage, étude statistique du style d'apprentissage

Dans ce chapitre, nous décrivons un méta-modèle conceptuel générique qui présente le processus de profilage comme étant l'activité d'estimer des variables non observables à partir d'indicateurs observables. A la base de ce méta-modèle, nous créons un modèle de dépendance permettant de profiler le style d'apprentissage d'un élève dans le cadre de l'enseignement à distance. Nous y décrivons l'étude empirique que nous avons menée pour caractériser les indicateurs mesurables du comportement de l'élève. Trois mesures originales du comportement y sont proposées et validées : l'énergie d'apprentissage, le degré d'attention, et le niveau d'adaptation de la séquence. Le chapitre montre statistiquement que l'adaptation de la séquence pédagogique au style d'apprentissage de l'élève influence son comportement et ses performances.

Chapitre 3 : Apprentissage supervisé du modèle de l'élève

A la base des résultats obtenus au chapitre précédent, nous présentons un modèle de l'élève qui sépare les connaissances et les caractéristiques psychologiques définissant les préférences d'apprentissage. La partie modèle de connaissances est utilisée par un séquenceur d'objectifs pédagogiques pour déterminer l'activité pédagogique suivante. La partie préférences est l'aspect central de ce chapitre et de cette thèse. Elle est le fondement de notre approche d'adaptation de la séquence pédagogique, de la nature de ses activités, du mode de navigation dans la séquence et du type de matériel présenté pour tenter de maximiser la performance de l'élève. Nous l'exploitons de deux manières complémentaires. D'un côté, en observant les comportements que l'élève exhibe face aux séquences qu'il a déjà suivies, le système apprend par renforcement son style d'apprentissage. De l'autre côté, étant donné les performances obtenues et la croyance actuel sur le style, le système estime probabilistiquement les métadonnées décrivant la séquence préférée à présenter à l'élève au prochain objectif. La séquence idéale n'existant pas, ce sera au module d'adaptation (Chapitre 5) de sélectionner la séquence optimale. Le composant de notre système qui se charge de ces deux tâches se nomme estimateur de style d'apprentissage. Il forme avec le séquenceur mentionné ci-dessus le moteur de décision de notre système adaptatif. Nous avons implémenté deux versions différentes de l'estimateur de style, toutes les deux probabilistes. La première s'appuie sur la technique des réseaux bayésiens. La seconde se fonde sur le formalisme de machines à vecteurs de support. Nous montrerons que sous certaines conditions, la seconde aboutit à de meilleurs résultats.

Chapitre 4 : Contenu adaptatif : modèle, méthode et implémentation

Ce chapitre traite de la conception, la création et l'implémentation du modèle de contenus de notre système adaptatif. Nous y proposons l'utilisation d'une méta-ontologie composée de trois modèles conceptuels : métier, pédagogique et de ressources. Nous proposons une méthode d'ingénierie pédagogique pour créer les contenus en instanciant la méta-ontologie. L'originalité de cette méthode se situe dans le fait que les phases enrichissent successivement l'ontologie décrivant le sujet d'étude. Le résultat est un ensemble d'objets d'apprentissage indexés par une ontologie. L'ontologie est la base du modèle de connaissances de l'élève (chapitre 2). En effet, les connaissances de l'élève sont représentées par un recouvrement probabiliste de l'ontologie.

Nous proposons SCORM CAM comme moyen technique pour structurer et décrire l'ontologie et les ressources. Nous améliorons les métadonnées LOM pour implémenter les informations nécessaires à l'adaptation. Nous montrons finalement la mise en œuvre de cette approche dans le projet européen UP2UML.

Chapitre 5 : Stratégie d'adaptation

Ce chapitre se focalise sur la stratégie d'adaptation. Nous y formulons le problème de l'adaptation comme un problème de recherche de séquence pertinente dans une base. Nous proposons un modèle d'adaptation de la séquence pédagogique en fonction de la navigation, l'activité, le média et le type de ressource. Nous établissons l'architecture du module d'adaptation qui implémente ce modèle. Ce module prend en entrée, sous forme de requête, les caractéristiques de la séquence idéale qui sort du moteur de décision (Chapitre 3). Il recherche la séquence optimale parmi celle disponibles dans le modèle de contenus. Il en compose le contenu pédagogique. Pour mesurer la distance entre deux séquences, nous proposons une nouvelle mesure de similarité sémantique qui exploite l'indexation des contenus par des métadonnées. Enfin, nous présentons la simulation que nous avons effectuée pour valider l'approche. Des mesures quantitatives de performance sont utilisées.

CHAPITRE

1. Etat de l'art

Résumé :

Les systèmes d'enseignement adaptatifs visent à mettre en adéquation les séquences pédagogiques et les profils des élèves. Ils utilisent, à cette fin, un ou plusieurs modèles pour représenter la connaissance. On rencontre principalement des modèles d'élèves et de contenus.

Dans ce chapitre, nous présentons une revue de la littérature des modèles et techniques d'adaptation. Nous finirons en montrant comment les nouveaux systèmes d'enseignement tentent d'utiliser les ontologies pour la conception de leurs modèles de connaissances.

1.1. Introduction

Les dernières années ont connu une explosion des systèmes d'aide à l'apprentissage humain qui exploitent les technologies du Web et du multimédia. Dans ce contexte, les techniques de l'intelligence artificielle sont utilisées pour la modélisation des élèves [Brusilovsky, 1999] et l'adaptation des séquences pédagogiques. L'objectif est de fournir à l'élève, à tout moment, l'information pertinente présentée d'une manière appropriée [Fischer, 2001]. Cependant une question s'impose : Quelle est la bonne information et qu'est ce qu'une présentation adaptée ?

Pour tenter d'y répondre, plusieurs modèles et prototypes de systèmes adaptatifs et intelligents d'enseignement par le Web¹ ont vu le jour. Ils sont le résultat du croisement des recherches sur les systèmes tuteurs intelligents² [Murray, 1999 ; Self, 1992 ; Sleeman, 1982 ; Wenger, 1997] et les systèmes hypermédia adaptatifs³ [Brusilovsky et al, 2003a].

Ils exploitent dans leur architecture divers modèles de connaissances [Zhang et al, 2003] pour offrir un apprentissage et des techniques d'enseignement personnalisés. Dans [Triantafillou et al, 2003], les auteurs montrent que l'architecture générale d'un AIWBES se résume à : un modèle de l'élève, un modèle de contenus et une stratégie d'adaptation [Piombo et al., 2003] (Figure 1-1).

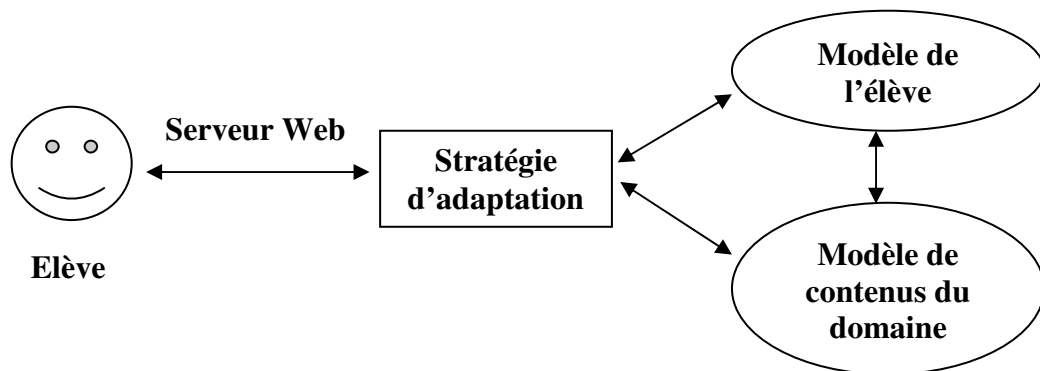


Figure 1-1 : Architecture générale des AIWBES

Dans [Alesso et al, 2004 ; Brooke et al, 2005], les auteurs ajoutent qu'à l'ère du Web2.0, l'interopérabilité des données stockées dans les différents modèles de connaissances est nécessaire. Cela permet à plusieurs systèmes de partager les informations recueillies sur un même élève. Chaque système peut alors faire des inférences sur les données partagées pour amorcer un raisonnement qui mène à la personnalisation des contenus et des services offerts [Mehta et al, 2005]. L'utilisation du Web sémantique [Alani et al, 2003] apporte une solution à ce besoin, par l'utilisation d'ontologies pour décrire le contenu des différents modèles de connaissances manipulés par les AIWBES.

¹ Adaptive and Intelligent Web-Based Educational Systems - AIWBES

² Intelligent Tutoring System - ITS

³ Adaptive Hypermedia Systems - AHS

Dans ce chapitre, nous présentons les différents modèles de connaissances exploités dans les AIWBES et leurs évolutions grâce à l'utilisation des technologies du Web2.0, notamment les ontologies.

1.2. Modélisation de l'élève

L'estimation des caractéristiques d'un élève est essentielle pour les applications qui exigent une adaptation, telles que la recherche documentaire sur le Web [Brusilovsky et al, 2004b], la visite de musées virtuels [Sparacino, 2003], l'accès aux pages Web pour des utilisateurs atteints d'handicaps [Alexandraki, 2004; Stephanidis, 2001], ou le commerce électronique [Pierrakos et al, 2003]. Une synthèse de solutions commerciales pour la modélisation des utilisateurs du web peut être trouvée dans [Fink et al, 2000].

Dans le contexte de l'apprentissage humain, le principe fondamental mis en œuvre consiste à estimer les besoins de l'élève pour adapter les contenus pédagogiques [Fisher, 2001].

Afin de faciliter leur comparaison, nous nous proposons de présenter les modèles de l'élève suivant trois critères :

- **Le type du modèle.** On distingue quatre types : expertise-partielle, différentiel, perturbation et social ;
- **Les données du modèle.** Deux catégories de données sont utilisées : le niveau de connaissances de l'élève ou les préférences de l'élève ;
- **L'acquisition du modèle.** Elle peut être statique, dynamique ou mixte.

1.2.1. Type du modèle

Dans la littérature quatre types de modèles sont utilisés pour représenter les caractéristiques d'un élève. Trois modèles se fondent sur la comparaison de l'élève avec un expert du domaine étudié (Expertise-partielle, Différentiel et perturbation). Le quatrième opte pour une classification des élèves (Social).

Le modèle expertise-partielle⁴

Le modèle de type expertise-partielle est le plus connu et le plus utilisé [Kass, 1989]. Il définit la connaissance de l'élève comme étant un sous ensemble de la connaissance d'un expert du domaine. On considère que l'élève aura acquis le niveau de connaissances souhaité lorsqu'il y aura recouvrement (Overlay) parfait des connaissances de l'expert par les connaissances de l'élève.

Dans ce modèle, il est nécessaire de pouvoir décomposer le domaine de connaissances étudié en différents éléments qui seront les points de mesure à contrôler. Cette modélisation est relativement générique, car elle peut représenter un ensemble de règles, de faits, de concepts ou autres éléments à étudier.

La Figure 1-2 montre une représentation d'un modèle de type expertise-partielle pour un domaine à étudier décomposé en 5 concepts. Par exemple, si le domaine est le « développement objet », *l'héritage* et le *polymorphisme* sont des concepts objets. L'expert

⁴ Overlay

est représenté avec un niveau de connaissances à 1, exprimant une maîtrise totale de chaque concept. L'élève modélisé montre quelques défaillances sur les concepts 1 et 3.

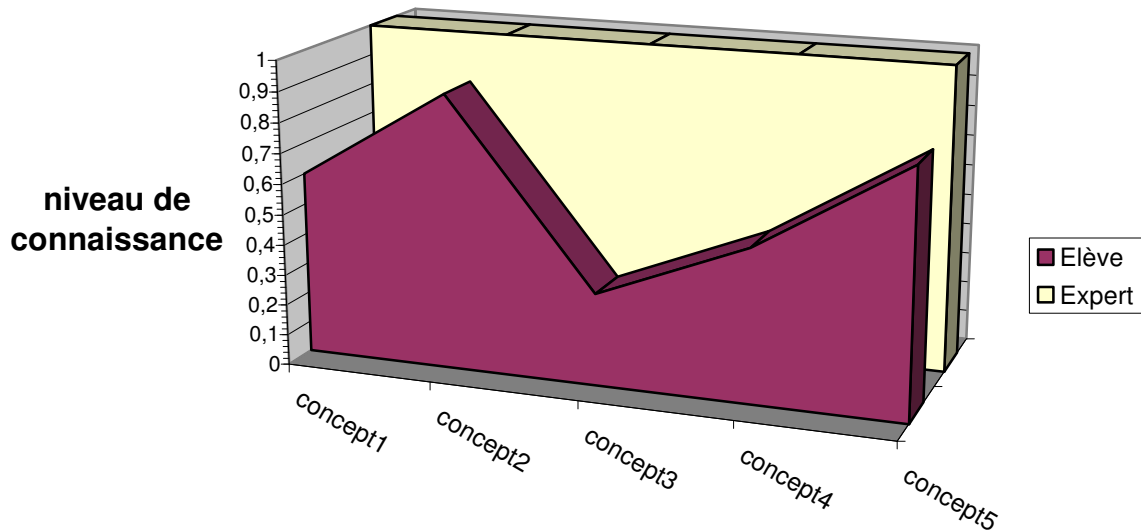


Figure 1-2 : Modèle expertise-partielle

Dans [Brusilovsky et al, 2003c], les auteurs montrent que ce type de modèle peut être utilisé dans les AIWBES en permettant la génération automatique d'un cours individualisé qui tient compte des objectifs et du niveau de connaissances de l'élève. Il sert à adapter dynamiquement le cours en fonction des succès ou échecs de l'élève dans la séquence pédagogique. Les auteurs développent le modèle dit *expertise-partielle numérique* dans lequel le niveau de connaissances de chaque concept est représenté par un nombre borné entre une valeur minimale (débutant) et une valeur maximale (expert). Un seuil de réussite est fixé. En dessous de celui-ci, le système continue à proposer des séquences pédagogiques portant sur le même concept.

La valeur associée à chaque élément du modèle peut être de différents types et avoir différentes significations [Mayo, 2001]. On trouve des systèmes dans lesquels le niveau de connaissances de chaque élément est représenté par une variable modale. Ainsi, une variable binaire représente pour chaque élément deux états : non acquis et acquis [De Bra, 1996]. La majorité des systèmes utilisent un modèle de type *expertise-partielle pondérée*. Dans ce cas, les variables possèdent plus de deux modalités. Il est alors possible de distinguer plusieurs niveaux de connaissances en utilisant une valeur qualitative : Bon, Moyen, Mauvais [Brusilovsky et al, 1998a; Grigoriadou et al, 2001].

D'autres systèmes considèrent que chaque élément est représenté par une variable continue bornée [Brusilovsky, et al, 1998b; De Bra, et al, 2001] offrant une mesure plus fine du niveau de connaissances. La valeur numérique associée à chaque élément correspond alors soit à la fréquence avec laquelle l'élément a été utilisé correctement ou incorrectement par l'élève [Kay, 2000b], soit à une probabilité comportant la notion d'incertitude dans la mesure [Millan, 2000]. Ce modèle de type *expertise-partielle probabiliste* est alors constitué d'un ensemble de variables, chacune représentant un élément du modèle dont la valeur est le reflet probable du niveau de connaissances acquis par l'élève pour cet élément [Henze et al, 1999; Specht et al, 2001]. La valeur de ces

variables peut être mise à jour en utilisant le théorème de Bayes pour calculer à partir d'observations, les probabilités à posteriori à partir de probabilités à priori.

Des systèmes plus récents utilisent plusieurs modèles de type expertise-partielle chacun spécialisé dans un type d'information à maintenir (connaissances, préférences, navigation, etc.) [Frasincar et al, 2004]. Par exemple dans [Castillo et al, 2003], les auteurs utilisent un premier modèle qui enregistre les croyances sur le niveau de connaissances de l'élève et un second modèle qui maintient les informations sur les interactions avec le système.

Ce type de modèle est très contraignant car il oblige l'élève à atteindre un seuil minimal de connaissance sur l'ensemble des concepts du domaine considéré. De plus, que faire si l'élève n'arrive pas à dépasser le seuil fixé ? L'échec pourrait être dû au style d'activités pédagogiques proposées à l'élève. Au lieu de lire des définitions, ne faudrait-il pas lui proposer d'étudier des exemples ? Dans ce cas, comment prendre cette décision pédagogique ? Dans cette thèse, nous proposons une approche qui permet de répondre à ces questions.

Le modèle différentiel

Le modèle différentiel est une variante du modèle expertise-partielle [Holt et al, 1994]. Il peut être considéré comme un modèle expertise-partielle sur un niveau de connaissances demandé qui est lui même un recouvrement du niveau de connaissances d'un expert du domaine. La Figure 1-3 montre que seuls les concepts 3 et 4 sont en dessous du niveau demandé. La séquence pédagogique qui suivrait devra donc se focaliser sur ces deux concepts.

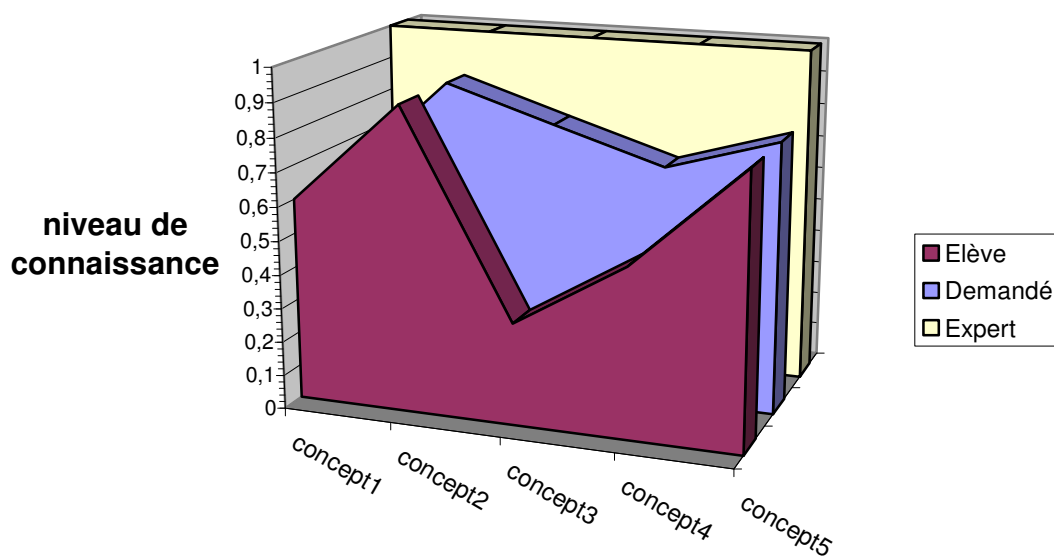


Figure 1-3 : Modèle différentiel

De ce fait, on peut dire que le modèle de type différentiel est plus flexible que le modèle de type expertise-partielle car il permet de focaliser les inférences sur le niveau de connaissances de l'élève uniquement pour les concepts visés. Ainsi, la séquence pédagogique qui lui serait proposée ne tiendrait compte que des concepts réellement nécessaires pour un certain objectif d'apprentissage [Burton et al, 1978]. La difficulté est

alors de définir quels sont les objectifs d'apprentissage significatifs pour l'élève dans le domaine d'apprentissage considéré. De plus, ce type de modèle n'apporte toujours pas de solution quant à l'analyse d'une erreur persistante de l'élève sur l'acquisition d'un concept particulier.

Le modèle perturbation

La considération des connaissances de l'élève comme un sous-ensemble de celles de l'expert est la principale limitation des modèles expertise-partielle et différentiel. En effet, ceci conduit à ignorer les mauvaises connaissances préalables et les erreurs de raisonnement lors de l'apprentissage. Le modèle perturbation tente de résoudre ces problèmes en proposant de représenter la connaissance de l'élève comme un sous-ensemble de la connaissance de l'expert complété avec un modèle des erreurs (de connaissances, de processus, etc.) connues a priori ou découvertes en cours de séquence par l'analyse de l'activité. Ainsi, le système sera capable de mieux réagir pour combler les lacunes. La Figure 1-4 donne une vue graphique du principe du modèle perturbation.

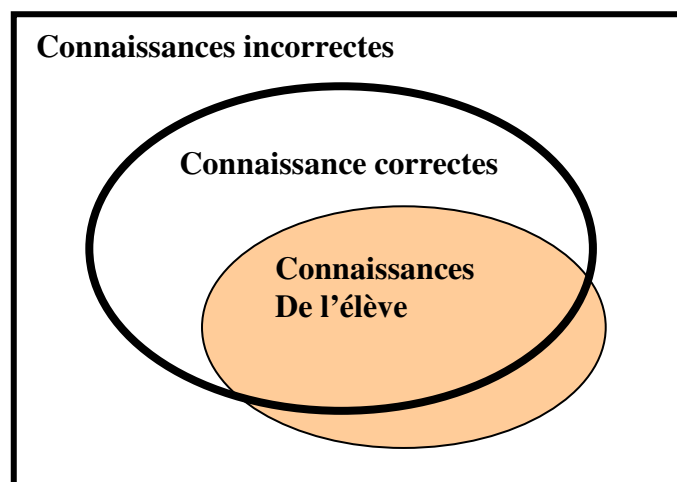


Figure 1-4 : Modèle perturbation

La Figure 1-5 présente un exemple d'erreur de raisonnement qu'il est possible de stocker dans un modèle de type perturbation [Burton, 1982].

$$\begin{array}{r} 307 \\ - 135 \\ \hline 232 \end{array}$$

Figure 1-5 : Réponse incorrecte dans un problème de soustraction de deux nombres

L'erreur peut être expliquée en observant que la mauvaise règle appliquée est : $0 \cdot n = n$. L'élève pense que lorsque le chiffre du haut de la colonne est 0 alors le résultat est égal au chiffre du bas de la colonne.

Le principal problème avec ce modèle est qu'en pratique il est extrêmement difficile, voir impossible de décomposer en éléments disjoints l'ensemble des erreurs qu'un élève peut commettre dans l'étude d'un domaine particulier, le rendant de ce fait peu exploitable [Virvou et al, 2000]. Dans ce contexte, il sera très difficile de l'utiliser pour corriger la cause réelle de l'erreur en proposant à l'élève une étude complémentaire du sujet qui fait défaut.

Le modèle social

Le modèle social consiste à classer un élève parmi un ensemble de stéréotypes prédéfinis. Le principe est d'utiliser ces stéréotypes comme guide pour le choix d'une séquence pédagogique. Les stéréotypes sont définis grâce à une analyse des réponses qu'un échantillon d'élèves donne à un questionnaire orienté. Par exemple, on peut définir une classification basée sur la personnalité des élèves en utilisant un test de personnalité psychologique [Milne et al, 1996]. Tout nouvel élève répondrait au questionnaire initial avant de débiter la séquence. L'analyse de ses réponses permettrait de calculer la probabilité qu'il appartienne à chaque stéréotype. Il est alors possible de l'associer au stéréotype le plus probable (Figure 1-6).

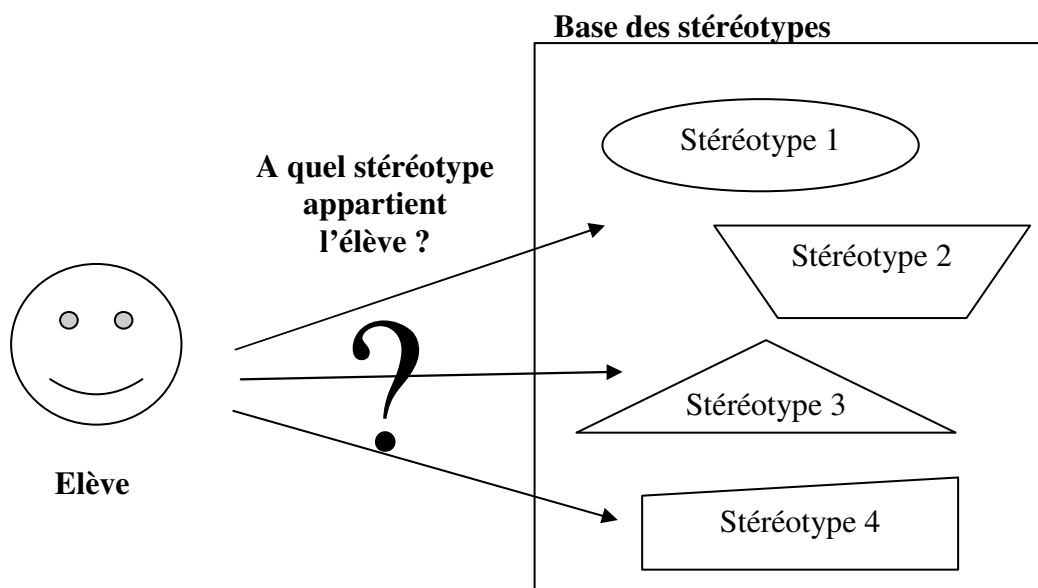


Figure 1-6 : Modèle social

Il existe deux sortes de modèles sociaux :

- Le premier type est dit « fixe ». Lorsqu'un élève a été associé à un stéréotype, il est alors impossible de réviser ce choix pour l'affecter à un autre stéréotype [Winter et al, 1999]. Cela est problématique si par exemple l'élève a prétendu dans le questionnaire avoir une certaine compétence, et qu'il s'avère au cours de la séquence qu'il n'en est rien. Il faudrait alors pouvoir lui réaffecter un nouveau stéréotype.

- Le second type est dit « par défaut ». Il consiste à considérer que le choix du stéréotype initial est simplement fait pour avoir des paramètres initiaux. Ensuite par des observations faites au regard des activités menées, on affine ces paramètres afin de les personnaliser [Kay, 2000a]. On est alors en mesure d'estimer à nouveau le stéréotype qui devient le plus probable.

La difficulté liée à ce genre de modèle est la complexité du nombre de cas différents à prendre en considération dans le cadre d'un système d'enseignement intelligent. Cela rend pratiquement impossible la création des stéréotypes associés qui doivent être tous disjoints. L'intersection de deux stéréotypes ne permettrait pas, pour l'élève qui s'y trouve, d'être associé à l'un des stéréotypes plutôt qu'à l'autre. Ensuite, à nouveau dans le cas où l'élève a du mal à suivre les activités pédagogiques qui lui sont proposées, comment identifier la cause avec ce modèle et quelle solution apporter ?

1.2.2. Données du modèle : connaissances et préférences

Les AIWBES, utilisent le modèle de l'élève pour concevoir dynamiquement la présentation et l'organisation du contenu pédagogique lors d'une séquence d'enseignement.

Dans [Cristea et al, 2004c], les auteurs précisent que concevoir un modèle de l'élève consiste à répondre en particulier à la question suivante :

- Quels sont les attributs qui devraient être « capturés » par le modèle de l'élève?

Dans [Grandbastien et Labat, 2006], les auteurs ajoutent que la conception d'un modèle de l'élève soulève deux problèmes : d'une part le choix des caractéristiques individuelles (cognitives, comportementales) et de leur mode de représentation informatique, d'autre part la mise en place de mécanismes de mise à jour de ce modèle à partir des événements logiciels (ou observables) recueillis durant l'interaction.

Pour répondre à ces questions, nous avons étudié quelques systèmes adaptatifs d'enseignement dans lesquels [Trella et al, 2002] il est possible de scinder le modèle de l'élève (Figure 1-7) en modèle de connaissances et de préférences (attitude).

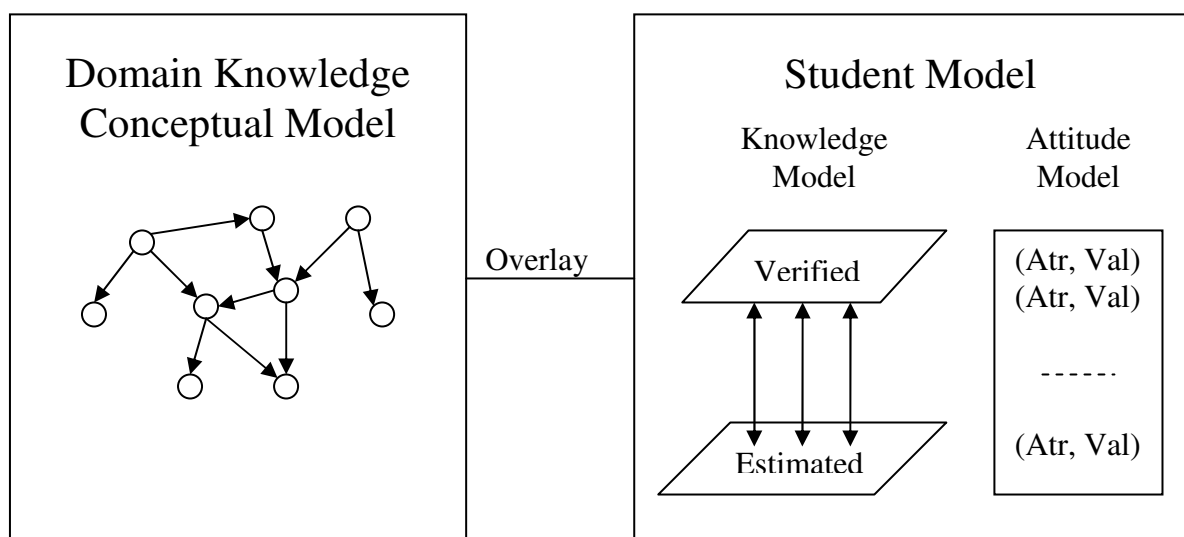


Figure 1-7 : Modèle de connaissances et modèle d'attitudes (préférences) de l'élève

Le principe de base est de considérer que le modèle de l'élève est de type expertise-partielle sur le modèle conceptuel de contenus. Ce dernier organise le domaine à étudier, généralement sous la forme d'un graphe dont les nœuds représentent les concepts et les arcs les notions de dépendances ou de composition entre les concepts.

Le modèle de connaissances de l'élève contient des informations sur le niveau de connaissances de l'élève au regard de chaque concept considéré dans le modèle conceptuel de contenus. Dans [Trella et al, 2002], les auteurs utilisent un modèle à deux couches (Figure 1-7) : la couche « verified » contient le niveau de connaissances des concepts qui ont été évalués d'une manière sommative, alors que la couche « estimated » contient une estimation du niveau de connaissances par la collecte et le traitement d'informations diverses reflétant le comportement de l'élève pendant la session d'apprentissage (ex : pages visités, temps d'accès aux pages). L'approche des auteurs est de considérer que les concepts, dont les pages n'ont pas été visités, ne peuvent être connus. Ce principe a déjà été utilisé par d'autres auteurs comme Brusilovsky dans le système ELM-ART-II [Weber et al, 1997] pour définir un modèle de connaissances à plus de deux niveaux (concepts : visited, evaluated, inferred, etc). Mais dans ce cas, on ne tient pas compte des connaissances que l'élève peut avoir acquises au préalable. La séquence d'enseignement qui lui sera proposée risque donc de ne pas être adaptée.

Le modèle de préférences (Figure 1-7) contient des informations qui décrivent des caractéristiques autres que le niveau de connaissances. On y trouve, des données comme l'âge, la capacité linguistique mais aussi des informations sur le matériel qu'il utilise comme le débit de sa connexion internet. Dans ce cas, les données sont souvent stockées sous la forme de couples *Attribut-valeur*. Dans [Schiaffino et al, 2000], les auteurs précisent que ces données sont à considérer comme l'ensemble des préférences d'un élève. Elles sont obtenues par un agent qui observe le comportement de l'élève dans son activité. Par exemple, lorsqu'un élève interroge une base documentaire, il est possible de capturer les attributs utilisés pour la recherche ainsi que des informations temporelles (ex : jour, mois, heure). En mémorisant chaque requête faite par l'élève sur une période donnée, l'utilisation de techniques de raisonnement de type CBR⁵ offre un moyen de prédire l'information à proposer à l'élève. On peut ainsi détecter, par exemple, qu'un utilisateur interroge tous les matins à 8h00 le site de météo France pour connaître le temps qu'il fait dans la région Midi-Pyrénées. Lors des futures connexions de l'utilisateur, le système sera capable de lui présenter cette information sans qu'il en fasse la demande.

Modélisation des connaissances

Ce type de modélisation a connu un essor important avec l'apparition des ITS. Le modèle des connaissances d'un élève est constitué d'attributs associés à chaque concept du domaine étudié. La valeur maintenue par chaque attribut est mise à jour en fonction des activités de l'élève dans la séquence. Ces attributs sont de différents types et apportent différentes informations sur l'élève. De nombreux systèmes d'enseignement adaptatifs utilisent ce type de modèle.

AHA! est un système hypermédia adaptatif orienté Web développé à Eindhoven University of Technology [De Bra et al, 2003a]. Le système utilise un modèle de l'élève de type expertise-partielle, dans lequel le niveau de connaissances pour un concept donné est représenté par un seul attribut appelé « knowledge ». La valeur de cet attribut est mise à jour de différentes façons. La première et la plus simple consiste simplement à enregistrer que l'élève a parcouru le contenu des pages associées au concept visé. AHA! offre une

⁵ CBR : Case-Based Reasoning

deuxième solution qui consiste à créer un questionnaire en ligne qui sera proposé à l'élève afin d'évaluer son niveau de connaissances. Cette solution permet une évaluation plus fine que simplement mémoriser si l'élève a visité ou non les pages reliées au concept.

Le système hypermédia adaptatif orienté Web ALEA⁶ est utilisé dans le cours de programmation logique et fonctionnelle à Slovak University of Technology [Kostelnik et al, 2003] depuis l'année universitaire 2002/2003. Pour chaque concept visité ALEA maintient le niveau de connaissances de l'élève qui lui est indiqué manuellement par ce dernier. L'élève peut par exemple par ce moyen indiquer la valeur « understood » qui précise que d'après lui, le concept abordé a été compris. Dans ce cas, le modèle d'élève utilisé est de type expertise-partielle pondéré. Cette approche est utilisée dans le système ISIS-Tutor [Brusilovsky et al, 1994] où à chaque concept est associé un état : not-ready-to-be-learned, ready-to-be-learned, in-work et learned. L'approche est aussi utilisée dans le système ELM-ART [Weber et al, 2001a] où les valeurs admises sont : visited, learned, inferred et known.

Les systèmes AHA ! et ALEA utilisent un modèle de connaissance unique qui ne permet pas une analyse des erreurs de l'élève. Il est donc difficile de connaître les causes des échecs et de prévoir ainsi les solutions pour y remédier.

Hera [Frasincar et al, 2004] est une méthodologie pour concevoir des systèmes Web d'information basés sur l'exploitation de plusieurs modèles de type expertise-partielle. Chaque modèle capture différents aspects du système. Le framework théorique LAOS⁷ [Cristea, 2004d], créé dans le but de proposer un cadre flexible de développement de systèmes auteurs hypermédiés adaptatifs pour le Web, est basé sur ce principe. Il a été utilisé pour concevoir le système MOT⁸ [Cristea et al, 2003b]. L'architecture de LAOS divise les éléments qui composent le système hypermédia adaptatif en couches. Cette décomposition est guidée par des aspects sémantiques et fonctionnels. La principale idée est de regrouper les éléments en considérant entre autre comme critères leurs contextes d'utilisation et leurs possibles réutilisations. Par exemple, le modèle conceptuel de contenus a été scindé selon ce principe en un modèle du domaine (DM) et un modèle d'objectifs et de contraintes (GM). Pour comprendre ce découpage, il est intéressant de prendre pour métaphore une présentation faite sur un sujet pour lequel la matière (le contenu) est extraite d'une encyclopédie. Le DM représente l'encyclopédie, alors que le GM représente le contenu du support de la présentation. A partir d'une encyclopédie, il est possible de construire plusieurs présentations dépendant de l'objectif pédagogique. La présentation ne contient pas l'encyclopédie complète mais simplement une partie qui est significative pour l'objectif visé.

Le modèle de connaissances de LAOS est du type différentiel. Il se compose des trois premières couches (Figure 1-9) du modèle global du framework :

- **Le modèle conceptuel de contenu (DM)** : il est composé d'un graphe de concepts (concept maps) auxquels sont associés des attributs. La signification du terme « concept » varie en fait en fonction des attributs qui leurs sont reliés. Les auteurs indiquent que grâce à ce principe, un concept peut en fait représenter une ressource (comme dans RDF [W3C]). La seule contrainte est qu'il n'est pas possible d'avoir dans le même

⁶ Adaptive LEArning

⁷ Layered adaptive hypermedia authoring model and its algebraic Operators

⁸ My Online Teacher

modèle des concepts qui représentent des ressources et des objectifs pédagogiques.

- **Le modèle d'objectifs et de contraintes (GM)** : ce modèle filtre, regroupe et restructure le modèle conceptuel de contenu (DM) en fonction d'un objectif pédagogique. Les nœuds du graphe qui le composent sont reliés par des arcs qui peuvent être des relations du type AND-OR. Les relations de type OR peuvent être pondérées donnant ainsi un poids différents à chaque arc. Il est alors possible d'utiliser une relation d'ordre ou de dépendance entre plusieurs concepts afin d'exprimer des contraintes pédagogiques pour la progression d'un élève dans le modèle de contenu (DM). La Figure 1-8 montre un exemple de modèle GM composé de cinq concepts et quatre relations. Ce graphe exprime par exemple que le système autorisera l'élève à étudier le concept C4 lorsque son niveau de connaissances sera suffisant sur C1 ou C2. Pour étudier le concept C5, l'élève devra maîtriser les concepts C2 et C3.

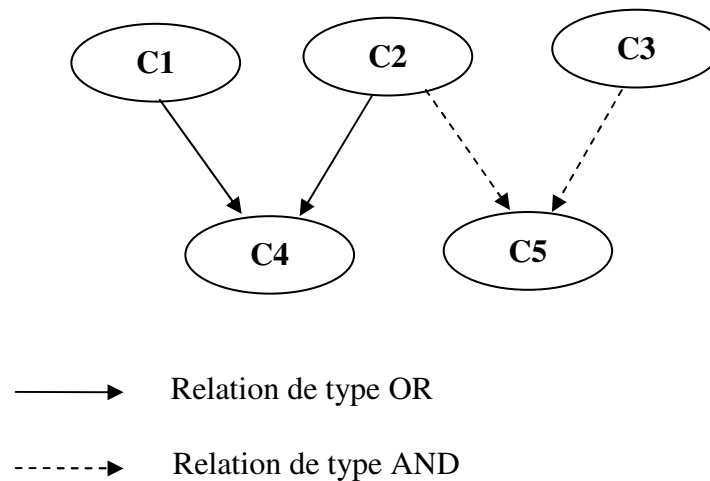


Figure 1-8 : Exemple de relations AND-OR dans le modèle GM

- **Le modèle de l'utilisateur (UM)** : il est composé d'un graphe de concepts bâti sur la base du modèle d'objectifs et de contraintes (GM). A Chaque nœud du graphe sont attachés différents attributs [Cristea et al, 2003c]. Le principal étant celui qui représente le niveau de connaissances de l'élève sur le concept correspondant.

Avec ce modèle de connaissances multicouche, il est possible d'envisager de trouver les concepts qui sont à l'origine d'un échec dans la séquence pédagogique. En revanche, il est toujours impossible de détecter que l'échec n'est peut être pas dû à un niveau de connaissance insuffisant mais simplement à l'inadaptation de la séquence pédagogique au style d'apprentissage de l'élève.

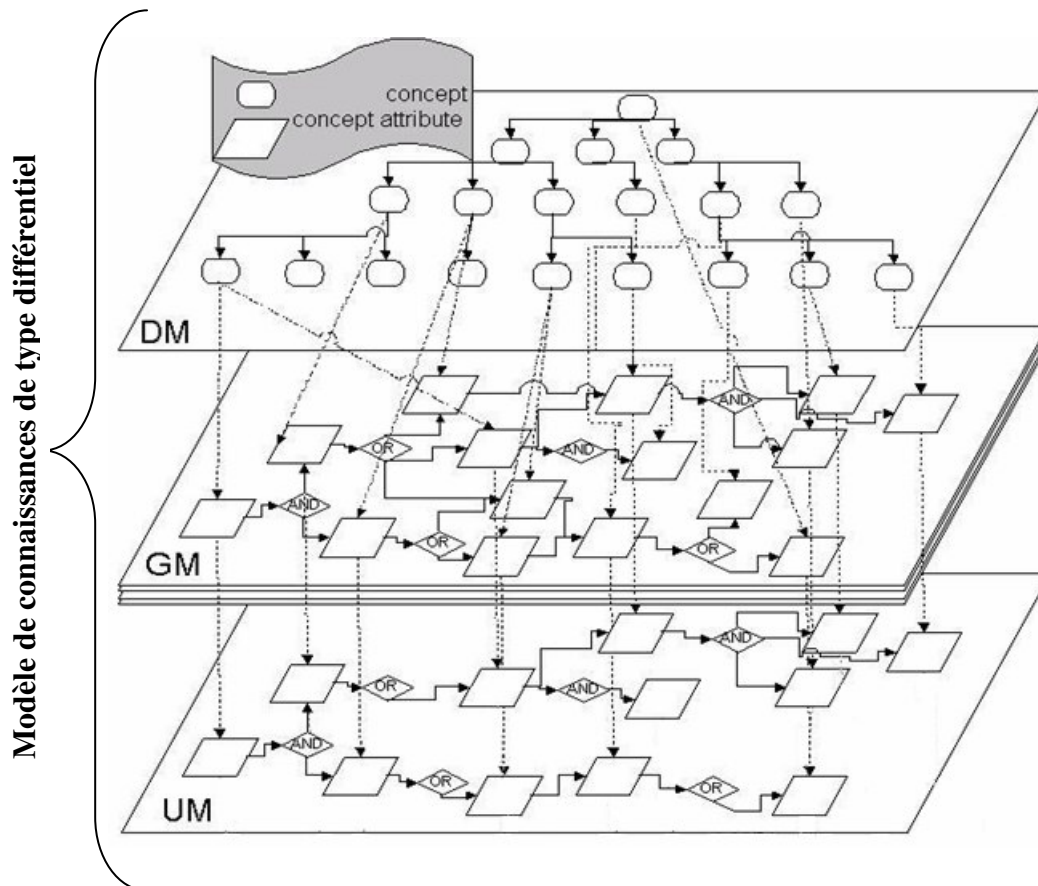


Figure 1-9 : le modèle de connaissances de LAOS

Déterminer le niveau de connaissances de l'élève dans le modèle pour chaque concept est une tâche complexe [Kristofic et al, 2005]. Suivant les recommandations énoncées dans le cadre du projet Européen ADAPT [Cristea et al, 2004c] pour aider à cette tâche, certains systèmes utilisent comme attributs :

- **le temps passé par (sur un ensemble de) concept** : la mesure offerte est dans ce cas simplement une estimation de la connaissance qualifiée par un état : « **visité** » ou « **non visité** ». L'état « **visité** » permettant de faire l'hypothèse que le concept est acquis ;
- **la performance mesurée par une évaluation sommative** : la mesure offerte est dans ce cas la performance de l'élève qui est plus précise que la précédente et est généralement une valeur numérique bornée dans un intervalle (ex : [0,20]).

La Figure 1-10 montre un exemple de représentation XML d'un modèle de connaissances de l'élève [Trella et al, 2002]. A chaque concept est associée une valeur qui définit le niveau de connaissances estimé (estimated) ou évalué (checked) de l'élève. Dans ce modèle, d'autres informations apparaissent comme l'identifiant de l'élève (s432), l'identifiant du cours (d03) et le dernier concept visité (c5).

```
<!DOCTYPE STUDENT_MODEL SYSTEM
"http://sirius.lcc.uma.es/medea/dtd/STUDENT_MODEL.dtd">
<STUDENT_MODEL courseid="d03" lastconcept="c5" studentid="s432">
  <ESTIMATED_STUDENT_MODEL>
    <CONCEPT id="c1" value="VERY WELL"/>
    <CONCEPT id="c2" value="VERY BAD"/>
    [...]
    <CONCEPT id="c5" value="WELL"/>
    <CONCEPT id="c6" value="VERY BAD"/>
  </ESTIMATED_STUDENT_MODEL>
  <CHECKED_STUDENT_MODEL>
    <CONCEPT id="c1" value="VERY BAD"/>
    <CONCEPT id="c3" value="REGULAR"/>
    <CONCEPT id="c5" value="REGULAR"/>
    <CONCEPT id="c6" value="VERY BAD"/>
  </CHECKED_STUDENT_MODEL>
</STUDENT_MODEL>
```

Figure 1-10 : Représentation XML d'un Modèle de connaissances de l'élève

Le modèle de connaissances d'un élève est dépendant du domaine enseigné. De ce fait, il devra être créé un modèle de ce type pour chaque matière étudiée.

Dans [Hibou et Labat, 2004], les auteurs proposent un modèle de connaissance de l'élève basé sur un réseau bayésien composé de deux couches (Figure 1-11) : la couche Domaine et la couche Activité. L'hypothèse que les auteurs posent est que la couche domaine dépend du profil de l'élève alors que la couche activité varie avec les activités pédagogiques sur lesquelles l'élève travaille.

La couche domaine se compose de nœuds « Knowledge » et « Know-How » qui sont utilisés pour définir le profil de l'élève. La structure et les tables de probabilités de cette couche sont spécifiées par des experts du domaine. Ils expriment ainsi des règles de conception pédagogique et didactique qui relient les concepts à apprendre « Knowledge » aux savoir-faire « Know-How ».

La couche activité est composée de nœuds « Item » qui représentent les éléments d'évaluation des activités. Les nœuds « Know-How » de la couche domaine sont reliés aux nœuds « Item » avec lesquels il existe une dépendance. La probabilité $P(item | Know - How)$ exprime le niveau de difficulté de l'exercice. Les auteurs avancent ainsi l'idée que le niveau de difficulté de la tâche dans un environnement pédagogique est directement lié à l'objectif d'enseignement. A travers l'observation des nœuds « Item », l'information sur le niveau de connaissance de l'élève est propagée à l'ensemble des nœuds de la couche domaine.

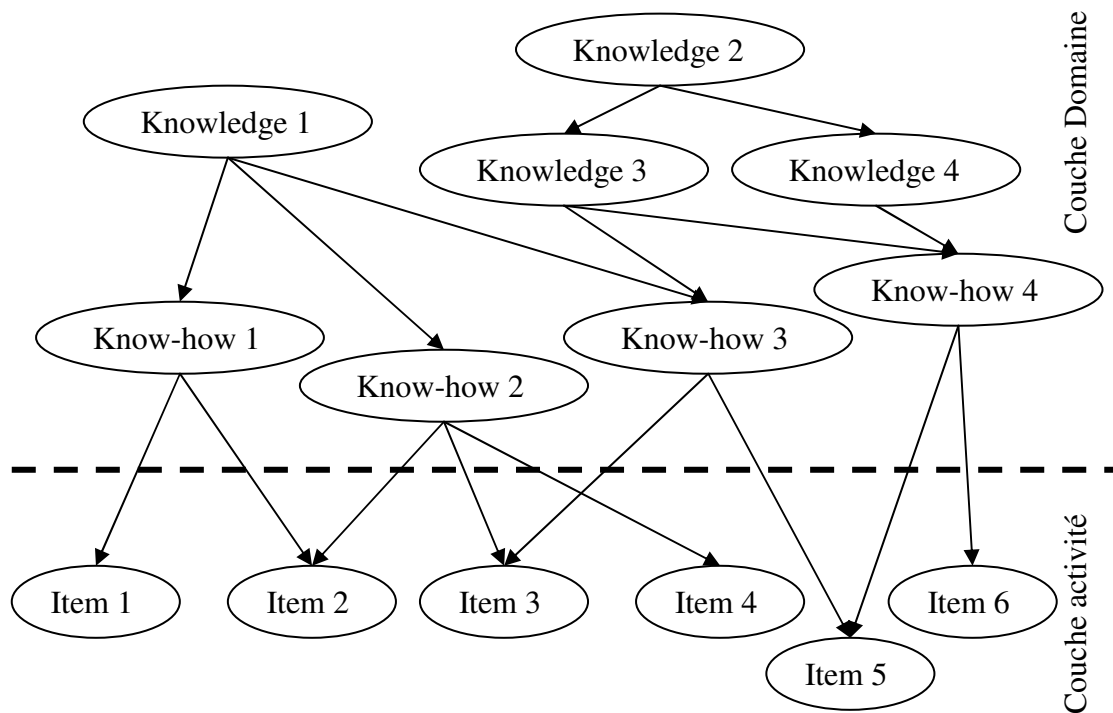


Figure 1-11 : couche domaine/couche activité

Il existe plusieurs versions de la couche domaine. Les nœuds sont identiques, mais la structure et les paramètres du réseau peuvent varier. Ces différentes versions de la couche domaine représentent les différents profils de l'élève que le système est capable de gérer. Le choix d'un profil parmi l'ensemble des profils est fait par une inférence *abductive* dans le réseau. Le principe est de sélectionner le réseau qui est le plus proche des observations faites par les experts. Considérons $Pr o_1, \dots, Pr o_m$ les m profils possibles et $I_1 = i_1, \dots, I_r = i_r$ les évidences sur les r variables observées de la couche des activités. Alors, pour chaque profil $Pr o_i$, on calcule la configuration la plus probable v_{pro}^{abd} sachant les évidences $I_1 = i_1, \dots, I_r = i_r$. Le nouveau profil associé à l'élève sera $Pr o_k$ tel que

$$k = \arg \max_{i \in \{1, \dots, m\}} P(V = v_{Pr o_i}^{abd} | I_1 = i_1, \dots, I_r = i_r).$$

Modélisation des préférences

Ce genre de modélisation a été utilisé dans les AHS. La nature personnelle des préférences ne permet pas de les modéliser de la même manière que les connaissances. Souvent, elles sont représentées par des couples *attribut - valeur* pouvant être déterminées par un vocabulaire précis.

Il est possible de définir ainsi la langue à utiliser de préférence pour les textes apparaissant dans les contenus pédagogiques [Klyne et al, 2003], ou bien encore, de préciser si l'on doit jouer la séquence pédagogique en incluant du son ou de la vidéo.

De même, on peut préciser des informations plus particulièrement tournées vers les caractéristiques du matériel utilisé par l'élève pour suivre la séquence pédagogique, comme la résolution de l'écran ou le nombre de couleurs. Le modèle de présentation proposé dans le cadre de la méthodologie Hera [Frasincar et al, 2004] est de type expertise-partielle sur le modèle de contenu qui spécifie les données qu'il est nécessaire de présenter. Il capture les préférences de l'élève pour traiter l'affichage des contenus. Par exemple, l'attribut « imageCapable » aura la valeur « yes » pour indiquer que le périphérique peut afficher des images.

Le principe est de manipuler ces attributs avec un ensemble de règles exprimées sous forme de structures conditionnelles [Cristea et al, 2004b] permettant d'obtenir l'adaptation de la séquence pédagogique aux préférences de l'élève. Les stratégies d'adaptation couramment mise en œuvre sont l'annotation, la gestion de l'apparition des liens, la variation de séquence de matériels pédagogiques et enfin la gestion de la visualisation de parties de contenus dans les pages [Brusilovsky, 2003b]. Les auteurs précisent de plus, que ces préférences sont là aussi pour indiquer comment l'élève préfère percevoir et traiter l'information. Elles peuvent inclure des informations qui caractérisent son style d'apprentissage, sa personnalité, son sexe, ses capacités ou handicaps physiques [Cristea et al, 2004c]. Dans [Trella et al, 2002], les auteurs maintiennent dans le modèle de préférences de l'élève des informations allant dans ce sens (Tableau 1-1)

Caractéristiques élève	Valeurs
Développement cognitif (capacité à formaliser et comprendre des concepts abstraits)	Elevé, moyen, bas
Motivation	Elevé, moyen, bas
Style d'apprentissage	Théorie, exercice
Expérience avec l'ordinateur	Beaucoup, normal, peu, aucun
Débit internet	Elevé, moyen, bas

Tableau 1-1 : Modèle de préférences de l'élève

Dans [Schuyten et al, 2007], les auteurs proposent un système d'apprentissage des statistiques. Le modèle de préférences utilisé se compose d'une première partie qui précise le mode préféré de représentation de l'information. Deux catégories sont considérées. Ceux qui préfèrent une information textuelle et ceux qui préfèrent une information graphique ou symbolique. La seconde partie du modèle décrit les aspects comportementaux de l'élève. On y trouve par exemple une estimation du degré d'intérêt pour les statistiques [Beenen et al, 1976 ; SMGS, 1968] ou bien l'aptitude à utiliser un ordinateur [Francis, 1993]. L'ensemble de ces informations est recueilli grâce à des questionnaires que l'élève remplit en début d'année. Ceci comporte deux principaux problèmes : la validité et la mise à jour des informations recueillies. En effet, il est généralement très difficile à un élève d'exprimer ses préférences. En plus, il ne lui est pas permis de modifier en cours d'année les informations fournies.

EDUCE [Kelly et al, 2004a,b] est un système adaptatif d'enseignement qui utilise la théorie des intelligences multiples (MI) [Gardner, 1983, 1993, 1999] pour créer dynamiquement le modèle de préférences d'un élève. La théorie MI identifie huit types distincts d'intelligence qui sont utilisés dans la résolution de problèmes, dans la création de matériels comme la composition, la musique ou la poésie ou d'autres activités éducatives. EDUCE utilise quatre de ces huit types d'intelligence.

- Logique/mathématique : mesure la capacité à détecter des modèles, à raisonner par déduction et à penser logiquement.
- Verbal/linguistique : mesure la maîtrise du langage pour exprimer des idées.
- Visuel/spatial : mesure la capacité à manipuler et créer des images mentales pour résoudre des problèmes.
- Musical/rythmique : mesure la capacité à reconnaître et composer des refrains, des introductions et des rythmes musicaux.

Ces quatre types d'intelligence constituent les informations de base pour modéliser les préférences de l'élève. EDUCE fabrique le modèle de préférences en observant, analysant et enregistrant les interactions de l'élève avec les différents types de matériels pédagogiques (ressources) spécifiques à chaque type d'intelligence (Figure 1-12).

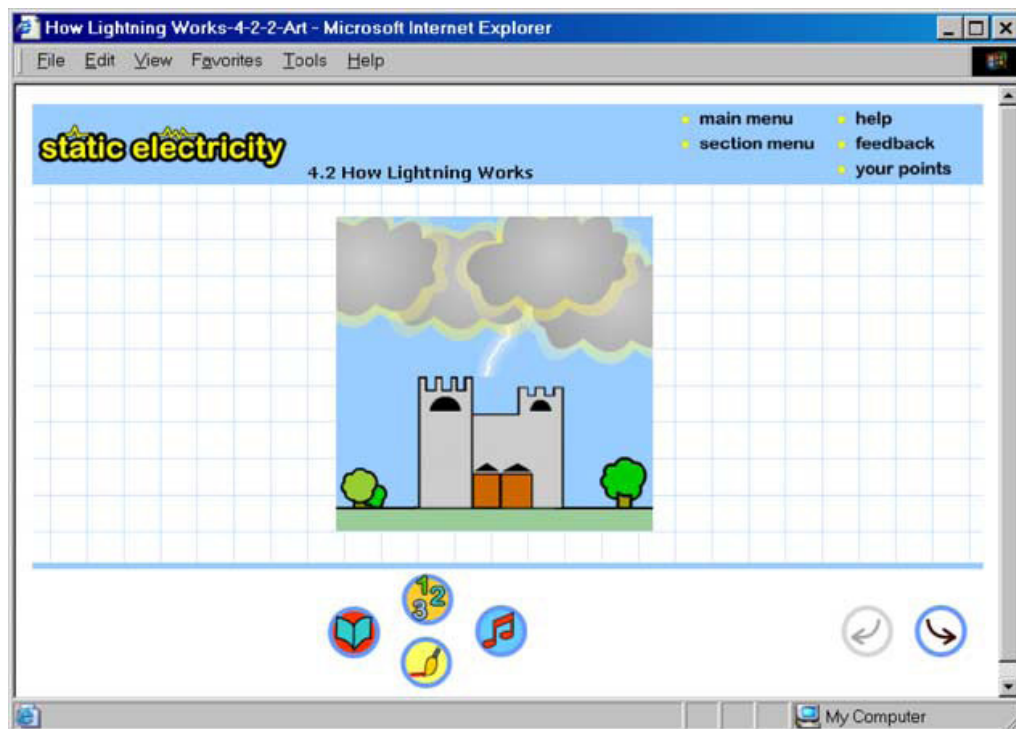


Figure 1-12 : exemple de ressource basée sur les principes de l'intelligence visuel/spatial

En particulier, le modèle décrit comment un élève utilise les différentes ressources en utilisant les critères suivants :

- Est-ce qu'un élève a passé peu de temps à utiliser une ressource ? Cette information exprime le degré d'intérêt pour la ressource. Un temps très faible indique que l'élève a juste jeté un coup d'œil à la ressource sans y apporter un grand intérêt avant de passer à la prochaine.
- Est-ce qu'un élève a passé beaucoup de temps à utiliser une ressource ? Si un élève passe beaucoup plus de temps que prévu à utiliser une ressource, cela veut certainement dire qu'il est intéressé par son contenu !
- Quelle est la ressource que l'élève a utilisée en premier ? Un élève avec des préférences très affirmées choisira sa ressource préférée en premier.
- Est-ce que l'élève utilise une seule ressource ou plusieurs ? Quelques élèves n'ont pas de préférences très affirmées. Ils vont donc utiliser différentes ressources pour examiner un concept de façons multiples.
- Est-ce que l'élève utilise la ressource plus d'une fois ? Revenir plusieurs fois sur la même ressource confirme la préférence de l'élève.
- Est-ce que l'élève fait une bonne évaluation après l'utilisation de la ressource ? Certains types de ressources peuvent encourager l'élève à pratiquer une évaluation et obtenir de meilleurs résultats par un meilleur niveau d'apprentissage.

L'ensemble de ces informations capture le temps que l'élève passe à utiliser une ressource, la première ressource utilisée, la ressource qui est souvent réutilisée et l'aide que la ressource apporte à l'élève dans le cadre d'une évaluation. Pour compléter ces informations recueillies dynamiquement, EDUCE maintient une mesure de chaque type d'intelligence, basée sur l'analyse des réponses faites au questionnaire MIDAS [Shearer, 1996]. Le modèle utilisé ne permet pas de mesurer un indice de confiance ou un pourcentage d'incertitude sur les valeurs obtenues pour l'ensemble des critères observés. Il est alors difficile de prendre une décision en cas d'échec de l'élève à une évaluation. Comment choisir ce qu'il faut changer dans l'évaluation des préférences de l'élève ?

Diverses études montrent que le modèle de préférences d'un élève est indépendant du domaine enseigné [Lin & al, 2003] et que donc il peut être partagé par plusieurs systèmes pour l'apprentissage de différentes matières. Ce genre de modèle adopte bien souvent une approche de type social ayant comme avantage de déduire les préférences d'un élève à partir des préférences d'autres. Mais dans ce cas, il faut que le modèle soit dans un format offrant une interopérabilité suffisante pour être partagé par différents systèmes.

1.2.3. Acquisition du modèle : statique, dynamique ou mixte

Dans cette section, nous essayons de répondre aux questions suivantes :

- Comment mesurer les valeurs des attributs du modèle de l'élève ?
- Quand effectuer la mesure ? en début de session, à la fin de la session ou bien en cours de session ?

Dans [Cristea et al, 2004c], les auteurs envisagent plusieurs solutions, telles que la saisie explicite en début de session par l'élève de quelques attributs le concernant (ex : background knowledge, learning preference) ou bien une saisie indirecte de données pour définir son style d'apprentissage par le biais d'un questionnaire comme le fait le système INSPIRE [Papanikolaou et al, 2003]. Une autre solution consiste à déléguer la tâche au système qui peut évaluer en cours de session les attributs indirects de l'élève par une analyse de la navigation (pages ou liens utilisés) ou bien le temps passé sur un concept donné (pages ou groupes de pages). Ce principe est utilisé par le serveur de gestion de modèles d'élèves CUMULATE présenté dans le cadre d'un système d'enseignement adaptatif distribué [Brusilovsky, 2004c]. Celui-ci qui enregistre, entre autre, les activités de l'élève au cours de la session (ex : lecture de page, réponse à une question, analyse d'un exemple).

On trouve donc dans la littérature que l'acquisition pertinente d'informations sur les élèves peut être faite par des méthodes automatiques [Webb et al, 2001] ou bien par un questionnement direct de l'élève [Waern, 2004, Kay, 2001] qu'il est possible de classer en trois familles suivant le mode d'acquisition : *acquisition statique*, *acquisition dynamique* et *acquisition mixte*.

Acquisition statique

L'acquisition statique consiste à demander explicitement à l'élève ses préférences en début de l'activité. C'est le moyen le plus simple qui consiste à représenter les préférences d'un élève sous la forme d'un vecteur de mots clefs pondérés [Marinilli et al, 1999].

La majorité des applications qui utilisent cette technique recueillent les données par l'intermédiaire d'un questionnaire qui est soumis à l'élève en début d'utilisation du système. Cette technique trouve vite ses limites car les élèves ont du mal à conceptualiser leurs besoins au travers d'activités réflexives et les questionnaires deviennent très vite complexes ou sont incapables de capturer l'ensemble des préférences de l'élève [Manber et al, 2000].

Certains auteurs tentent d'apporter une solution à ce problème en demandant à l'élève de donner son sentiment sur un ensemble de documents [Billsus et Pazzani, 1999]. Les réponses sont alors utilisées pour déduire les préférences de l'élève.

Acquisition dynamique

L'acquisition dynamique consiste à observer l'élève en situation, pendant l'activité. Le système infère le modèle à partir de la trace des situations [Broisin, Vidal et al., 2006]. Dans [Ardissono et al, 1999], les auteurs mettent à jour l'appartenance probabiliste d'un élève à des stéréotypes par cette technique.

L'utilisation des méthodes dynamiques pour deviner les objectifs et les intérêts d'un élève sont plus intéressantes car elles permettent de garder une certaine fluidité de l'interaction sans détourner l'élève de sa principale tâche. En effet, il n'est pas souhaitable de parasiter l'activité de l'élève par des questions ayant pour but de renseigner son modèle afin de personnaliser l'interaction.

Dans [Kristofic et al, 2005], les auteurs préconisent que la meilleure façon de découvrir les caractéristiques de l'élève pour renseigner son modèle est de le faire par l'intermédiaire de fichiers de logs qui sont créés tout au long de la session. Ces fichiers peuvent contenir par exemple les attributs caractérisant les actions de l'élève, tels que :

- **timestamp** : définit l'heure de l'action. Il est utilisé pour ordonner les actions dans le temps et calculer le temps passé sur chaque concept ;
- **user identification** : identifie de façon unique chaque élève. Il est alors possible d'associer les actions à chaque élève afin d'envisager une personnalisation de l'analyse ;
- **type** : offre un raisonnement sur plusieurs niveaux. Les auteurs ont défini des types pour quelques actions élémentaires (avec la possibilité de les compléter par la suite)
 - **Login et Logout** qui délimitent la session d'apprentissage d'un élève,
 - **ConceptVisit et FragmentDisplay** qui contiennent des informations sur les concepts visités ainsi que les contenus élémentaires présentés pour un concept donné.

- **object** : spécifie l'identification de l'objet sur lequel l'action de l'élève a porté. Par exemple pour l'action **ConceptVisit** l'attribut objet contient l'identification du concept visité par l'élève.

Les systèmes d'enseignement adaptatifs AHA ! et ALEA exploitent aussi des fichiers de logs par des techniques de fouille de données⁹ pour renseigner le modèle de l'élève.

Acquisition mixte

L'acquisition mixte combine à la fois une phase d'initialisation statique pour bâtir un premier modèle de l'élève et une phase dynamique de mise à jour du modèle en observant les interactions de l'élève avec le système. Dans [Smyth et Cotter, 2000], les auteurs expliquent qu'un élève est prêt à donner ses préférences avant qu'il n'ait des objectifs autres que de se faire connaître par le système, donc en début d'échange. A partir du moment où l'élève est engagé vers un objectif d'apprentissage, il souhaite ne pas le quitter. C'est pourquoi on doit utiliser les interactions de l'élève sans que celui-ci en ait conscience pour essayer d'en déduire ses objectifs et ses préférences afin de conserver la fluidité de l'interaction avec le système d'enseignement.

1.2.4. Discussion

Les modèles de connaissances présentés ci-dessus montrent uniquement le niveau de connaissances d'un élève dans un domaine particulier. Il est difficile de les exploiter pour identifier les causes éventuelles d'échecs et ainsi d'y remédier par l'ajustement du contenu de la séquence pédagogique. Les modèles de préférences maintiennent pour la plupart simplement les caractéristiques du support préféré de l'information. Ils admettent que ces préférences peuvent être identifiées par des questionnaires posés à l'élève. Ces deux types de modèles ne gèrent pas l'incertitude et l'évolution de l'information qu'ils maintiennent.

Dans cette thèse, nous proposons un modèle adaptatif de l'élève. Il se compose d'un modèle de connaissances capable d'émettre une hypothèse sur la cause éventuelle d'un échec et de proposer ainsi une solution pour remédier au problème. Il comporte aussi un modèle de préférences capable d'identifier les préférences d'un élève par observation de son comportement dans la séquence pédagogique. Ces deux modèles utilisent des structures qui offrent la possibilité de représenter un niveau d'incertitude sur les informations qu'ils manipulent.

⁹ data mining

1.3. Modélisation de contenus

Le modèle de contenus des systèmes hypermédia d'enseignement est composé d'un ensemble d'éléments qui représentent des fragments élémentaires de connaissances du domaine étudié [Brusilovsky, 2003d]. Ils sont nommés différemment en fonction des systèmes hypermédia d'enseignement (concepts, éléments de connaissances, sujet) et sont souvent organisés de façon hiérarchique.

Les systèmes hypermédia d'enseignement manipulent des concepts de différents types. Par exemple, dans le contexte de l'enseignement de la programmation par des exercices, un concept peut contenir de la connaissance générale (ex : programming scheme) ou de la connaissance spécifique (un exercice avec un exemple de code source comme application spécifique d'un programming scheme).

Le système ALEA possède un modèle de contenu hiérarchique qui suit ce principe. Il est composé de concepts auxquels les auteurs ont associé les types suivants : Text, Programming scheme, Exercise, Test. Chaque concept est ensuite décomposé en fragments élémentaires (Text, Exercise definition, Exercise hint, Exercise solution, Source code).

Le framework MEDEA [Trella et al, 2002] qui est l'acronyme Espagnol de "Methodology and Tools for the Development of Intelligent Environments of Teaching and Learning" est composé d'un modèle de contenus organisé sous la forme d'un réseau sémantique de connaissances. D'un point de vue conceptuel, le domaine à étudier est défini par un réseau sémantique de concepts et des données pédagogiques qui visent à aider l'enseignement des concepts. Par exemple, chaque concept possède comme attribut une valeur qui indique quel est le seuil minimum au dessus du quel on estime que l'élève maîtrise le concept étudié. D'un point de vue de l'implémentation du modèle, MEDEA exploite trois éléments de base qui sont : *les concepts*, *les relations* entre les concepts et *les types d'évaluation* pour représenter le niveau de connaissances de l'élève. La Figure 1-13 montre un exemple de représentation XML de ce modèle de contenu.

```

<!DOCTYPE DOMAIN_MODEL SYSTEM
"http://sirius.lcc.uma.es/medea/dtd/DOMAIN_MODEL.dtd">
<DOMAIN_MODEL id="domain_model01" name="Logic of proposals ">
<EVALUATION_TYPES>
  <EVALENUM id="EvalEnum" default_minimum_mark="Passed">
    <ENUMERATED id="Passed"/>
    <ENUMERATED id="Failed"/>
  </EVALENUM>
  <EVALREAL id="EvalReal" lower_boundary="0" upper_boundary="10"
  default_minimum_mark="5"/>
</EVALUATION_TYPES>
<CONCEPTS>
  <CONCEPT id="t1" idref_evaluation="EvalEnum" name="Introduction"
  difficulty="Low"/>
  <CONCEPT id="t2" idref_evaluation="EvalEnum" name="Formal syntax"
  difficulty="Low"/>
  <CONCEPT id="t3" idref_evaluation="EvalEnum" name="Semantic"
  difficulty="Low"/>
  [...]
  <CONCEPT id="c32" idref_evaluation="EvalReal" name="CDN"
  difficulty="High"/>
</CONCEPTS>
<RELATIONS>
  <RELATION id="r1" id_origin_concept="c21" id_destiny_concept="c1"
  type="prerequisite"/>
  <RELATION id="r4" id_origin_concept="c4" id_destiny_concept="c2"
  type="is_a "/>
  [...]
  <RELATION id="r73" id_origin_concept="c31" id_destiny_concept="t6"
  type="belongs_to"/>
  <RELATION id="r74" id_origin_concept="c32" id_destiny_concept="t6"
  type="belongs_to "/>
</RELATIONS>
</DOMAIN_MODEL>

```

Figure 1-13 : Représentation XML d'un modèle de contenus

On y distingue dans la première section (**EVALUATION_TYPES**) un type d'évaluation énuméré (EvalNum) qui prend deux valeurs (Passed et Failed) ainsi qu'un type d'évaluation numérique (EvalReal) borné dans l'intervalle [0,10]. La seconde section (**CONCEPTS**) caractérise les concepts en indiquant leur type d'évaluation ainsi que leur niveau de difficulté. La troisième section (**RELATIONS**) décrit les relations qui peuvent exister entre les concepts. Il existe plusieurs types de relation : *prerequisite*, *part of*, *is a*, *belongs to*, *is useful to understand*, *is similar to* et *is opposite to*. Chaque type de relation peut définir un graphe acyclique de concepts qui peut être utilisé pour adapter la navigation de l'élève dans les concepts du domaine à étudier grâce à un modèle de l'élève de type overlay sur le modèle de contenu [Cristea et al, 2004c]. Cette représentation souffre d'un manque important pour un système d'enseignement. Elle ne permet pas d'associer à un concept des activités pédagogiques. On ne définit donc pas le matériel pédagogique que l'élève devrait utiliser pendant la session d'enseignement.

Dans [Brusilovsky et al, 2003c], les auteurs ont conçu le modèle de contenu du système DCG¹⁰ sur la base du modèle présenté ci-dessus. Chaque nœud représente un élément de connaissances (*concepts, sujet, règles, etc.*) alors que les arcs représentent des opérateurs booléens (*ET, OU*). Si deux nœuds A et B sont connectés avec un troisième nœud C à l'aide d'un arc de type ET, cela indique que les concepts associés aux nœuds A et B doivent être acquis par l'élève pour qu'il puisse étudier le concept associé au nœud C. Dans le cas où l'arc est du type OU alors la maîtrise de l'un des deux concepts suffit pour passer à l'étude du concept associé au nœud C.

Les auteurs ont de plus ajouté un autre attribut aux arcs afin d'augmenter leur sémantique. En effet, les arcs peuvent aussi être qualifiés par un second attribut qui peut prendre les valeurs : *temporel, causal, analogie, pré-requis, agrégation* ou *généralisation*. Ainsi lorsque deux nœuds A et B sont reliés à un troisième nœud C avec un arc de type ET-AGREGATION, cela signifie que C a comme sous-concepts A et B. Alors que s'ils sont connectés avec un arc de type OU-GENERALISATION, cela indique que C est un concept général qui peut avoir comme instances possibles A ou B.

Une utilisation simpliste de la structure consiste à faire apparaître dans le modèle une seule sémantique pour les arcs. Il est alors possible, par cette structure, d'adapter l'ordonnancement des concepts¹¹ en reliant les nœuds par des arcs de type ET-PREREQUIS ou OU-PREREQUIS. Chaque élément du modèle (nœud et arc) est relié à un matériel pédagogique qui constitue la ressource que l'élève manipulera lors de la session. Ces ressourcesinstancient différents moyens pour enseigner le concept (*introduction, explication, exemple, exercice* ou *test*) avec différents médias (*texte, image, animation, vidéo, etc.*).

Le système INSPIRE¹² [Papanikolaou et al, 2003] exploite cette structure pour son modèle de contenus (Figure 1-14).

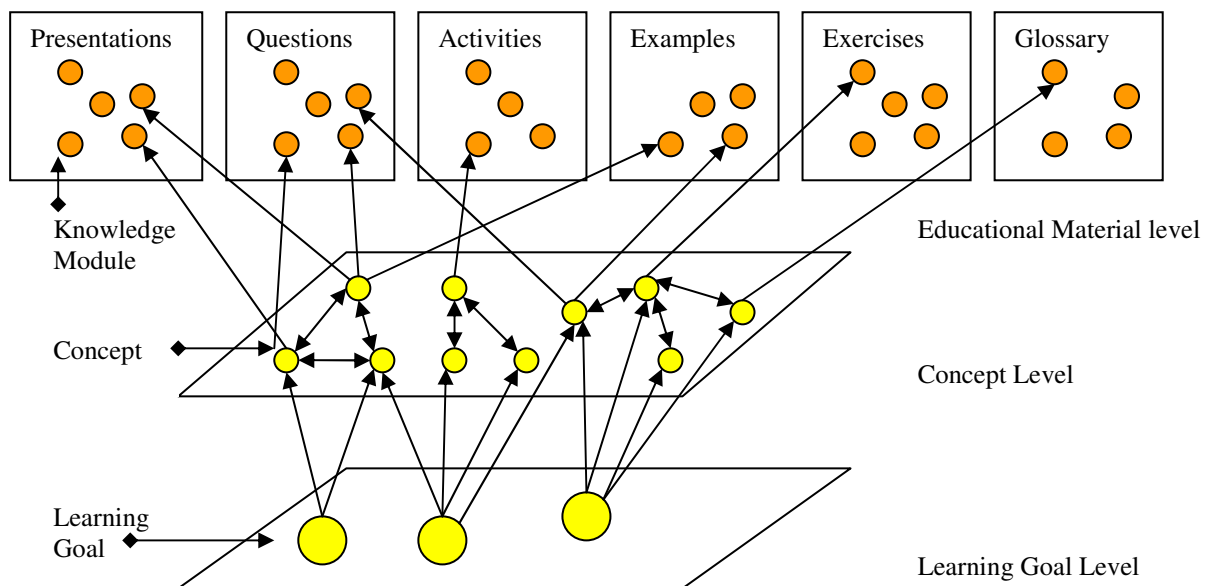


Figure 1-14 : modèle de contenu du système INSPIRE

¹⁰ Dynamic Course Generation

¹¹ curriculum sequencing

¹² INtelligent System for Personalized Instruction in a Remote Environment

Il s'agit d'un système hypermédia d'enseignement adaptatif¹³. Il a comme objectif de rendre l'élève maître de son parcours d'apprentissage. Le modèle de contenus du système est basé sur la notion d'objectifs d'apprentissage que l'élève est autorisé à choisir pour débiter son étude. INSPIRE fournit à l'élève différentes sortes d'activités pédagogiques et de ressources. Le modèle du domaine est représenté par trois niveaux hiérarchiques : **Objectif d'apprentissage**, **concepts** et **matériels pédagogiques**.

Un objectif pédagogique correspond à un sujet du domaine à étudier, qui peut être choisi par un élève. Chaque objectif est associé à un ensemble de concepts du domaine à étudier ; et chaque concept est lui-même relié à un ensemble de matériels pédagogiques considérés comme des modules de connaissances dans le domaine étudié. Les modules de connaissances constituent de multiples représentations externes des concepts tels que :

- la **Présentation théorique** d'un concept : définition, description, conclusion
- des **Questions** d'introduction ou d'évaluation d'un concept
- des **Exemples** pour un concept : instance concrète d'un concept, exemples d'application du concept, analogie
- des **Exercices** sur un concept
- des **Activités** : simulation sur ordinateur, expérimentation, étude de cas.

Les concepts peuvent être aussi reliés entre eux, formant à nouveau un graphe de dépendances entre objectifs dont les arcs représentent des relations de type ET. Cela indique que les concepts associés aux nœuds A et B doivent être acquis par l'élève pour qu'il puisse étudier le concept associé au nœud C.

Les pages développées pour chaque concept incluent les modules de connaissances qui aident l'élève pour atteindre trois niveaux de performance :

- Performance de niveau **Rappel** : inclut les modules de connaissances qui introduisent le concept et amène l'élève à spéculer sur les nouvelles idées introduites telles que les présentations théoriques du concept, les questions (introduction ou self-évaluation), les instances du concept (exemple réel ou analogie du concept).
- Performance de niveau **Utilisation** : inclut les modules de connaissances qui conseillent ou aident l'élève à appliquer le concept dans un cas spécifique tels que les exemples d'application, les exercices, les activités de simulation sur un ordinateur, les expérimentations.
- Performance de niveau **Découverte** : inclut les modules de connaissances qui stimulent l'élève à trouver de nouveaux principes ou procédures telles que les activités de simulation sur un ordinateur, les expérimentations et les études de cas. A ce stade, l'élève doit rendre un livrable au tuteur et bien souvent des activités collaboratives lui sont proposées.

Chaque niveau de performance inclut des éléments de présentations, des éléments de pratiques et des éléments de tests.

¹³ AEH : Adaptive Educational Hypermedia

De plus, le matériel pédagogique créé pour chaque concept comporte :

- une *page d'introduction* qui présente sommairement les idées fondamentales couvertes par le concept,
- une *page de synthèse* qui reprend les points particuliers associés au concept,
- une *page d'évaluation sommative* qui mesure le niveau de connaissances de l'élève sur le concept ainsi que sur ses pré-requis.

INSPIRE suit les recommandations du projet ARIADNE [ARIADNE, 2000] en décrivant les modules de connaissances par des métadonnées. Celles-ci utilisent trois types de descripteurs : *attributs pédagogiques*, *sémantique de la ressource*, *information générale sur la ressource*. Le contenu de ces descripteurs est utilisé par le système pour développer une stratégie d'adaptation qui consiste à proposer des contenus adaptés au niveau de connaissance de l'élève.

Discussion

Les modèles de contenus étudiés ci-dessus sont largement orientés vers la gestion du niveau de connaissance d'un élève. Leur sémantique n'inclut que rarement la notion d'activité pédagogique. En effet, le principe fondamental de séquences composées d'une série d'activités variées (présentation théorique, question, exemple ...) est ignoré. Ceci les rend inappropriés pour adapter l'enseignement aux styles d'apprentissage.

Dans cette thèse, nous proposons un modèle de contenus qui offre une décomposition structurelle du domaine à enseigner en partant des tâches métier et allant jusqu'à celui de l'objet d'apprentissage. Cette décomposition fait apparaître à différents niveaux, les adaptations envisagées dans les contenus, en fonction des styles d'apprentissage.

1.4. Stratégies d'Adaptation

Les systèmes adaptatifs ont la capacité de s'ajuster à chaque élève en analysant des fichiers de logs ou directement les analyses des actions des élèves [Gaudioso et al, 2002]. Ils utilisent essentiellement deux techniques d'adaptation : *adaptation de la présentation* et *adaptation de la navigation* [Brusilovsky, 2001].

Dans certains systèmes tel que le système TANGOW¹⁴ [Carro et al, 1999], on utilise un modèle de l'élève de type social contenant, entre autre, les attributs d'âge et de capacité linguistique pour aider l'élève dans sa tâche. Le principe est de supposer que les élèves ayant des caractéristiques proches auront un comportement similaire face à une tâche d'apprentissage.

La capacité linguistique est estimée par un outil de test multimédia [Barker et al, 2002]. Cet outil évalue les qualités de lecture et d'écoute des élèves en utilisant le test SMOG. SMOG est un test standard utilisé dans la formation continue pour adulte en Grande Bretagne, basé sur les procédures Cloze [Taylor, 1953; Bormuth, 1968; Vaughn, 1995]. Il est efficace pour classer les élèves dans des groupes selon leur capacité linguistique.

¹⁴ Task-based Adaptive learner Guidance On the WWW

1.4.1. Adaptation de la présentation

Les systèmes qui adoptent l'adaptation de la présentation offrent le choix entre différents médias pour présenter le matériel pédagogique. C'est pourquoi, en plus du traditionnel média (texte) utilisé dans la plupart des systèmes éducatifs, il est proposé d'utiliser des vidéos, du son, des animations, etc [Brusilovsky, 2001].

Adapter la présentation peut aussi consister à modifier la manière de présenter une information à l'élève. Par exemple, il est possible de modifier le comportement d'un système lorsque l'élève réalise une évaluation formative [Trella et al, 2002]. Les réponses aux questions peuvent être fournies à l'élève une après l'autre à la suite de chaque question ou bien en un seul bloc à la fin du questionnaire. Trella préconise de corrélérer ce genre d'adaptation au niveau de motivation de l'élève. En effet, proposer la réponse après chaque question sera la solution adoptée lorsque l'élève montre peu de motivation à l'activité. Le principe étant d'essayer de stimuler sa curiosité en lui proposant une correction progressive du questionnaire. La difficulté reste néanmoins de mesurer correctement le niveau de motivation de l'élève.

1.4.2. Adaptation de la navigation

Les systèmes qui mettent en œuvre l'adaptation de la navigation ont une structure de contenus hiérarchique. Des hyperliens offrent à l'élève de naviguer entre les chapitres ou les sections. Deux techniques sont utilisées pour implémenter l'adaptation de la navigation : *annotation adaptative* et *curriculum sequencing*.

Généralement, un modèle de l'élève de type expertise-partielle à plusieurs niveaux est utilisé pour implémenter l'adaptation de la navigation [Weber et al, 2001b]. La majorité de ces approches se focalisent sur le niveau de connaissances de l'élève dans le domaine. Elles ne considèrent aucun autre facteur, tel l'aspect cognitif pour différencier les élèves.

Annotation adaptative

Cette technique consiste à changer l'ordre, la validité ou l'apparence des hyperliens. Les systèmes AHA ! et ALEA utilisent des techniques de traitement des données pour offrir à l'élève une navigation adaptée dans les contenus. En utilisant par exemple les techniques du « Sequential patterns mining » ou « Traversal patterns mining » pour analyser les données de fichiers de logs, ils découvrent l'ordre dans lequel les élèves préfèrent majoritairement visiter certains concepts. Cette information est utilisée pour recommander à un élève en cours de session le prochain concept à visiter.

Le système Hera [Frasincar et al, 2004] préconise une adaptation de la navigation basée sur le niveau d'expertise de l'élève dans le domaine étudié. Cette adaptation est statique, c'est à dire qu'elle n'évolue pas en cours de séquence. Deux modèles sont à considérer ici. Le modèle conceptuel qui spécifie les données (contenus sous forme hiérarchique) du domaine à étudier et le modèle d'application qui capture la logique de l'application (navigation entre les concepts). En début de séquence, le modèle d'application est conçu comme un modèle de type différentiel sur le modèle conceptuel. Le modèle d'application ne contient alors que les concepts que l'élève est considéré pouvoir suivre en fonction de son niveau d'expertise (débutant, normal, expert). Le modèle d'application est vu alors comme un sous-ensemble du modèle conceptuel.

Cela se traduit à l'écran par une modification de l'affichage des hyperliens auxquels le système ajoute des symboles colorés pour indiquer à l'élève ce qu'il est préférable de faire dans le contexte actuel (généralement en fonction de son niveau de connaissances sur le concept étudié). Le système ELM-ART [Weber et al, 2001a] ajoute à chaque lien présent à l'écran un feu bicolore (identique aux feux de circulation) qui conseille l'élève sur le parcours à suivre dans la page. Un feu de couleur rouge indiquera à l'élève qu'il ne lui est pas conseillé de suivre ce lien, alors qu'un feu de couleur verte l'engagera à découvrir le contenu de la page pointée.

Ce genre d'annotation ne donne pas à l'élève une image de son niveau de connaissances ou de sa progression dans le système sur le sujet étudié. Le système INSPIRE [Papanikolaou et al, 2003], pour combler cette lacune, complète l'information (qui dans ce cas est une lampe torche : éteint = déconseillé, allumée = conseillé) par une coupe dont le niveau de liquide est proportionnel au niveau de connaissances dans le sujet du lien. Une coupe vide indiquera que le système considère l'élève comme un novice dans le sujet, alors qu'une coupe pleine sera le reflet d'un niveau d'expert sur le domaine.

L'ensemble des systèmes analysés montre que seul le niveau de connaissance de l'élève est considéré pour proposer une adaptation. Il n'est pas pris en considération le style d'apprentissage qui pourrait influencer sur les choix proposés à l'élève : étudier des aspects théoriques ou bien avoir une activité pratique.

Ordonnement de programme¹⁵

Cette technique largement utilisée dans le domaine des ITS consiste à prendre des décisions sur les chapitres ou sections vers lesquels l'élève doit être dirigé en fonction de pré-requis et de ce qu'il a déjà effectué. L'idée est de générer un cours individualisé pour chaque élève en sélectionnant la meilleure activité pédagogique (présentation, exemple, question ou problème) pour l'apprentissage d'un concept donné. Ce choix ayant pour but ultime d'amener l'élève dans les meilleures conditions pour atteindre l'objectif pédagogique.

Par exemple, dans [Weibelzahl et al, 2002] l'élève subit un pré-test en début de chapitre pour vérifier s'il possède les pré-requis nécessaires à la compréhension du concept étudié. Dans le cas où l'élève n'obtient pas une note supérieure à un certain seuil, le système considère que l'élève ne possède pas les pré-requis et lui refuse donc l'accès aux contenus. A la fin du chapitre, l'élève subit une évaluation sommative afin de vérifier si l'objectif pédagogique visé est atteint. Ces informations sont ensuite stockées dans le modèle de l'élève qui est du type expertise-partielle.

¹⁵ Curriculum sequencing

Dans [Brusilovsky et al, 2003c], les auteurs présentent le système DCG¹⁶. Ce système est capable de générer un cours individualisé pour chaque élève en tenant compte de l'objectif pédagogique à atteindre (un concept ou un sujet que l'élève doit apprendre). La génération prend en compte le niveau de connaissances de l'élève. Le cœur de l'architecture du système est la représentation explicite de la structure des concepts du domaine étudié qui est dissociée du matériel pédagogique à utiliser. DCG utilise la structure des concepts (représentée par un ensemble de règles) comme une feuille de route pour générer le plan du cours. En tenant compte du concept que l'élève désire étudier ainsi que des concepts marqués comme acquis dans son modèle de l'élève (initialisé par un pré-test), un composant logiciel recherche la route qui connecte les concepts connus par l'élève à celui à étudier. L'élève voit les séquences pédagogiques reliées à chaque concept figurant dans le plan du cours généré. A chaque moment, l'élève peut tester son niveau de connaissances sur le concept courant en faisant une évaluation sommative. Le résultat du test est enregistré dans le modèle de l'élève comme étant le niveau de connaissances de l'élève pour le concept courant. Si l'élève n'est pas capable d'atteindre un score minimum au test, un nouveau plan peut être construit en cours de session.

Le système INSPIRE [Papanikolaou et al, 2003] est construit sur ce principe. Il a comme objectif d'aider les élèves dans la phase d'apprentissage en proposant des contenus adaptés à leur niveau de connaissances dans le domaine étudié [Bransford et al, 1999]. Ainsi, il propose aux novices dans le domaine étudié des contenus simplifiés au début de leur interaction avec le système, avant de les enrichir progressivement en fonction de leur performance en cours de session d'apprentissage. Basé sur la notion d'objectif d'apprentissage que l'élève sélectionne, INSPIRE génère les leçons correspondant à ce besoin d'apprentissage en fonction du niveau de connaissances de l'élève ainsi que de son style d'apprentissage [Papanikolaou et al, 2003].

1.4.3. Modèle d'adaptation à plusieurs couches

Les modèles d'adaptation dans les systèmes adaptatifs hypermédia sont traditionnellement composés d'un ensemble de règles IF-THEN qui sont déclenchées à l'apparition de certains événements (ex : accès à une page ou un concept). La rédaction de ces règles est fastidieuse et complexe pour un pédagogue non initié.

Dans [Cristea et al, 2003a], pour solutionner ce problème, les auteurs ont introduit un modèle d'adaptation (AM) (Figure 1-15) et un modèle de présentation (PM) dans le Framework LAOS.

¹⁶ Dynamic Course Generation

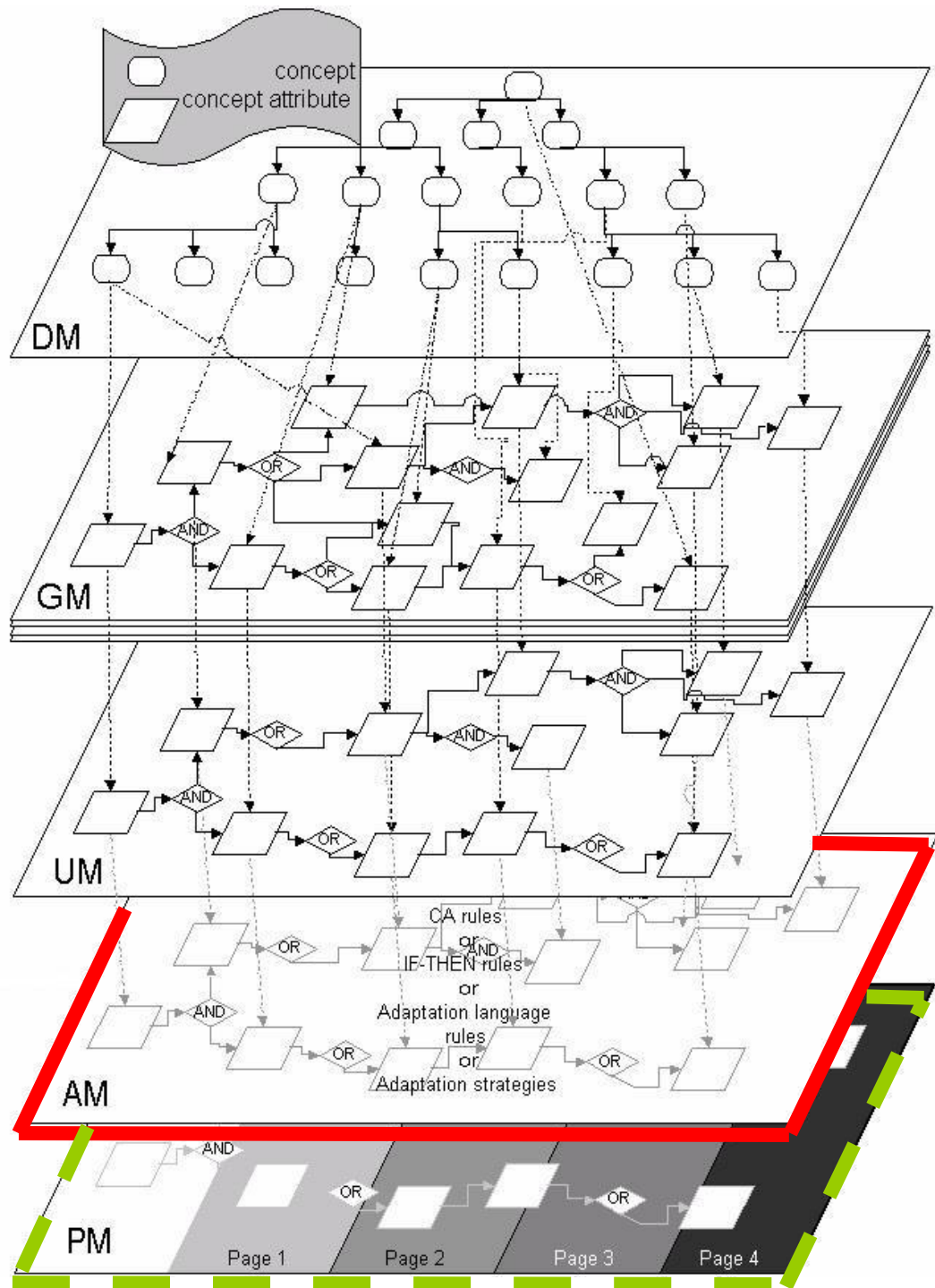


Figure 1-15 : Modèles d'adaptation (AM) et (PM) du framework LAOS

Le modèle d'adaptation (AM)

Le modèle d'adaptation se compose de trois couches¹⁷ (Figure 1-16). Cette section présente brièvement chacune d'elles.

<p>Première couche : technique/règles d'adaptation directe</p> <ul style="list-style-type: none"> • Supporte l'adaptation de la navigation et de la présentation • Techniques généralement basées sur le test d'un seuil calculé par combinaison de plusieurs couples attribut-valeur
<p>Deuxième couche : Langage de programmation d'adaptation</p> <ul style="list-style-type: none"> • Supporte l'adaptation par objectifs • Technique basée sur un langage de haut niveau manipulant des règles de la première couche
<p>Troisième couche : Stratégies d'adaptation</p> <ul style="list-style-type: none"> • Supporte l'adaptation par objectifs • Technique qui enveloppe les deux couches précédentes en permettant la création de fonctions

Figure 1-16 : Les trois couches de AM

1. Technique/règles d'adaptation directe :

Cette couche concerne les techniques traditionnelles d'adaptation des systèmes hypermédias adaptatifs [Brusilovsky, 2001].

- adaptation des contenus :
 - insertion/enlèvement des fragments d'écran,
 - ordonnancement des fragments d'écran,
 - adaptation de la présentation des fragments.
- adaptation des liens
 - ordonnancement des liens,
 - activer/désactiver/cacher des liens
 - génération de liens

Pour illustrer le modèle d'adaptation du framework LAOS, les auteurs fournissent l'exemple suivant. Considérons qu'on désire gérer la présentation des 7 sous-concepts du concept « NN Introduction » exprimé dans le GM de LAOS (Figure 1-17).

¹⁷ LAG : Layered Adaptation Granulation

NN-intro

Author: vadian

[\[refresh\]](#) [\[home\]](#) [\[student view\]](#)

- (1) **[root]** (AND)
 - (1) **title** (root)
 - (2) **[NN Introduction]** (AND)
 - (1) **Title** (NN Introduction)
 - (2) **Keywords** (Von Neuman architecture; infor...)
 - (3) **text** (Artificial neural networks con...)
 - (4) **[How Artificial Neural Networks came into being]** (AND)
 - (1) **title** (How Artificial Neural Networks...)
 - (5) **[The biological neural network]** (AND)
 - (1) **Title** (The biological neural network)
 - (2) **Keywords** (biological neural network;)
 - (3) **text** (Structurally speaking, BNN's a...)
 - (6) **[The Von Neuman computer versus the human brain]** (OR)
 - (1) (10%, least imp) **Title** (The Von Neuman computer versus...)
 - (2) (40%, somehow imp) **keywords** (number of elements; power cons...)
 - (3) (50%, most important) **text** (A comparison is made on two ac...)
 - (7) **[The biological neuron]** (AND)

Figure 1-17 : Concepts du modèle GM de LAOS

En utilisant l'écriture standard des règles d'adaptation dans cette couche, il faut écrire sept lignes de structures alternatives (Figure 1-18) :

```

IF NN Introduction.access == 'yes' THEN
    NN Introduction.Title.available = 'yes';
IF NN Introduction.access == 'yes' THEN
    NN Introduction.Keywords.available = 'yes';
IF NN Introduction.access == 'yes' THEN
    NN Introduction.text.available = 'yes';
IF NN Introduction.access == 'yes' THEN
    NN Introduction.How Neural ....available = 'yes';
IF NN Introduction.access == 'yes' THEN
    NN Introduction.The biological neural....available = 'yes';
IF NN Introduction.access == 'yes' THEN
    NN Introduction.The von Neuman....available = 'yes';
IF NN Introduction.access == 'yes' THEN
    NN Introduction.The biological neuron.available = 'yes';
  
```

Figure 1-18 : exemple de règles d'adaptation directes

2. Langage de programmation d'adaptation :

Cette couche consiste à regrouper les éléments de la couche précédente pour construire des modèles typiques d'adaptation (règles d'adaptation de haut niveau). Pour cela un langage de programmation pour fabriquer ces modèles de stratégies d'adaptation a été créé [Cristea et al, 2003a]. Il se nomme « langage adaptatif¹⁸ ». Il est à la base des modèles d'adaptation élaborés dans le cadre du projet Minerva ADAPT [ADAPT, 2002]. Les éléments utilisés sont : **while-do**, **for-do**, **generalize** et **specialize**.

L'exemple précédent devient (Figure 1-19) :

```

IF NN Introduction.access == 'yes' THEN (
    UM.NN Introduction.i = 1;
    FOR 7 DO (
        NN Introduction[i].available = 'yes';
        UM.NN Introduction.i += 1
    )
)

```

Figure 1-19 : Structure FOR- DO

Le système MOT, offre une interface de création de ces différents modèles d'adaptation (Figure 1-20).

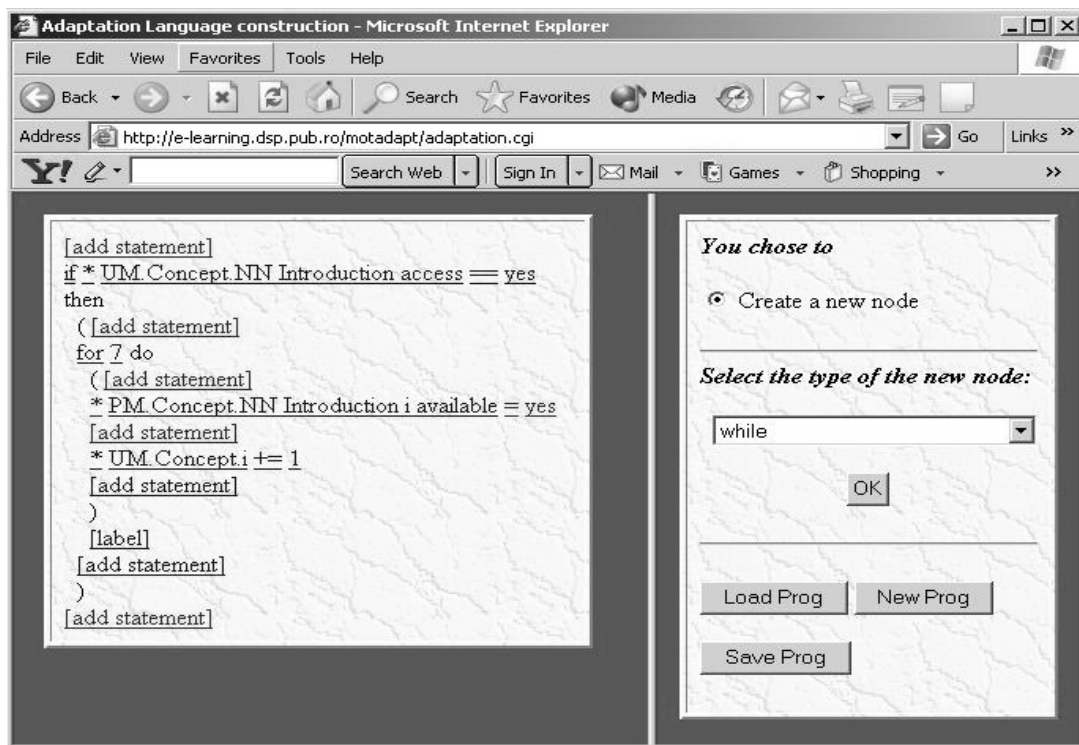


Figure 1-20 : Le modèle d'adaptation de LAOS dans MOT

¹⁸ adaptive language

3. Stratégies d'adaptation :

Cette couche consiste simplement à utiliser les structures des précédentes couches (Structure alternative et structures itératives) dans un niveau plus élevé de programmation. Les auteurs ont ajouté un composant de plus haut niveau appelé : *stratégie adaptative*¹⁹.

Ce composant a été utilisé pour la création de stratégies d'adaptation pédagogique basées sur le style d'apprentissage de l'élève [Bajraktarevic et al, 2003; Cristea et al, 2002].

Le principal inconvénient de ce modèle est qu'il est écrit dans des langages propriétaires. Son utilisation par d'autres systèmes est de fait largement compromise.

Le modèle de présentation (PM)

Le modèle de présentation prend en considération les propriétés physiques et environnementales de la présentation. Il fournit un pont pour l'actuel composant de génération de code pour différentes plateformes (ex : HTML, SMIL [SMIL W3C]). Le modèle de présentation dans le framework LAOS est similaire au modèle de l'élève dans le sens où il est aussi un modèle de type différentiel sur les modèles conceptuel de contenu (DM) et d'objectifs et de contraintes (GM). Cette structure permet d'attacher à chaque concept spécifique certains types de représentation pour l'écran, tandis qu'en même temps elle offre la possibilité de définir des types de représentation qui dépendent des valeurs courantes du modèle de l'élève.

Ce modèle n'offre pas la possibilité d'exprimer un niveau d'incertitude dans le choix d'une représentation. Il n'est pas possible aussi de faire évoluer ce choix au cours du temps en fonction d'évènements, comme la performance de l'élève à des activités d'évaluation formatives ou sommatives.

1.4.4. Incertitude dans le choix de l'adaptation

Dans les AEHS, la tâche qui consiste à choisir un type de présentation pour un élève ou à déterminer la prochaine activité pédagogique à proposer à l'élève est critique. Cette difficulté est principalement liée à l'incertitude qui plane sur la validité des critères qui serviraient à faire le choix [Mayo et al, 2001].

Dans [Zukerman et al, 2001a], les auteurs listent les techniques utilisées pour développer des modèles statistiques prédictifs d'élève. Les modèles probabilistes obtenus permettent de manipuler l'incertitude dans l'évaluation :

- des plans des élèves [Albrecht et al, 1998],
- des objectifs des élèves [Carberry, 2001],
- de la connaissance des élèves [Conati et al, 2002].

Küper et Kobsa [Kupper et al, 2003] présentent l'approche de génération de plans pour atteindre des objectifs en relation avec les capacités des élèves. Dans [Millan et al, 2002], les auteurs mesurent ces capacités avec différents niveaux de granularité en utilisant les réseaux bayésiens.

¹⁹ adaptive procedure or adaptive strategy

Dans [Castillo et al, 2004], les auteurs confirment cette idée en précisant que l'adaptation de ressources pédagogiques au style d'apprentissage de l'élève est extrêmement difficile, notamment à cause de l'incertitude des informations recueillies sur l'élève par les instruments psychométriques disponibles. Castillo préconise l'utilisation de réseaux bayésiens pour répondre à cette problématique. Il a conçu le système GIAS²⁰ dont la fonction d'adaptation a pour principal objectif d'aider l'élève dans l'exploration de nombreuses ressources pédagogiques reliées à une tâche d'apprentissage. Cette aide consiste à recommander à l'élève les ressources qui sont les plus adaptées à son style.

Les techniques d'adaptation mises en œuvre se focalisent sur l'optimisation de la sélection du sujet du cours et des ressources pédagogiques. Le système se base sur les objectifs de l'élève, son niveau de connaissances, son style d'apprentissage et ses préférences. Il exploite pour cela deux modèles : *le modèle de contenus* et *le modèle de l'élève*.

- *Le modèle de contenus* est organisé en trois couches (Figure 1-21) : *les objectifs, les sujets* et *les ressources*. Chaque objectif est associé à un ensemble de sujets qui sont eux mêmes associés à un ensemble de ressources pédagogiques. Il est possible de définir entre les sujets des relations d'agrégation.

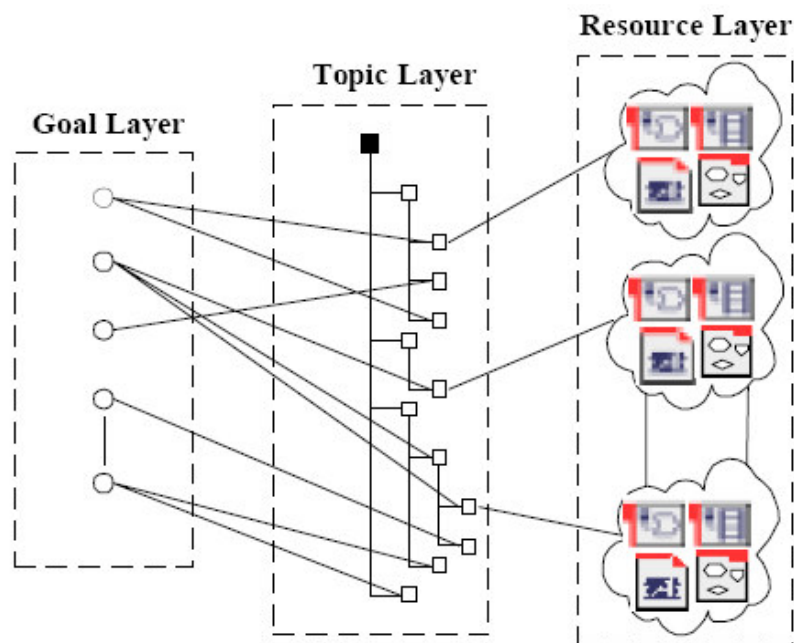


Figure 1-21 : Modèle de contenu du système GIAS

Les ressources pédagogiques existent en différents formats multimédia (texte, vidéo, animation, etc.) pour offrir une adaptation de l'activité pédagogique au style d'apprentissage de l'élève.

²⁰ Geometry Intelligent Adaptive System

- *Le modèle de l'élève* se compose de quatre modèles :

- **Profile Model** : il maintient les informations personnelles sur les élèves (nom, âge, style d'apprentissage).
- **Cognitive Overlay** : il enregistre les croyances du système sur la connaissance de l'élève sur les concepts définis dans le modèle de contenus. De ce fait, l'élève reçoit une qualification pour chaque concept : *novice, intermédiaire, expert*.
- **Predictive Model** : il enregistre les préférences de l'élève au sujet des ressources pédagogiques.
- **Course Overlay** : stocke les informations obtenues par le système sur les interactions de l'élève avec le cours (ex : nombre de visites par sujet, ou ressource pédagogique, performance obtenue à une évaluation sommative, etc.)

Le modèle de style d'apprentissage utilisé est celui créé par Felder et Sylverman [Felder et al, 1988]. L'acquisition du modèle se fait de façon explicite en début de session. L'élève répond au questionnaire « index of learning styles » créé par Felder et Soloman pour que le système évalue ses préférences.

Lorsque l'élève désire accéder au contenu d'un sujet, le système utilise le modèle prédictif pour classer les ressources accessibles. Le choix des ressources pour un sujet consiste à corrélérer les caractéristiques de la ressource avec le niveau de connaissances, le style d'apprentissage et les préférences de l'élève.

Ce processus est composé des étapes suivantes (Figure 1-24) :

1. **Filtrage** : les ressources sont filtrées en prenant comme critère le niveau de difficulté de la ressource (modèle de contenu) et le niveau de connaissances de l'élève (modèle de l'élève). Pour un élève caractérisé de « **novice** », le système choisira les ressources ayant un niveau de difficulté « **bas** ».
2. **Prédiction** (Figure 1-22) : par le modèle prédictif, chaque ressource filtrée est classée comme « appropriée » ou « non appropriée » pour l'élève en fonction de son style d'apprentissage (modèle de l'élève) et des caractéristiques de la ressource (modèle de contenu) (Figure 1-23).

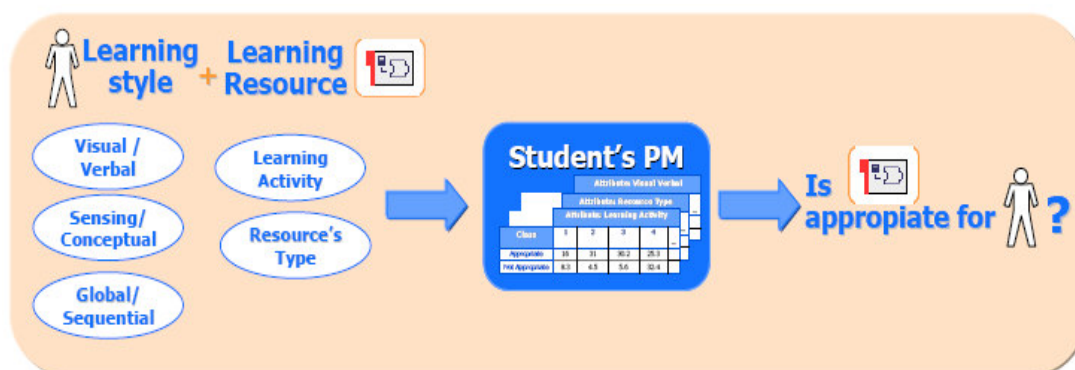


Figure 1-22 : Etape de prédiction

Attributes	Values
Characterizing the student's learning style	
<i>VisualVerbal</i>	$VVi, VV \in \{\text{Visual, Verbal}\}, i \in \{\text{mild, moderate, strong}\}$
<i>SensingConceptual</i>	$SCi, SC \in \{\text{Sensing, Conceptual}\}, i \in \{\text{mild, moderate, strong}\}$
<i>GlobalSequential</i>	$GSi, GS \in \{\text{Global, Sequential}\}, i \in \{\text{mild, moderate, strong}\}$
Characterizing the learning resource	
<i>Learning Activity (LA)</i>	<i>Lesson objectives/Explanation/Example/Conceptual Map/Synthesis Diagram/ Glossary / Summary /Bibliography /HistoricalReview /Inter.Activity</i>
<i>Resource Type (RT)</i>	<i>Text/HTML Text/Picture/Animated Picture/ Animated Picture with Voice/ Audio /Video /Software</i>

Figure 1-23 : Attributs et leurs valeurs

Les auteurs utilisent pour cette étape un classificateur probabiliste (Naïve Bayes Classifier). En guise d'illustration, la prédiction peut consister à estimer si une ressource implémentant une activité pédagogique de type « **Exemple** » avec un support multimédia de type « **Texte** » (LA = exemple, RT = Texte) est adaptée à un élève dont le style d'apprentissage est (VV=Visuel, SC=conceptuel, GS=Séquentiel).

Les compteurs du classificateur probabiliste sont mis à jour à chaque fois que l'élève parcourt un des liens d'une des deux listes de ressources pour le sujet choisi. Les probabilités sont alors mise à jour en conséquence.

3. **Décision** : une page HTML est envoyée à l'élève incluant les deux listes de ressources :
 - « *ressources conseillées pour l'étude* » : liste des ressources classées comme « appropriée »
 - « *autres ressources pour l'étude* » : liste des ressources classées comme « non appropriée »
4. **Adaptation** : toutes les fois qu'un nouvel exemple est observé et évalué, le modèle prédictif est adapté en conséquence.

L'adaptation est principalement basée sur les trois caractéristiques suivantes : la performance réalisée par l'élève à un test, la croyance sur le niveau de connaissances de l'élève et l'estimation de son style d'apprentissage. Le principal problème dans ce système est qu'il est difficile de trouver la cause d'un éventuel échec lors d'une évaluation : est-ce à cause d'une mauvaise croyance sur son niveau de connaissance ou bien sur son style d'apprentissage ?

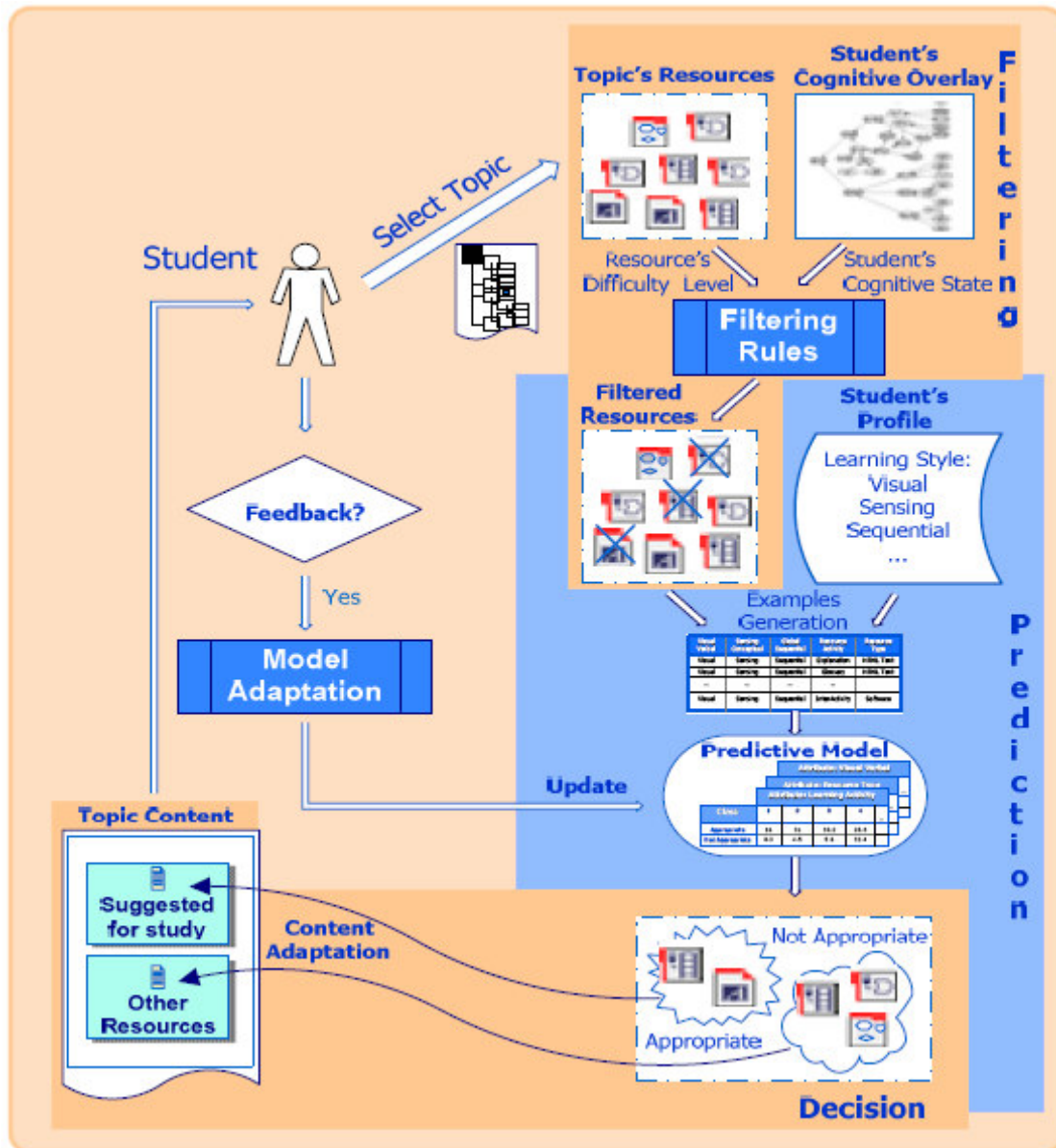


Figure 1-24 : processus de génération des sujets

1.4.5. Discussion

Les stratégies d'adaptation rencontrées visent essentiellement à adapter la présentation et la navigation. Peu de systèmes tentent d'y ajouter une adaptation de la pédagogie.

Dans cette thèse, nous proposons une stratégie d'adaptation de la présentation et de la navigation, mais aussi de la démarche pédagogique au style d'apprentissage. Cette adaptation est faite sur les activités pédagogiques ainsi que sur les objets d'apprentissage qui composent la séquence pédagogique.

1.5. Style d'apprentissage

La personnalisation de l'apprentissage humain remonte aux travaux sur les interactions entre aptitudes et traitements qui prônent l'adaptation de l'instruction aux caractéristiques de l'individu [Cronbach et Snow, 1977]. A cette fin, de nombreux travaux de recherche se sont focalisés sur l'identification des dimensions des différences individuelles [Snow, 1989]. Ces recherches ont mené à la naissance de la théorie du style d'apprentissage. Ce style est défini comme étant l'ensemble des processus mentaux qu'utilise l'individu pour percevoir et traiter d'une manière optimale l'information. La personnalisation implique alors, l'estimation des différences dans les styles d'apprentissage, les approches pour apprendre et les niveaux de développement intellectuels des élèves [Felder et al, 2005a]. Cette approche peut être utilisée dans les applications éducatives multimédia pour présenter différemment l'information basée sur les caractéristiques individuelles des apprenants [Barker, 2002] telles que les objectifs, les préférences et la connaissance [Brusilovsky, 2001] maintenues dans un modèle de l'élève multidimensionnel [Zukerman et al, 2001b].

1.5.1. Différentes approches et modèles

Les travaux sur le style d'apprentissage ont donné de multiples méthodes et instruments permettant de catégoriser les apprenants en fonction de leurs différences : Learning Styles Theory [Kolb, 1985], Index of Learning Styles [Felder et al, 1988], Learning Styles [Honey et Mumford, 1992], Student Learning Style Scales [Grasha, 1996], Multiple Intelligences [Gardner, 1999], Auditory Visual Tactile Learning Styles [Sarasin, 1998].

Riding a esquissé une classification des principaux modèles de styles d'apprentissage [Riding et al, 2001] qu'il a classé en trois groupes :

- Modèles axés sur le *processus d'apprentissage*,
- Modèles axés sur les *préférences pédagogiques*,
- Modèles axés sur les *habiletés cognitives* ainsi que les *stratégies d'apprentissage*.

Cependant, il démontre que tous ces modèles ne proposent pas un composant mental de base (construct). Pour lui, ce composant de base, non évolutif, existe et s'appelle le style cognitif. Celui-ci est à la base de toutes les activités cognitives humaines y compris l'apprentissage. Riding soutient que l'individu développe des stratégies d'apprentissage qui reflètent les processus sous-jacents à son style. Le modèle qu'il propose se fonde sur les habiletés cognitives et représente l'approche préférée et habituelle d'un individu pour organiser et se représenter l'information. Ce modèle est composé uniquement de deux dimensions :

- *wholist - analytic* : modélise la tendance d'un individu à organiser l'information comme un tout ou bien comme un ensemble de parties,
- *verbaliser - imager* : modélise la tendance d'un individu à se représenter l'information pendant la phase de réflexion sous forme verbale ou bien sous forme d'images mentales.

Riding propose une méthode²¹ pour évaluer la position d'un individu sur les deux dimensions du modèle.

Felder R. considère que l'ensemble des différences individuelles qui comprennent les préférences personnelles indiquées pour l'apprentissage et certaines activités pédagogiques, mais aussi les différences intellectuelles et psychologiques individuelles, font référence au style d'apprentissage d'un individu [Felder et al, 1988]. Pour lui, le style d'apprentissage d'un élève doit être considéré en termes de présentation, organisation, traitement et assimilation d'information [Felder, 1993].

Il propose un modèle composé de quatre dimensions en émettant l'hypothèse que le style d'apprentissage d'un élève peut être défini en répondant aux quatre questions suivantes [Felder, 1993] :

- *déductif - inductif* : avec quelle organisation de l'information l'élève est-il le plus à l'aise ?
- *visuel - verbal* : à travers quelle modalité sensorielle l'information est-elle réellement perçue ?
- *séquentiel - global* : comment l'élève progresse-t-il dans l'activité de compréhension ?
- *actif - réfléchi* : comment l'élève préfère-t-il traiter l'information ?

Felder propose un questionnaire²² composé de 44 questions (Annexe 1) pour évaluer la position d'un élève sur les quatre dimensions.

1.5.2. Style d'apprentissage et performances de l'élève

Depuis l'apparition des modèles sur le style d'apprentissage [Riding et al, 2001], plusieurs études ont été menées afin de montrer si l'adéquation de la séquence pédagogique au style d'apprentissage améliorerait la performance de l'élève.

Les résultats semblent être mitigés [Redmond et al, 2003 ; Miller, 2004], mais la majorité des auteurs s'accordent à dire que la prise en compte du style d'apprentissage dans la conception d'une séquence pédagogique ne peut qu'être un facteur positif pour le soutien de l'élève dans l'activité qui lui est proposée [McLoughlin, 1999]. Par exemple, on a constaté que l'intérêt porté par un élève pour suivre une activité pédagogique, augmentait si cette dernière était en accord avec son style d'apprentissage [Rumetshofer & al, 2003] notamment en incorporant dans les contenus des composants multimédia [McLoughlin & al, 2002].

De même, la dernière version de ELM-ART, système adaptatif d'apprentissage par le Web du langage de programmation LISP [Brusilovsky, 2004a] a montré que l'adaptation de la navigation avait une incidence directe sur la motivation des élèves.

Néanmoins, Triantafillou [Triantafillou et al, 2003] précise que la majorité des élèves, s'ils apprécient d'un côté que le système cherche à adapter la séquence pédagogique à leur style d'apprentissage, aiment malgré tout d'un autre côté pouvoir garder la maîtrise sur le fonctionnement du système. Ils apprécient de pouvoir sélectionner un lien de navigation même si le système ne le conseille pas.

²¹ Cognitive Styles Analysis

²² Index of learning style

Riding a montré que dans certaines situations [Riding et al, 2001], la performance d'un élève était affectée par une interaction entre son style d'apprentissage et les facteurs suivants :

- la structure utilisée pour le matériel pédagogique (position du titre, présence de sous titre, ajout de présentations générales et leur position),
- son mode de présentation (texte, image, multimédia),
- son type de contenu (concret, abstrait).

Une autre étude [Parkinson & al, 2002] conduit à la conclusion que les AIWBES ont un énorme potentiel à offrir aux élèves sur le plan pédagogique lorsque leur conception prend en compte le style d'apprentissage d'un individu. Certains auteurs pensent que la motivation d'un élève pour suivre une séquence pédagogique peut être augmentée si elle est conçue dans le but de répondre à plusieurs styles d'apprentissage [Zhang, 2004].

Néanmoins, la prise en compte du style d'apprentissage d'un individu dans le processus de formation n'est pas le seul facteur à considérer. Il a été aussi mis en évidence que le type de navigation, contrainte ou conseillée devait être corrélée au niveau de connaissances de l'élève pour obtenir une performance maximale dans la phase d'apprentissage [Brusilovsky, 2004a]. Un élève avec aucun pré-acquis préférera une séquence dirigée, alors qu'à l'inverse un élève avec certaines connaissances sur le domaine aura tendance à vouloir naviguer librement dans les contenus pédagogiques.

1.5.3. Style d'apprentissage et documents adaptatifs

L'adaptation aux styles d'apprentissage est une tâche qui doit être prise en compte lors de la phase de conception des séquences pédagogiques [Moallem, 2001] ainsi que lors de leur enchaînement durant la phase d'enseignement à proprement dite. On distingue donc la création de contenus multimédia adaptatifs et leur séquençement et présentation adaptatifs. Plusieurs travaux ont porté sur ces deux aspects séparément. Cependant, il y en a peu qui traite le problème d'une manière globale.

Hui Min Lee [Hui Min Lee et al, 2004] conclut dans son étude que la qualité de l'enseignement se trouve améliorée lorsque l'auteur de la séquence pédagogique cherche à exploiter les différentes possibilités offertes par les supports hypermédias (séquences vidéo, audio,...) afin de contenter les caractéristiques et préférences individuelles des élèves.

Mayer [Mayer, 2003] a mené plusieurs expériences qui ont montré par exemple l'importance que pouvait revêtir la position dans la page d'une légende associée à une image. Plus la distance qui les sépare est grande, plus il sera difficile pour l'élève de comprendre le contenu. De même, il a montré que l'ajout d'éléments multimédia non pertinent dans une page produisait un effet négatif sur l'apprentissage de l'élève en perturbant son niveau de concentration.

La majeure partie des travaux concernant les documents multimédia adaptables aux styles d'apprentissage porte sur les médias à utiliser lors de l'adaptation pour la présentation. Très peu se consacre à la relation entre le style et la nature de l'activité et des tâches d'apprentissage. Moallem précise que la conception du cours doit aussi être réalisée en fonction des pédagogies que l'auteur vise [Moallem, 2001] et pas seulement sur des considérations liées aux médias utilisés comme support de l'information.

La tâche de créer des contenus adaptatifs est très délicate et fastidieuse du fait de la complexité des critères à prendre en considération. Les règles d'adaptation doivent être spécifiées pour chaque utilisation de chaque composant pédagogique, pour chaque situation particulière. Pour aider les auteurs dans leur entreprise, certaines recherches prônent la création de logiciels auteurs [Cristea, 2004a] tel que le système MOT qui offre des fonctions d'assemblage de composants pédagogiques en précisant pour chacun la stratégie d'adaptation à utiliser. L'adaptation peut être programmée par l'auteur du cours grâce à un langage d'adaptation [Cristea et al, 2004b] intégré au système. De même, eCAD est un système d'aide à l'analyse et la conception [Timothy et al, 2004] d'un cours formulé sur la base d'un ensemble d'objectifs pédagogiques respectant la taxinomie de Bloom [Bloom, 1956].

Les systèmes adoptant les styles d'apprentissage dans les deux phases de conception et enseignement sont rares. Dans [Cristea et al, 2003a], les auteurs présentent un modèle de document à trois couches qui offrent une adaptation de la navigation ainsi que de la présentation en considérant le style d'apprentissage. Castillo [Castillo et al, 2003] montre un début d'intérêt pour les réseaux bayésiens dans la modélisation de prédiction de tâches, en se basant sur les résultats des travaux de Felder [Felder, 1993].

Dans cette thèse, nous tentons de traiter cette problématique en proposant un modèle de dépendances statistiques entre le style d'apprentissage et les métadonnées décrivant les contenus pédagogiques. Les contenus sont indexés en termes de capacité à convenir aux différents styles.

1.6. Web2.0 : Modèles de l'élève et de contenus

Web2.0 est un cadre conceptuel pour l'évolution du web. Il considère deux dimensions : le Web sémantique et le Web social [Gandon, 2006]. L'objectif du Web sémantique est de rendre les ressources du Web accessibles et utilisables par des programmes et agents logiciels, grâce à des métadonnées formelles définies explicitement par les auteurs [Brooke et al., 2005]. L'idée est de rendre possible des inférences allant bien plus loin qu'une simple analyse linguistique [Alesso et al, 2004] en utilisant des représentations formelles standards comme RDF, RDFS et OWL [Shen et al, 2006]. Le Web sémantique est basé sur l'utilisation d'ontologies. Une ontologie est une description de concepts clés dans un domaine incluant les règles, propriétés et relations existants entre ces concepts [Alani et al, 2003].

Le Web social est le nom donné à la partie des activités en ligne qui requièrent une participation collaborative des utilisateurs [Foell et al, 2007]. Dans ce contexte, les outils, qui offrent une génération coopérative de contenus, sont appelés des logiciels sociaux [Tepper, 2003]. Ces outils ont en commun une structure sociale égalitaire ; ils agrègent les contributions de tous les utilisateurs dans des représentations partagées d'une croyance collective [Rodriguez et al, 2007]. Certains de ces outils, exploitent des modèles de connaissances, afin par exemple de réaliser une classification des utilisateurs. Dans [Bermudez et al, 2006], les auteurs préconisent l'utilisation d'une ontologie pour permettre la réutilisation et l'extension de modèles décrivant le profil d'une communauté d'utilisateurs.

Dans cette partie, nous commençons par présenter différents travaux réalisés dans le domaine des ontologies, avant de montrer leurs utilisations dans des systèmes d'enseignement. Nous terminons par l'étude de quelques systèmes qui exploitent les ontologies pour modéliser le profil des élèves.

1.6.1. Les ontologies

Une ontologie définit les termes servant à décrire et représenter un champ de connaissances. Il ne suffit pas de s'entendre sur la façon de nommer les choses mais aussi sur le sens qu'on accorde aux mots qu'on utilise. Pour ce faire, les ontologies sont utilisées par les personnes, les bases de données et les applications qui ont besoin de partager des informations de domaine. Un domaine étant un sujet ou un champ de connaissances particuliers, comme la médecine, la fabrication d'outils, l'immobilier, la réparation automobile, la gestion financière, etc. Les ontologies incluent les définitions, utilisables par l'ordinateur, des concepts élémentaires du domaine et leurs relations [W3C Ontology Web].

Dans [Daconta et al, 2003], les auteurs précisent qu'on trouve différents niveaux de sémantique dans les ontologies (Figure 1-25). Elles portent alors différents noms :

- Taxinomie,
- Thésaurus,
- Modèle conceptuel,
- Théorie locale de domaine.

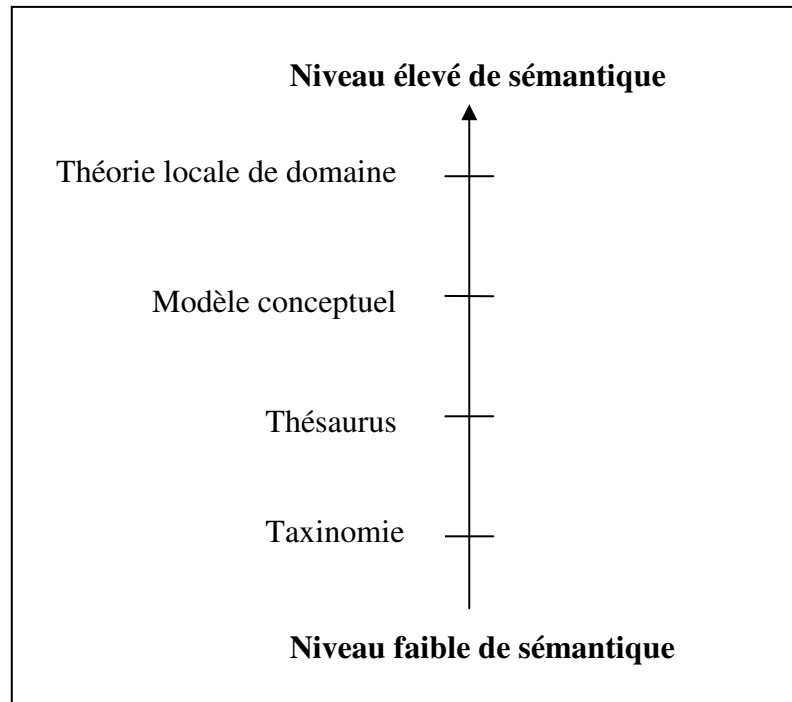


Figure 1-25 : différents niveaux sémantiques dans les ontologies

Les taxinomies contiennent des données structurées pour lesquelles la sémantique des relations est mal définie (ex : généralement on trouve comme relations est-une-sous-classe-de ou est-une-partie-de). Les thésaurus quand à eux, sont constitués de vocabulaire contrôlé ayant entre eux des relations sémantiques et génériques (ex : équivalence, hiérarchique et associative). Les modèles conceptuels permettent la création d'une hiérarchie de classes et la définition de propriétés (comme avec les langages UML²³ ou OWL²⁴). Les théories locales de domaine utilisent la logique modale ou de 1^{er} ordre et sont directement interprétables sémantiquement. Elles représentent le plus haut niveau de sémantique pour les ontologies.

Leur relative complexité, rend difficilement exploitable les ontologies par un être humain. Des travaux récents [Yuxin et al, 2006] dans ce domaine montrent qu'il est possible de visualiser graphiquement une ontologie en utilisant comme méthode de modélisation les « concept map ». Les auteurs proposent pour ce faire une conversion des éléments du langage OWL vers les entités manipulées par les « concept map »(Tableau 1-2).

²³ Unified Modeling Language

²⁴ Ontology Web Language

OWL ontology elements	Concept map units	Graphical Types
Class	Concept	Node
Property	Relation	Arc
Domain class	Concept	Source node
Range class	Concept	Target node
Instance	Specific examples	Node
Statement	Proposition	Node

Tableau 1-2 : correspondance entre éléments OWL et entités concept map

Une ontologie concerne les concepts, le vocabulaire, la taxinomie, les relations entre concepts et les inférences sur les concepts d'un domaine. Elle permet donc de préciser formellement un vocabulaire commun dans une perspective de partage. D'importants travaux ont été réalisés sur les ontologies par le groupe de travail WebOnt²⁵ du W3C. L'enjeu est de rendre le Web « interprétable » par des agents logiciels. On parle alors de Web sémantique.

Le Web sémantique²⁶ consiste à marquer manuellement ou automatiquement les informations véhiculées sur le Web avec des annotations sémantiquement riches qui peuvent être interprétées automatiquement par des agents ou d'autres programmes [Cristea, 2004e]. Cette interprétation repose sur la manipulation d'ontologies [Mizoguchi, 2004].

La pile du Web sémantique (Figure 1-26) a été proposée par Berners-Lee [Berners-Lee, 2003]. Elle montre que le Web Sémantique est construit sur la capacité de XML de définir des schémas de balisage personnalisés et sur la flexibilité de l'approche RDF pour représenter les données. Les agents logiciels devant faire des tâches de raisonnement sur les documents du Web, le langage de description doit aller au-delà des sémantiques de base de RDFS. OWL a été conçu pour répondre à ce besoin comme un langage d'ontologie pour le Web.

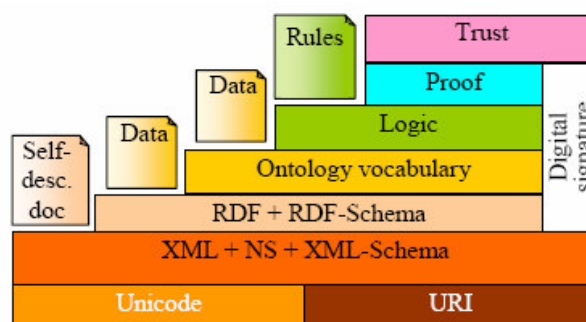


Figure 1-26 : pile du Web sémantique

²⁵ <http://www.w3c.org/2001/sw/WebOnt/>

²⁶ <http://www.w3.org/2001/sw/>

A la base de la pile du Web sémantique, on trouve les ressources, identifiées par leur identificateur unique (URI). XML est un langage utilisé pour créer une description riche d'un domaine particulier [Daconta et al, 2003] avec la notion d'espace de noms pour éviter les conflits de vocabulaire. Mais il ne fournit aucune contrainte sémantique pour créer une description. XML-S est un langage pour restreindre la structure des documents XML, mais aussi pour étendre XML avec des types de données. RDF est un modèle de données pour les objets (« ressources ») et les relations entre eux, fournissant des sémantiques simples pour ce modèle de données qui peuvent être représentées en XML. RDFS est un langage orienté objet qui permet la création du vocabulaire RDF. OWL ajoute plus de vocabulaire pour décrire les propriétés et les classes. Il est ainsi possible d'ajouter des relations entre les classes, de définir des cardinalités, d'avoir un typage de propriétés plus riche, etc.

Le niveau de sémantique augmente dans la pile de bas en haut. Les ontologies sont construites à partir d'un vocabulaire explicite et structuré. En particulier les ontologies rendent la connaissance réutilisable par la création de classes (concept généraux), d'instances (concepts particuliers), de relations entre les concepts, de propriétés qui décrivent les concepts (avec leurs valeurs), de fonctions, de contraintes et de règles qui manipulent les concepts.

Des systèmes hypermédias adaptatifs récents tentent d'atteindre un niveau sémantique élevé avec des représentations RDF, RDFS et OWL. Parmi ceux-là citons : Personal Reader [Dolog et al, 2003], GEAHS [Jacquot et al, 2004], Hera [Frasincar et al, 2003], DLRS [Maneewatthana et al, 2004].

Le W3C [W3C Ontology Web] a identifié 6 cas d'utilisation représentatifs d'ontologies Web :

- Les collections multimédias
- L'administration d'un site web d'entreprise
- La documentation d'un concept
- Les agents et services
- L'informatique omniprésente
- Les Portails Web

Les ontologies peuvent servir pour l'annotation sémantique des collections multimédias : images, sons ou autres objets non textuels [Piombo et al, 2007]. Il est encore plus difficile pour les machines d'extraire une sémantique significative d'objets multimédias que d'un texte en langue naturelle. Ainsi, ces types de ressources sont habituellement identifiés par des légendes ou des balises de métadonnées. Par contre, puisque les personnes décriraient différemment ces objets non textuels, il importe que les installations de recherche offrent plus qu'une simple correspondance de mots-clés. Idéalement, les ontologies devraient saisir une connaissance supplémentaire à propos du domaine pouvant servir à améliorer la récupération des images. Les ontologies multimédias peuvent être de deux types : spécifiques du média ou spécifiques du contenu.

Les ontologies spécifiques du média peuvent avoir des taxinomies de différents types de média, et décrire les propriétés des différents médias. Par exemple, une vidéo peut inclure des propriétés pour identifier la longueur du clip et les changements de plan.

Les ontologies spécifiques du contenu décrivent le sujet de la ressource, tels que la mise en scène ou les participants. Comme ces ontologies ne s'adressent pas à un média particulier, elles peuvent être réutilisées par d'autres documents portant sur le même domaine. Cette réutilisation améliorerait une recherche simple d'informations concernant un domaine particulier, indépendamment du format de la ressource. Les recherches, où le type du média est important, peuvent combiner les ontologies spécifiques du média et celles spécifiques du contenu. Comme exemple de collection multimédia, prenons une archive d'images de meubles anciens. Une ontologie du mobilier ancien serait très utile pour effectuer des recherches dans une telle archive. On peut utiliser une taxonomie pour classer les différents types de meubles. Il serait aussi utile que l'ontologie puisse exprimer une connaissance de définition. Par exemple, si un indexeur sélectionne la valeur *Georgien tardif* comme *style/période* d'un coffre ancien à tiroirs, on devrait pouvoir inférer que l'élément de donnée « *date.création* » a une valeur comprise entre 1760 et 1811, et que l'élément de donnée « *culture* » a pour valeur *britannique*. La disponibilité de ce type de connaissances de fond augmente beaucoup les possibilités de l'indexation comme de la recherche. Une autre caractéristique qui peut se révéler utile est la prise en charge de la représentation d'une connaissance implicite. Comme exemple de connaissances de ce type, un coffre à tiroirs du *Georgien tardif* qui sans information supplémentaire sera supposé fabriqué en *acajou*. Cette connaissance est cruciale pour les requêtes sémantiques réelles, par exemple, une requête d'utilisateur pour un meuble de rangement ancien en *acajou* pourrait correspondre aux images de coffres à tiroirs du *Georgien tardif*, même sans mention du bois dans l'annotation de l'image.

Les grandes entreprises ont habituellement beaucoup de pages Web à administrer pour leur site Web. Les contenus concernent des sujets comme les communiqués de presse, les offres de produits et d'études de cas, les procédures d'entreprise, les présentations et comparatifs de produits internes, les documents techniques et les descriptions de processus. Les ontologies peuvent servir à indexer ces documents et fournir de meilleures méthodes de récupération. Bien que nombre de grandes entreprises disposent d'une taxonomie d'organisation des informations, souvent cela ne suffit pas. Une seule ontologie est souvent limitative car les catégories constituantes sont vraisemblablement contraintes à celles représentant une seule vue et une seule granularité d'un domaine.

La possibilité de travailler simultanément avec plusieurs ontologies enrichirait les descriptions. En outre, la possibilité d'effectuer une recherche sur les valeurs de différents paramètres est souvent plus utile qu'une recherche par mots-clés avec des taxonomies. Un site Web utilisant des ontologies est susceptible d'intéresser :

- Un vendeur recherchant une garantie de disponibilité du produit concernant la poursuite d'une vente ;
- Un technicien recherchant des cas d'expertise technique particulière et des détails de l'expérience acquise ;
- Un chef de projet recherchant une expérience acquise et des modèles pour soutenir un projet complexe, pendant la phase de proposition comme pendant l'exécution.

Pour chacun de ces types d'utilisateur, le problème habituel est qu'ils ne peuvent pas partager de terminologie avec les auteurs des contenus demandés. Le vendeur peut ignorer le terme technique d'une caractéristique souhaitée, ou les techniciens dans différents domaines de compétence peuvent employer des termes différents pour le même concept. Pour ces problèmes, il serait utile que chaque classe d'utilisateurs dispose de différentes ontologies de termes, mais que les ontologies soient inter-reliées, de façon à pouvoir effectuer automatiquement des traductions.

Le cadrage des requêtes au bon niveau d'abstraction soulève un autre problème. Un chef de projet recherchant un expert en systèmes d'exploitation devrait, par exemple, pouvoir trouver un employé expert à la fois d'Unix et de Windows. Une grande entreprise de services peut offrir un large éventail de compétences. Mais pour réaliser des contrats importants, il faut parfois assembler ces savoir-faire autrement et innover. Souvent, aucun projet antérieur seul ne correspondra. Le défi tient dans le raisonnement sur la manière de réassembler les modèles et documents passés en configurations nouvelles, tout en satisfaisant à un ensemble de conditions préalables.

Un corps volumineux de documents techniques, comme ceux trouvés dans l'industrie aérospatiale peut être de plusieurs types, comprenant la documentation de conception, la documentation de fabrication et la documentation d'essai. Ces ensembles de documents ont chacun une structure hiérarchique différente d'un ensemble à l'autre. Il existe aussi un ensemble d'axes implicites reliant les ensembles de documentation. Par exemple, dans des documents de conception aéronautique, un élément « longeron d'aile » pourrait être présent dans chaque ensemble. Les ontologies peuvent servir à bâtir un modèle informationnel [Bateman et al, 2006] permettant l'exploration de l'espace d'informations selon les éléments représentés, les associations entre éléments, les propriétés des éléments et les liens vers la documentation qui les décrit et les définit (c'est-à-dire, la justification externe de l'existence de l'élément dans le modèle). Le projet Poncelet [Habel et al, 2006], qui vise l'enseignement de la géométrie projective par l'exploitation de ressources pédagogiques multimédia accessibles par le Web, s'inscrit dans cette logique. Il utilise une ontologie qui est considérée comme le référentiel sémantique de toutes les ressources utilisées dans le projet. Cette ontologie a été créée avec le logiciel MOTPlus en utilisant le langage OWL. Cela veut dire aussi que l'ontologie et la taxonomie ne sont pas indépendantes des éléments physiques qu'elles représentent mais qu'elles peuvent être développées et explorées en tandem. Un exemple concret de ce cas d'utilisation est celui de la documentation de conception dans le domaine aérospatial utilisée habituellement par :

- Les mécaniciens d'entretien à la recherche de toutes les informations relatives à une pièce particulière (par exemple, un longeron d'aile) ;
- Les ingénieurs concepteurs examinant les contraintes de réutilisation d'un sous-assemblage particulier.

Pour gérer ce type d'utilisation, il est important de pouvoir définir des contraintes. Celles-ci peuvent être utilisées pour améliorer la recherche ou pour vérifier la cohérence.

Une autre utilisation courante de ce type d'ontologie est de permettre la visualisation et l'édition de tableaux montrant des instantanés de l'espace d'informations centrés sur un concept particulier (par exemple, une classe ou une instance). Ce sont généralement des diagrammes d'activités-règles, ou des diagrammes entités-relations

Le Web sémantique peut offrir aux agents la capacité de comprendre et d'intégrer des ressources d'information diverses. Un exemple spécifique est celui d'un planificateur d'activités sociales qui utilisera ces données pour organiser la soirée de l'utilisateur selon ses préférences (quels types de films aime-t-il, quelle est sa cuisine favorite, etc.). La tâche de planification de cette soirée dépendra de la richesse des services offerts et des besoins de l'utilisateur. Pendant le processus de détermination/filtrage du service, des services de notation et de critiques peuvent également être consultés pour approcher au mieux les préférences de l'utilisateur (par exemple, les critiques et les notations de films ou de restaurants pour trouver le plus adapté). Ce type d'agent a besoin d'ontologies de domaines pour représenter les termes spécifiques aux restaurants, aux hôtels, etc., et d'ontologies de services pour représenter les termes employés par les services de notation et de critique. Ces ontologies permettront la saisie des informations nécessaires aux applications pour discriminer et évaluer les préférences de l'utilisateur. Ces informations peuvent être fournies par des sources diverses, tels que des portails, des sites à service spécifique, des sites de réservation, et du Web en général. Dans [Pasin et al, 2006], les auteurs exploitent ce principe en utilisant une ontologie, écrite avec le langage OWL, comme filtre de recherche sur le Web pour trouver des ressources pédagogiques liées au domaine de connaissances abordé.

L'informatique omniprésente est un paradigme émergent de l'informatique personnelle, caractérisé par un glissement d'une machinerie informatique dédiée vers des capacités informatiques envahissantes intégrées dans nos environnements quotidiens. Les appareils miniaturisés, portables et sans fil en sont des exemples. La prolifération et la nature sans fil des appareils imposent aux architectures de réseaux de prendre en charge une configuration automatique ad hoc. Une autre raison au développement d'une configuration automatique tient au fait que cette technologie vise les consommateurs ordinaires. Une technologie-clé de réseaux ad hoc est la découverte de service, une fonctionnalité au travers de laquelle des services (les fonctions offertes par divers appareils tels que les téléphones portables, les imprimantes, les capteurs, etc.) peuvent être décrits, annoncés et détectés par d'autres. Tous les mécanismes actuels de découverte de service et de description de capacité (JINI Sun, UPnP Microsoft) se fondent sur des systèmes de représentation ad hoc et s'appuient fortement sur une normalisation. C'est-à-dire, sur une identification a priori de toutes les choses susceptibles d'être communiquées ou discutées.

La question-clé et le but de l'informatique omniprésente est l'interopérabilité fortuite, une interopérabilité dans des conditions non chorégraphiées, c'est-à-dire que les appareils non conçus pour fonctionner obligatoirement ensemble (tels ceux construits pour différents usages, par des fabricants différents, à des époques différentes, etc.) devraient être capables de détecter leurs caractéristiques mutuelles et d'en tirer profit. Pouvoir comprendre les autres appareils et raisonner sur leurs services/fonctionnalités est nécessaire, puisque les scénarios à terme de l'informatique omniprésente mettent en scène des dizaines voire des centaines d'appareils, et qu'une normalisation a priori des scénarios d'utilisation n'est pas concevable.

Les scénarios d'interopération sont par nature dynamiques (c'est-à-dire que les appareils apparaissent et disparaissent à n'importe quel moment, car leurs propriétaires les transportent d'une pièce ou d'un bâtiment à un autre) et ne donnent lieu à aucune intervention humaine pour la (re)configuration. L'utilisation des services implique des phases de découverte, de passage de contrat et de composition. Le passage de contrat de services peut impliquer une représentation d'informations concernant la sécurité, la vie privée et la confiance, ainsi que les détails d'une compensation (un fournisseur de services peut recevoir une rémunération pour les services rendus). En particulier, un but est d'exprimer les politiques de sécurité d'entreprises ou d'organisations dans une forme neutre pour les applications, qui permet ainsi une représentation des contraintes entre les divers mécanismes d'exécution (par exemple, les pare-feux, les filtres/scanners, les dispositifs de surveillance du trafic, les routeurs au niveau de l'application et les répartisseurs de charge). On utilisera donc un langage d'ontologie pour décrire les caractéristiques des appareils, les méthodes pour accéder à ces appareils, la politique d'utilisation de l'appareil établie par son propriétaire et les autres contraintes et conditions techniques qui affectent l'incorporation de l'appareil dans un réseau de type informatique omniprésente.

Un portail Web est un site Web qui offre un contenu d'informations concernant un thème commun, par exemple une ville ou un domaine d'intérêt particulier. Un portail Web permet aux personnes intéressées par ce thème de recevoir des nouvelles, de trouver et de parler à quelqu'un, de bâtir une communauté et de trouver des liens vers d'autres ressources Web d'intérêt commun. Toutefois, un simple index de domaines sera peut-être insuffisant pour rechercher le contenu demandé par les membres de la communauté. Pour un meilleur partage de l'information, le portail Web peut définir une ontologie pour la communauté. Cette ontologie peut fournir une terminologie pour décrire le contenu et des axiomes pour définir des termes d'après les autres termes de l'ontologie. Par exemple, une ontologie pourrait inclure une terminologie avec les termes journal, publication, personne et auteur. Elle pourrait ensuite inclure des définitions telles que tous les journaux sont des publications, ou les auteurs de toutes les publications sont des personnes. Combinées à des faits, ces définitions permettent d'inférer d'autres faits forcément vrais. À leur tour, ces inférences permettent aux utilisateurs d'obtenir des résultats de recherche auprès du portail impossibles à obtenir avec des systèmes de repérage conventionnels. Cette technique repose sur les fournisseurs d'information utilisant le langage d'ontologie Web pour saisir des relations d'ontologies de qualité. C'est un autre défi que de minimiser l'effort de créer les ontologies de qualité. Un langage d'ontologie aura vraisemblablement un impact plus grand s'il facilite la saisie non intrusive des métadonnées au cours du processus de création de l'information.

[Ardissono et al, 2005] propose la création d'une méta-ontologie, classification de haut niveau, des ontologies utilisées dans le domaine de la modélisation des utilisateurs (Figure 1-27).

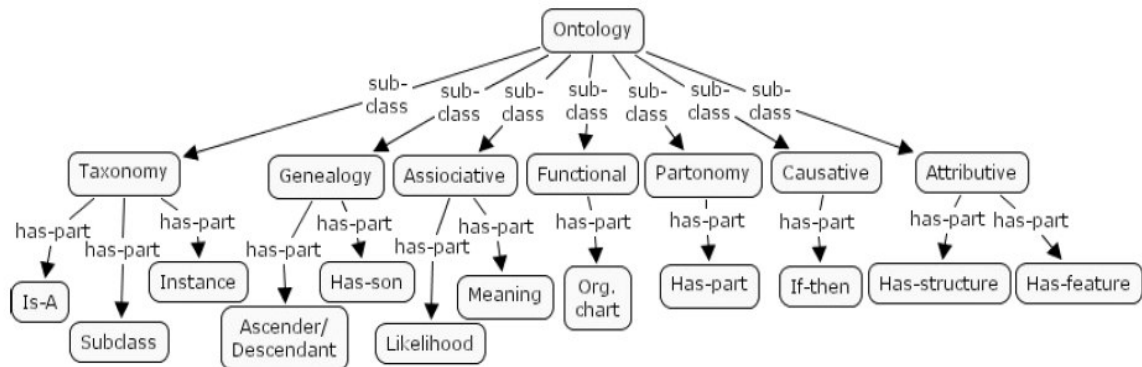


Figure 1-27 : classification des ontologies dans le domaine de la modélisation des utilisateurs

Le rôle de cette Méta-Ontologie sur la modélisation des utilisateurs (UMMO-User Modeling Meta-Ontology) est multiple. En premier, elle peut servir à présenter dans une vue synthétique l'état de l'art sur les ontologies dans ce domaine. En second, elle peut servir à structurer les différentes approches, articles et systèmes dans le domaine de la modélisation de l'utilisateur en les indexant par les termes de cette méta-ontologie pour être capable ensuite de les comparer ou bien de les relier entre eux. Enfin, les auteurs proposent de l'utiliser comme base du développement d'un système hypermédia adaptatif d'enseignement en considérant que ces systèmes se fondent aujourd'hui sur l'indexation avancée de concepts et que donc le développement d'ontologie devrait être la première étape à réaliser pour concevoir ce genre de système.

Au regard du nombre croissant d'ontologies, certains auteurs orientent leurs recherches vers la conception d'une approche pouvant être utilisée comme outil d'évaluation et de validation d'une ontologie [Gangemi et al, 2006]. D'autres répondent au besoin d'enrichir l'ontologie de manière semi-automatique, en exploitant la richesse du vocabulaire d'une certaine langue cible pour compléter les termes figurant dans l'ontologie d'origine [Pazienza et al, 2006].

[Berners-Lee, 2001] explique qu'une ontologie typique pour le Web se compose d'une taxinomie sur les concepts du domaine traité et d'un ensemble de règles d'inférences sur les concepts précités. D'autres auteurs s'intéressent à l'analyse de l'usage du web pour découvrir des patterns à base de schémas ou d'ontologies [Quafafou et Naouali, 2005].

Le paragraphe suivant présente l'utilisation des ontologies dans les systèmes d'enseignement notamment émergents dans le contexte du Web2.0.

1.6.2. Ontologies et systèmes d'enseignement

Notons que dans le domaine de l'éducation, il existe plusieurs initiatives et travaux pour utiliser les ontologies. Le projet KP-Lab²⁷ se concentre sur la création d'un système d'apprentissage facilitant des pratiques innovatrices de partage, de création et de traitement de la connaissance dans l'éducation et sur les lieux de travail. Pour atteindre ces objectifs, des outils et des modèles utilisant des ontologies ont été conçus. Le portail O4E²⁸ [O4E] est le fruit d'un projet entre les universités de Winston-Salem, Pittsburgh et Saint-Petersburg. Le but principal du projet est de rassembler et disséminer des ressources Web dans le domaine de l'application des ontologies et des technologies appropriées pour des perspectives éducatives modernes. Le portail est « topic map driven ». Pour structurer le domaine, une ontologie O4E a été développée. Le langage utilisé est XTM (eXtended Topic Map). Cette ontologie est utilisée pour indexer les ressources offertes par le portail. Il est possible d'enrichir le contenu du portail en fournissant des fichiers XTM créés avec l'éditeur TM4L²⁹. Cet éditeur d'ontologie adopte le paradigme « topic map » [Dicheva et al, 2006b]. Il considère ainsi qu'une ontologie est constituée par des « topics » reliés par un ensemble de relations prédéfinies, qu'il est possible d'étendre pour répondre à une problématique particulière.

Dans [Cristea, 2004e], l'auteur exprime que l'exploitation d'ontologies dans le domaine des systèmes adaptatifs et intelligents d'enseignement est tout à fait indiquée car elle autorise une meilleure interopérabilité et échange de données. Plusieurs recherches ont porté sur l'exportation de contenus entre les logiciels auteurs et les systèmes adaptatifs et intelligents d'enseignement. Citons, MOT vers WHURLE [Stewart et al, 2004]; MOT vers AHA! [Stach et al, 2004], Interbook vers AHA! [De Bra et al, 2003b], AHA! vers Claroline [Arteaga et al, 2004]. Ces travaux ont montré que l'étape la plus importante était la définition d'une sémantique commune entre tous les systèmes intervenant dans la fabrication et la délivrance de contenus pédagogiques. L'objectif étant qu'un matériel pédagogique et sa stratégie d'enseignement créés sur un certain système puissent être délivrés sur un autre. Dans [Schwartz, 2003], l'auteur précise que le Web sémantique est un moyen pour offrir un environnement dans lequel des systèmes d'informations connectés à internet pourront échanger des spécifications sur la connaissance et les actions qui manipulent cette connaissance.

La nouvelle génération de logiciels sociaux comme les « blogs » et « wikis », ou les systèmes collaboratifs de filtrage/recommandation, d'annotation partagée et de navigation sociale deviennent courant dans les offres d'enseignement à distance [Dron, 2006]. Le filtrage collaboratif a depuis longtemps été utilisé dans l'enseignement [Recker et al, 2000]. Les techniques de navigation sociale ont été utilisées dans de nombreux contextes d'apprentissage [Brusilovsky et al, 2004d]. L'annotation considérée comme l'action d'ajouter un commentaire, une remarque, une explication ou une note à un document ou une partie de ce dernier faite dans un cadre collaboratif est d'un grand intérêt pour les communautés éducative, professionnelle et scientifique [Piombo et al., 2007]. Le réseau d'excellence PROLEARN³⁰ a reconnu la nécessité d'employer les logiciels sociaux dans le cadre de la formation professionnelle. Il a esquissé quelques recommandations pour une nouvelle génération de plateformes d'enseignement appelée « plateforme d'apprentissage adaptatif collaboratif³¹ » [Klamma et al, 2006]. Une des idées est de relier

²⁷ <http://www.kp-lab.org/>

²⁸ Ontologies for Education

²⁹ <http://compsci.wssu.edu/iis/NSDL/download.html>

³⁰ <http://www.prolearn-project.org>

³¹ CALP :Collaborative Adaptive Learning Platforms

les élèves par le contenu [Nodenot et al., 2005]. Afin de permettre l'échange efficace entre ces derniers, le système doit avoir des possibilités sémantiques basées sur l'utilisation des ontologies. Dans SITIO [Gomez et al, 2007], une plateforme sémantique sociale de recommandation comporte de telles capacités. Ce type d'échange structuré permet d'appliquer un raisonnement formel et des stratégies d'inférence pour classifier et relier l'information.

Nombreux sont les outils d'annotation qui utilisent aujourd'hui les ontologies [Piombo et al., 2007]. OntoElan [Chebotko et al, 2004] permet d'annoter une vidéo en se basant sur une ontologie qui décrit les concepts à utiliser. ConAnnotator [HU et al, 2006] propose une activité d'annotation collaborative basée sur une ontologie écrite en OWL. SMARTNotes [Bouzidi et al, 2006] est un outil Web d'annotation collaborative qui offre aux élèves et à leur tuteur de collaborer autour d'un document de cours HTML. Il utilise un modèle d'annotation basé sur une description RDF qui caractérise uniquement la sémantique de l'annotation sur 4 concepts : Highlighting, Consign, Conversation et Indication. MemoNote [Azouaou et al, 2006b ; Azouaou et al, 2006c] est un outil d'annotation sémantique dédié aux enseignants. Il offre le moyen de garder une trace de toutes les annotations qu'un enseignant peut faire lors de ses différentes activités (préparer un cours, dispenser un cours, préparer un examen...). Une ontologie permet d'adapter le type d'annotation aux différentes activités de l'enseignant. Cette ontologie prend en considération aussi bien des informations sur les matériels et logiciels utilisés par l'élève que des informations contextuelles comme le lieu et la date de l'activité ou le niveau (bachelor 1, master 1...) et le type d'apprentissage (cours, exercice, simulation...) visés par l'activité suivie par l'élève [Azouaou et al, 2006a].

1.6.3. Ontologies et « profiling »

Le « profiling » est l'activité qui consiste à analyser le comportement d'une personne afin d'en déduire son profil. Le profil est à considérer comme un ensemble de données qui modélisent certaines caractéristiques de la personne.

Dans un environnement professionnel, une entreprise cherche généralement à profiler les compétences de ses employés. Les compétences étant, donc dans ce cas, les caractéristiques à modéliser. Ceci dans le but d'identifier que chaque employé possède bien les compétences requises pour les activités métiers qu'il doit accomplir. Le système eCCO³² [Pernici et al, 2006] utilise une ontologie pour décrire des concepts à associer au profil d'un employé. On trouve par exemple comme concepts :

- **Connaissance** : l'ensemble de savoir-quoi, savoir-comment et savoir-pourquoi.
- **Objet de connaissance (KO)** : un petit ensemble autonome de connaissances dans un domaine donné.
- **Qualification** : KO mis en œuvre dans une action, KO + verbe d'action (AV). Être capable de faire quelque chose.
- **Compétence** : une qualification dans un contexte (Cx) spécifique d'une activité, KO+AV+Cx. Être capable de faire quelque chose avec le respect d'un contexte spécifique d'une activité.
- **Performance** : un ensemble de comportements observables produisant un résultat.

³² eCompetences and Certifications Observatory

- Profil métier : un ensemble de compétences reliées au concept de performances.

En utilisant un réseau sémantique, le système est capable par exemple d'indiquer à un employé les profils métiers qui sont proches de ses compétences et performances. Mais l'inconvénient majeur de ce système est que l'employé doit renseigner lui-même un questionnaire pour affecter des valeurs aux différents concepts de l'ontologie utilisés pour concevoir le réseau sémantique.

Dans le contexte des systèmes adaptatifs d'enseignement, l'activité de profiler consiste à définir les valeurs des éléments manipulés par les différents modèles, comme par exemple : le modèle de connaissances et le modèle de l'élève. Certains auteurs préconisent l'utilisation d'ontologies différentes pour chacun de ces modèles [Sosnovsky, 2006], [Castello et al, 2006], [Morales et al, 2006], [Ounnas et al, 2006].

OntobUM [Razmerita, 2003] est un serveur de modèles d'élèves stockés au format RDF/RDF-S. Le modèle d'un élève est construit sur la base d'informations données manuellement par l'élève. Des services intelligents complètent automatiquement ces données par d'autres informations estimées par un certains nombres d'heuristiques. L'architecture du système se compose de trois ontologies différentes (Figure 1-28) :

- Ontologie de l'élève : structure les différentes caractéristiques d'un élève et leurs relations.
- Ontologie du domaine : structure les concepts spécifiques au domaine traité et à l'application utilisée.
- Ontologie de log : définit la sémantique des interactions de l'élève avec le système.

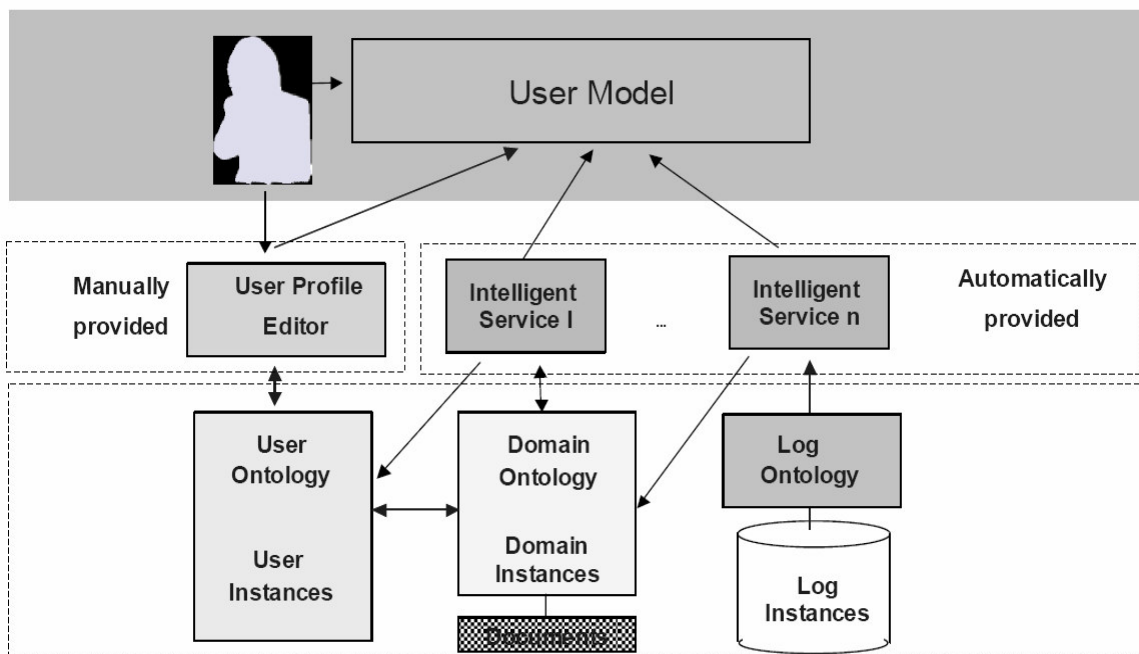


Figure 1-28 : architecture de OntobUM

Le modèle de l'élève est structuré suivant les spécifications IMS LIP³³. Le paquet IMS LIP est décomposé en onze groupes de données (identification, objectif, qualifications, certification et diplômes, accessibilité, activité, compétence, intérêt, affiliation, sécurité et relation) auxquels les auteurs ont ajouté le concept de « comportement » de l'élève. En se basant sur les activités menées dans le système par un élève, OntobUM profile le comportement des élèves. Par exemple, en fonction du nombre d'interventions dans le système et de documents lus, l'élève est stéréotypé comme « readers », « writers » ou « lurkers ». La classification est faite par des heuristiques qui ne sont pas fondées sur un modèle formel. La première faiblesse de ce système est qu'il ne procure aucun moyen pour estimer le niveau de validité du profil attribué à l'élève et revoir ainsi le cas échéant son classement. La seconde est qu'il demande à l'utilisateur de répondre manuellement à un questionnaire pour initialiser certaines parties du modèle. Ce principe ayant montré ses limites dans la modélisation de certaines composantes du modèle de l'élève comme les préférences par exemple. Il est difficile pour un élève d'exprimer ce qu'il préfère.

D'autres systèmes proposent de limiter les demandes à l'élève à profiler. Par exemple dans [Song et al, 2005], les auteurs proposent pour découvrir les centres d'intérêts d'un élève, d'analyser uniquement un ensemble de documents qu'il aura disposé dans un répertoire. L'analyse consiste à classer les documents en fonction d'une ontologie proposant, sous forme de hiérarchie arborescente, un ensemble de centres d'intérêts. Les mots contenus dans les documents servent de base à leur classification. Les auteurs utilisent la méthode LSI³⁴ [Noriaki, 2002] pour définir l'espace sémantique latent à chaque document. Il est alors possible de construire dans le modèle de l'élève la partie estimant ses centres d'intérêts comme étant une sorte de modèle overlay sur l'espace sémantique latent fourni par les documents analysés et l'ontologie de référence. Mais cette technique est difficilement exploitable pour découvrir d'autres caractéristiques reliées aux préférences de l'élève comme le style d'apprentissage. Ceci étant simplement dû au fait que l'ontologie utilisée est trop restrictive, puisqu'elle se limite aux centres d'intérêts.

Plus récemment, le Web 2.0 développe l'idée de former des groupes de travail pour apprendre. Partant du principe que pour travailler ensemble il faut avoir des points communs, certains auteurs [Bermudez et al, 2006] ont montré le besoin de profiler des communautés d'utilisateurs en se basant sur des ontologies de domaines. Les utilisateurs faisant partie d'une communauté, ont forcément des points communs. On retrouve ici l'idée développée par le modèle social. Dans [Cantador et al, 2006], les auteurs proposent de créer un réseau sémantique social multicouche basé sur les similitudes découvertes pour chaque utilisateur. Dans ce système, un utilisateur i est représenté par un vecteur de préférences $u_i = (u_{i1}, u_{i2}, \dots, u_{iN})$. L'ontologie est composée de N concepts. Les éléments $u_{ij} \in [0,1]$ mesurent le degré d'intérêt de l'élève i pour le concept c_j de l'ontologie du domaine considéré. En analysant la structure de l'ontologie du domaine et en considérant le vecteur de préférences de chaque utilisateur, les auteurs décomposent l'espace de concepts en clusters. L'indice de similarité entre les préférences d'un utilisateur $u_i = (u_{i1}, u_{i2}, \dots, u_{iN})$ et un cluster C_r est calculé par :

$$sim(u_i, C_r) = \frac{\sum_{c_j \in C_r} u_{ij}}{|C_r|}$$

³³ Information Management Systems Learner Information Package (www.imsproject.org/aboutims.html)

³⁴ Latent Semantic Indexing

Dans cette formule, c_j représente le concept associé à l'élément u_{ij} du vecteur de préférences et $|C_r|$ est le nombre de concepts inclus dans le cluster. Le cluster avec le plus fort indice de similarité est alors assigné à l'utilisateur. Ainsi, les groupes de communautés d'utilisateurs sont créés. Cette technique est certes intéressante pour profiler des communautés d'utilisateurs, mais peut-on pour autant conclure que tous les utilisateurs qui se passionnent pour les reptiles auront tous le même style d'apprentissage ? Il semble évident que non. Donc, il apparaît difficile d'utiliser cette classification dans le contexte d'un système d'enseignement pour adapter des contenus aux profils des différentes communautés d'utilisateurs.

1.7. Conclusion

Les modèles de connaissances présentés ci-dessus ne peuvent attribuer les causes d'éventuels échecs qu'aux seules déficiences en matière de connaissances. Or, les causes peuvent être multiples et liées à la méthode pédagogique.

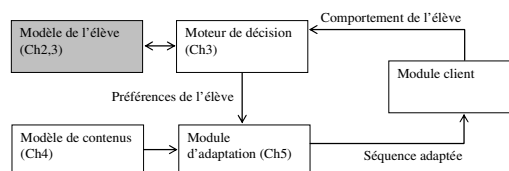
Ils permettent difficilement d'envisager de rejouer la séquence avec une nouvelle stratégie d'enseignement afin de mener l'élève à atteindre l'objectif d'apprentissage. Leur seule utilité est de permettre la réponse à la question : Quelles sont les connaissances acquises par l'élève ?

Les modèles de préférences quant à eux, maintiennent la plupart du temps simplement des informations sur la préférence de l'élève quant au support de l'information : Texte, son, image, vidéo, etc.

Les modèles rencontrés ne gèrent pas l'incertitude qui est associée à cette information : ils sont statiques [Frasincar et al, 2004] ! Ils ne permettent pas de prendre en compte l'évolution des préférences de l'élève au cours de son apprentissage, ni l'erreur qu'il aurait pu commettre dans ses réponses lors de l'initialisation du modèle. Car ces systèmes partent du principe que l'élève est capable de comprendre les différents aspects qui peuvent être adaptés à lui et identifier clairement ses préférences. De plus, dans ce cas il est aussi fait l'hypothèse que l'ensemble des préférences d'un élève peut être découvert soit par un questionnaire ou bien par un pré-test.

Depuis la naissance du Web2.0, les systèmes d'enseignement tentent d'augmenter la sémantique associée à leurs modèles et leur capacité à les partager. Différents types d'ontologies (domaine, utilisateur, comportement...) ont été proposés.

Dans cette thèse, nous proposons une approche globale qui se fonde sur un modèle probabiliste des préférences de l'élève. Ce modèle permet par inférence de déterminer les métadonnées adaptées pour une activité pédagogique. Ces métadonnées servent à la sélection des contenus adaptés dans une base indexée par deux ontologies : une première du domaine et une seconde pour la pédagogie. Nous réalisons ainsi, une instance d'application Web2.0 pédagogique. Le prochain chapitre décrit l'étude statistique des préférences de l'élève et leur dépendance aux styles d'apprentissage. Les chapitres suivants exploitent ce résultat pour concevoir les composants dans notre système d'enseignement adaptatif.



CHAPITRE

2. Activité de profilage : étude statistique du style d'apprentissage

Résumé :

Les systèmes adaptatifs et intelligents d'enseignement par le Web utilisent un modèle pour décrire les caractéristiques de l'élève. La détermination de ces caractéristiques se fait généralement par l'observation du comportement de l'élève au cours de la séquence d'enseignement. La méthode de transformation des observations en informations pédagogiques exploitables se nomme *profilage*.

Dans ce chapitre, nous présentons un méta-modèle conceptuel générique décrivant le processus de profilage comme étant l'estimation de variables non observables à partir d'indicateurs observables. Nous instancions ce modèle dans le cadre de l'enseignement à distance pour profiler le style d'apprentissage d'un élève. Pour valider cette théorie, nous avons mené une étude empirique pour caractériser les indicateurs mesurables du comportement de l'élève. En effet, nous proposons trois mesures originales du comportement : l'énergie d'apprentissage, le degré d'attention et le niveau d'adaptation. Ces indicateurs sont usités pour définir des variables qui seront utilisées pour inférer le style d'apprentissage.

2.1. Introduction

Le modèle de l'élève est un composant fondamental de tout système adaptatif d'enseignement. Il permet de décrire les caractéristiques de l'élève. Ce genre de modèle contient des informations utilisées pour adapter les séquences d'enseignement. Ces systèmes doivent identifier pour chaque élève : ses objectifs d'apprentissage [Clifford, 2000], son niveau de connaissances [De Bra, & Calvi, 1998], ses préférences [Höök et al, 1997], ses stéréotypes [Zakaria et Brailsford, 2002], ses préférences cognitives [Chen et al, 2002] et ses styles d'apprentissage [Stach et al, 2004].

L'identification de ces caractéristiques se fait généralement par l'observation du comportement de l'élève au cours de la séquence d'enseignement. La méthode de transformation des observations en informations pédagogiques exploitables se nomme *profilage*.

C'est ainsi que dans l'ensemble des informations présentées ci-dessus, certaines sont considérées comme des données que l'on peut recueillir par l'analyse du choix de l'élève. Par exemple, l'élève peut indiquer dans une liste, les sujets qu'il désire consulter. Le système est alors capable de garder une trace des choix faits au cours de différentes sessions et de construire ainsi une classification des élèves par stéréotype [Messina et al, 2006]. Ce genre de classification serait ensuite utilisé pour composer dynamiquement le prochain contenu à présenter à l'élève.

On peut aussi, en observant le contenu des documents que visualise l'élève [Semeraro et al, 2005], conclure qu'il préfère le français comme langue et la vidéo comme support quand c'est possible. Ces informations peuvent être utilisées pour adapter l'information qui lui sera proposée par la suite. Il est aussi possible de mesurer son niveau de connaissances sur un domaine particulier en lui proposant de réaliser une évaluation sommative.

Malgré tout, certaines données restent non observables comme par exemple le style d'apprentissage. Il est pourtant nécessaire d'arriver à capturer cette information car elle est souvent à la base des méthodes d'adaptation dans les systèmes hypermédias adaptatifs d'enseignement.

Nous considérons donc que :

- Le profil est un ensemble de couples paramètre-valeur ;
- Les paramètres d'un profil ne sont pas tous mesurables ;
- Pour tout paramètre non mesurable, il existe un ensemble d'indicateurs observables permettant de l'estimer.

Dans ce contexte, nous définissons le profilage comme étant le processus qui consiste à :

1. Définir un ensemble de paramètres non observables, leurs dépendances et leurs domaines de définition ;
2. Définir l'ensemble des indicateurs observables, leurs dépendances et leurs domaines de définition ;
3. Définir les relations d'influence entre les paramètres non observables et les indicateurs observables.

Dans ce chapitre, nous nous proposons d'établir un méta-modèle conceptuel décrivant de façon générique le processus de profilage. Ce méta-modèle nous sert à créer un modèle de dépendance entre des indicateurs observables et des paramètres non observables pour profiler le style d'apprentissage d'un élève. Nous y proposerons, notamment, trois indicateurs originaux du comportement de l'élève : l'énergie d'apprentissage, le degré d'attention et le niveau d'adaptation. Ensuite, nous présentons les résultats de deux études empiriques que nous avons menées pour valider ce modèle. La première étude nous a permis de valider notre questionnaire d'évaluation du style d'apprentissage selon le modèle proposé par Felder. La seconde nous a permis de valider les paramètres et indicateurs que nous proposons ainsi que leurs dépendances. Nous généralisons, ces résultats obtenus, à un ensemble étendu de caractéristiques pour décrire le style d'apprentissage, le matériel et l'activité pédagogique ainsi que le comportement de l'élève.

2.2. Méta-modèle conceptuel du processus de profilage

Reprenant la définition du profilage, nous proposons un modèle conceptuel qui manipule deux catégories d'informations :

- **Les paramètres non observables** : l'ensemble des paramètres non observables se compose de caractéristiques qu'il est impossible de mesurer directement. Ce sont généralement des caractéristiques intrinsèques. Par exemple, mesurer le style d'apprentissage d'un élève ne peut pas se faire directement lors d'une séquence pédagogique. C'est une caractéristique intrinsèque de l'élève. De même, les centres d'intérêts d'un utilisateur de e-commerce peuvent être vus comme des caractéristiques intrinsèques. Les mesurer directement n'est pas possible.
- **Les indicateurs observables** : l'ensemble des indicateurs observables se compose de caractéristiques dont les valeurs peuvent être mesurées concrètement ou bien contrôlées par le système. Par exemple, il est possible de mesurer le temps qu'un élève passe dans une activité pédagogique, de mémoriser son parcours dans les contenus ou bien encore d'enregistrer qu'il a joué toutes les animations qui étaient proposées dans la séquence. Toutes ces caractéristiques peuvent aussi être contrôlées par le système en imposant à l'élève un parcours, ou en lui en présentant uniquement une partie spécifique, éventuellement dans un temps limité. De la même façon, il est possible de déterminer que l'utilisateur d'un site de e-commerce ne visualise que des pages décrivant des équipements motos et particulièrement de la marque Suzuki.

Pour établir le modèle, nous définissons deux sortes de relations d'*influence* (ou de *dépendance*) entre les paramètres et les indicateurs:

- Influence des paramètres non observables sur les indicateurs observables (NObs-Obs): cette relation exprime que les valeurs des indicateurs observables sont influencées par des caractéristiques inhérentes non observées (c'est à dire les paramètres non observables). Par exemple, le choix de l'élève lors du parcours des contenus pédagogiques sera influencé par son style d'apprentissage. De même, lorsqu'un internaute parcourt uniquement les pages d'un site de e-commerce présentant le burgman 650 Suzuki, on peut considérer qu'il est influencé par son centre d'intérêt : le scooter.
- Influence entre indicateurs observables (Obs-Obs): il est, en effet, possible que des dépendances existent entre indicateurs observables. On peut ainsi exprimer que le type d'activité pédagogique dépendra du niveau de difficulté de la tâche d'apprentissage envisagée.

Il est important de noter ici que nous considérons inexistantes les influences entre paramètres non observables (NObs-NObs). Nous justifions cela par le fait que les variables non observables sont des constantes psychologiques qui causent les indicateurs observables. Les constructs étant des composants fondamentaux, ils ne devraient pas être causés par autres choses !

En pratique, ce modèle peut être généralisé en incluant les influences NObs-NObs. Cela peut par exemple permettre de capturer l'influence de l'intérêt sur la satisfaction, tous deux non observables.

La Figure 2-1 présente une représentation UML du processus de profilage en utilisant un diagramme de classe. Dans ce schéma, on note que les indicateurs observables et les paramètres non observables sont considérés comme des sortes de variables.

Une variable est caractérisée par trois attributs :

- nom : il s'agit de l'identificateur de la variable ;
- type : représente le domaine de définition de la variable (ex : entier, réel, énumération...);
- valeurs : ensemble des valeurs admises pour la variable (ex : [1..10], {image, texte, son}, [1,10]).

Ce méta-modèle peut être décrit avec un langage comme OWL, donnant le début d'une méta-ontologie très abstraite décrivant le profilage. On pourrait y greffer les variables observables et non observables dans un cas donné ainsi que la description des méthodes de transformations. Précisément, la section suivante présente une instance de ce méta-modèle par ce qui nous intéresse dans cette thèse : le style d'apprentissage.

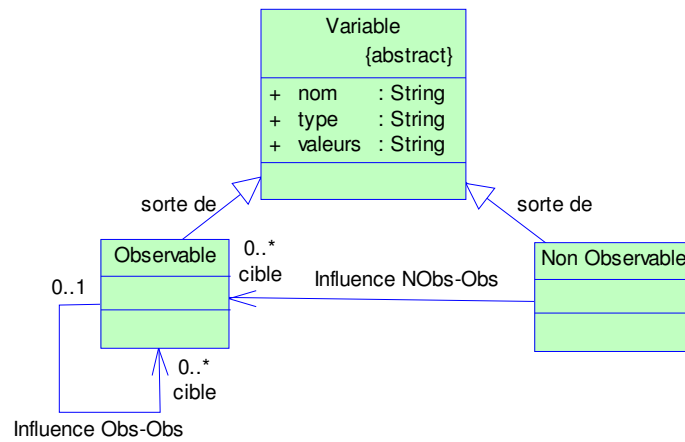


Figure 2-1 : diagramme de classes du processus de profilage

2.3. Modèle de dépendance du style d'apprentissage

Le style d'apprentissage est une caractéristique psychologique des individus. Il détermine la manière dont ils perçoivent et traitent mentalement l'information [Felder et al. 1988, 1993; Kolb, 1985; Honey et Mumford, 1992; Grasha, 1996; Riding et al. 2001 ; Riding et al, 1998].

Les psychologues cognitivistes ont proposé divers modèles de style. Souvent, il est représenté en utilisant des dimensions multiples estimées statistiquement. Une revue peut en être trouvée dans [Riding et al, 2001].

Chaque modèle suggère un instrument pour mesurer le style d'un individu. La plupart d'entre eux utilisent des questionnaires. Cependant, ils partagent tous la même méthode : estimer le modèle, qui est l'information cachée, à partir de données observables représentant les réponses de l'utilisateur.

En utilisant le méta-modèle de la Figure 2-1, nous proposons un modèle de dépendance pour décrire le style d'apprentissage. Nous commençons par étudier les indicateurs observables et les paramètres non observables avant de terminer par les dépendances qui les relient.

2.3.1. Indicateurs observables

Les indicateurs observables de notre modèle incluent les variables qui décrivent l'activité pédagogique et le matériel pédagogique. Notre représentation du matériel pédagogique est fondée sur l'hypothèse suivante : « Un matériel (ou objet) pédagogique est conçu pour atteindre un objectif d'apprentissage simple et clairement identifié ».

Par conséquent, nous représentons une activité d'apprentissage par un ensemble de métadonnées décrivant la nature et l'objectif de la tâche. Le matériel pédagogique correspondant est décrit en utilisant un autre ensemble de métadonnées telles que les médias utilisés, la structure, la nature et d'autres caractéristiques.

Soit \mathcal{T} l'ensemble des métadonnées décrivant l'activité pédagogique,

$$\mathcal{T} = \{T_1, T_2, \dots, T_k\};$$

et \mathcal{M} l'ensemble des métadonnées décrivant le matériel pédagogique

$$\mathcal{M} = \{M_1, M_2, \dots, M_n\};$$

Ces métadonnées sont représentées par des variables discrètes. Nous considérons que ces variables prennent les valeurs suivantes :

$$T_i \in \{t_j^i\}_{j \in [1, \text{cad}(T_i)]},$$

$$M_i \in \{m_j^i\}_{j \in [1, \text{cad}(M_i)]}.$$

Nous représentons également le comportement de l'élève par un ensemble de paramètres tels que la performance, le temps passé, et les modèles de navigation. Nous notons \mathcal{P} l'ensemble de variables décrivant de tels paramètres,

$$\mathcal{P} = \{P_1, P_2, \dots, P_m\};$$

avec

$$P_i \in \{p_j^i\}_{j \in [1, \text{cad}(P_i)]}.$$

2.3.2. Paramètres non observables

Dans notre cas, les paramètres non observables sont les seules variables associées aux dimensions du style d'apprentissage. Soit \mathcal{D} l'ensemble de variables représentant ces dimensions,

$$\mathcal{D} = \{D_1, D_2, \dots, D_q\};$$

où D_i sont des variables discrètes avec

$$D_i \in \{d_j^i\}_{j \in [1, \text{cad}(D_i)]}.$$

2.3.3. Dépendances entre indicateurs observables et paramètres non observables

Les indicateurs observables et les paramètres non observables sont reliés par différentes dépendances. Le choix de ces dépendances est fondé sur des hypothèses pédagogiques décrites dans les paragraphes suivants.

Premièrement, nous considérons que la pertinence du matériel pédagogique pour atteindre un objectif d'apprentissage donné dépend du style d'apprentissage de l'élève.

En d'autres termes, l'objectif d'apprentissage et le style d'apprentissage ont une influence causale sur le choix du matériel pédagogique. En utilisant le symbole $\xrightarrow{\text{influence}}$ pour représenter la dépendance, nous pouvons écrire :

$$\{\mathcal{T}, \mathcal{D}\} \xrightarrow{\text{influence}} \mathcal{M}$$

En instanciant le modèle conceptuel de la Figure 2-1 et en ajoutant la relation précédente, on obtient le diagramme de la Figure 2-2.

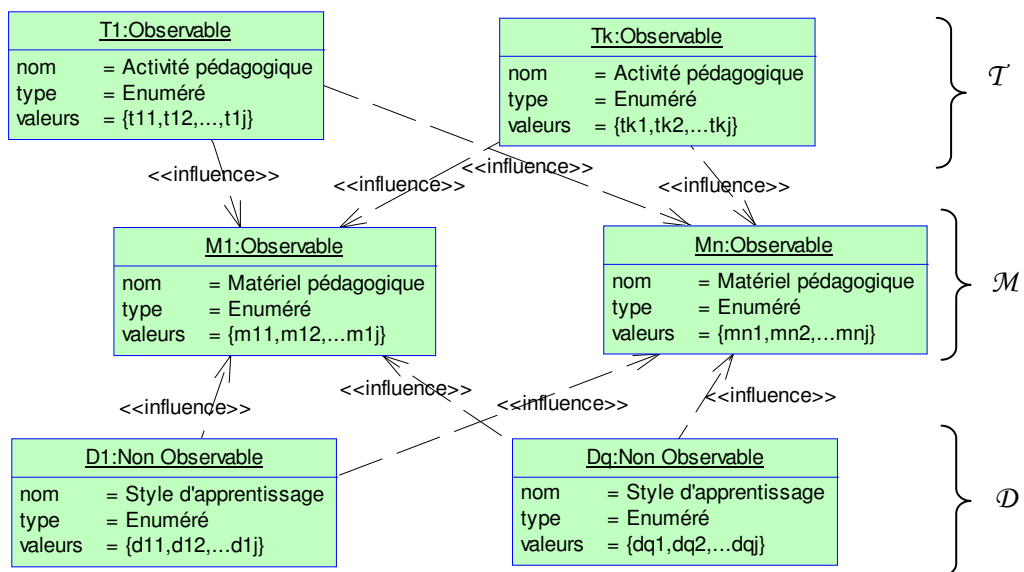


Figure 2-2 : dépendances entre \mathcal{D} , \mathcal{T} , et \mathcal{M} .

Une autre hypothèse pédagogique consiste à considérer que le comportement d'apprentissage (la performance de l'élève, l'effort de concentration, ...) dépend de l'adéquation de l'activité et du matériel pédagogique au style d'apprentissage de l'élève. Pour un objectif donné, fournir à l'élève le matériel pédagogique adapté à son style d'apprentissage mène à un meilleur comportement (performance maximale, effort minimal...). Ceci signifie que l'activité pédagogique, le matériel pédagogique et le style d'apprentissage ont ensemble une influence causale sur le comportement :

$$\{T, M, D\} \xrightarrow{\text{influence}} P$$

Développer cette relation mène à l'instance représentée sur la Figure 2-3.

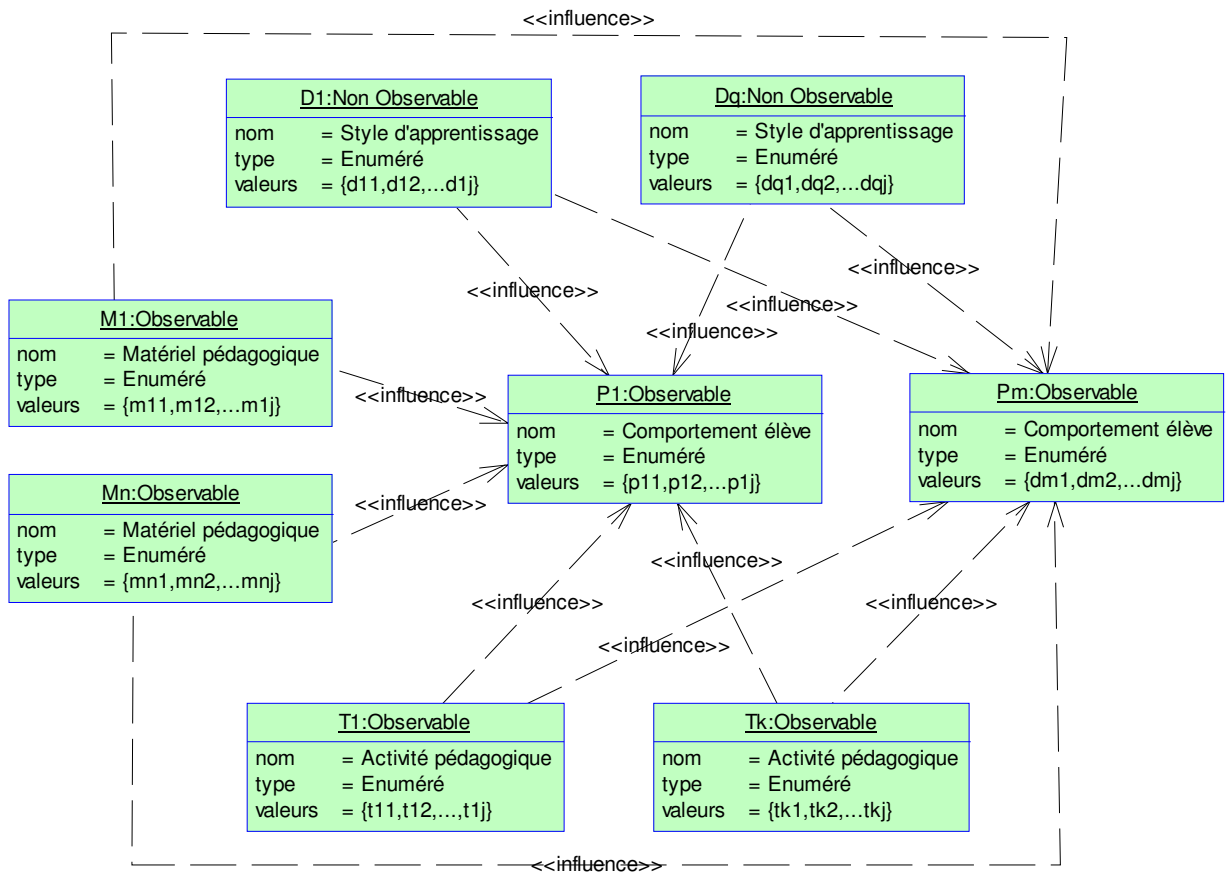


Figure 2-3 : relations entre D, M, T et P

Pour valider ces hypothèses pédagogiques, nous avons mené deux études empiriques. La première étude cherche à valider notre questionnaire pour mesurer le style d'apprentissage d'un élève. La seconde quant à elle, tente de valider que le comportement d'un élève est influencé par le niveau d'adaptation de la séquence pédagogique à son style d'apprentissage.

2.4. Mesure du style d'apprentissage : Questionnaire ILS

Cette mesure est basée sur l'index des styles d'apprentissage (ILS³⁵) établi par Felder et Solomon [Felder R.M. et al 1988, 1993]. L'ILS considère que les approches des personnes pour percevoir et traiter mentalement l'information peuvent être classées selon quatre dimensions. Nous avons traduit en français, avec l'aide de psychologues, le questionnaire original en anglais (Annexe 1) de Felder-Soloman composé de 44 questions modales. Les nouvelles questions ont de plus été adaptées pour rester proches de préoccupations scolaires ou d'adolescents.

L'hypothèse que nous cherchons à vérifier est :

H_0 : La nouvelle version Française du questionnaire ILS permet de mesurer le style d'apprentissage des élèves de cinquième selon le modèle proposé par Felder.

2.4.1. Echantillon de la population pour l'étude

L'étude que nous avons menée s'est adressée à des étudiants de première année psychologie de l'université du Mirail de Toulouse en janvier 2004. Des élèves volontaires ont pu participer à l'étude. L'échantillon de la population a été composé au final de 435 élèves.

Le sujet principal de cette étude était de montrer que la version Française du questionnaire ILS (Annexe 2) permet de mesurer avec une bonne confiance le style d'apprentissage de l'élève.

2.4.2. Mesures de l'étude

Dans cette étude, nous avons utilisé un seul instrument de mesure : la version Française du questionnaire ILS adaptée pour les élèves de cinquième.

Le questionnaire comporte 44 questions fermées à deux modalités (a et b) qui forment 4 groupes de 11 questions (Annexe 1). Chaque groupe de questions définit une dimension pour le modèle cognitif de l'élève qui est donc composé de 4 dimensions (Tableau 2-1):

Dimensions	Extrémité inférieure (a)	Extrémité supérieure (b)	Questions ILS
Réflexion	Actif	Réfléchi	Q1, Q5, Q9, Q13, Q17, Q21, Q25, Q29, Q33, Q37, Q41
Raisonnement	Inductif	Déductif	Q2, Q6, Q10, Q14, Q18, Q22, Q26, Q30, Q34, Q38, Q42
Sensorielle	Verbal	Visuel	Q3, Q7, Q11, Q15, Q19, Q23, Q27, Q31, Q35, Q39, Q43
Progression	Séquentiel	Global	Q4, Q8, Q12, Q16, Q20, Q24, Q28, Q32, Q36, Q40, Q44

Tableau 2-1 : Question ILS par dimension du style

³⁵ ILS : Index of Learning Style

Pour localiser l'élève sur une dimension, il suffit de comptabiliser le nombre de réponses a et de réponses b sur les 11 questions correspondantes et d'en faire la soustraction afin d'obtenir un nombre positif. Soient A et B respectivement ces nombres. La différence A-B permet de situer le style de l'élève (Figure 2-4). Un nombre négatif indique que l'élève est proche de l'extrémité b et inversement. On attribut à ce style un degré de confiance calculé par $d = |A - B|$. Le Tableau 2-2 donne la sémantique de cette mesure.

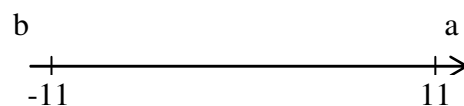


Figure 2-4 : dimension du style

Degré de confiance	Signification
1-3	incertain
5-7	modéré
9-11	fort

Tableau 2-2 : degré de confiance de la mesure du style

La première dimension D_1 (Figure 2-5) représente la dimension sensorielle. Elle mesure la manière avec laquelle l'élève préfère percevoir l'information. Son domaine de définition est un continuum qui varie de visuel à verbal.

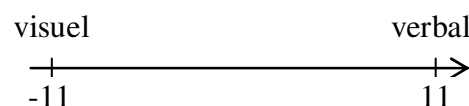


Figure 2-5 : dimension sensorielle (D_1)

Dans un contexte d'enseignement, connaître cette capacité sensorielle aide à choisir les médias appropriés pour présenter l'information à l'élève. Une personne ayant $D_1 = -11$, préférera des documents en grande partie illustrés avec des images, des graphiques, etc. Alors qu'un élève ayant $D_1 = 11$ préférera le texte ou le son.

La seconde dimension D_2 (Figure 2-6) représente la dimension de progression. Elle définit la manière, globale ou séquentielle, que la personne préfère pour progresser (naviguer) dans une séquence pédagogique. Un élève global préférerait plutôt choisir librement son parcours dans la séquence. Alors qu'une personne séquentielle se sentira mieux quand elle est contrainte à progresser pas à pas. Un matériel pédagogique devrait, donc, fournir une navigation et une progression linéaire contrainte pour les élèves séquentiels et une navigation sélective pour les élèves globaux.

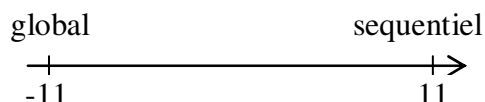


Figure 2-6 : dimension de progression (D_2)

La troisième dimension D_3 (Figure 2-7) représente la dimension de la réflexion. Elle varie de réfléchi à actif. Les personnes actives réussissent mieux en faisant des activités pratiques individuelles ou collectives. Les personnes réfléchies préfèrent les activités qui nécessitent un effort personnel d'analyse. Pour atteindre un objectif d'apprentissage donné, un matériel pédagogique approprié pour les élèves actifs devrait donner plus de poids à la pratique, au travail de groupe et à l'expérimentation. Inversement, pour être plus efficace avec un élève réfléchi, le matériel pédagogique devrait être la plupart du temps basé sur la théorie, les démonstrations et d'autres activités personnelles.

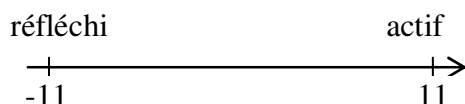


Figure 2-7 : dimension de réflexion (D_3)

La dernière dimension D_4 (Figure 2-8) représente le raisonnement. Elle varie de déductif à inductif. Les personnes déductives préfèrent progresser de la théorie vers la pratique. Elles sont considérées comme penseurs abstraits. Alors que, les personnes inductives préféreraient progresser des faits, exemples, et pratiques vers la théorie. Elles sont considérées comme penseurs concrets. Un matériel pédagogique adapté aux élèves déductifs devrait commencer par des définitions et des théories et progresser vers la pratique.

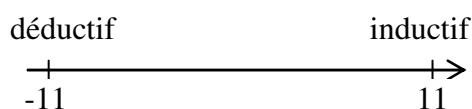


Figure 2-8 : dimension du raisonnement (D_4)

2.4.3. Conception et mise en œuvre de l'étude

Des enseignants psychologues du laboratoire travail et cognition (UMR 5551) de l'université du Mirail, ont administré le questionnaire papier dans leurs classes. Les réponses ont été saisies et analysées par le logiciel SPHINX.

2.4.4. Résultats de l'étude

L'étude que nous avons menée, nous a permis de montrer le taux de fiabilité de l'instrument de mesure utilisé dans l'expérimentation : la version Française du questionnaire ILS. Ce résultat est détaillé ci-dessous.

Fiabilité et validité du questionnaire

Nous avons calculé le coefficient α de Cronbach pour évaluer la fiabilité des mesures offertes par le questionnaire. Dans [Felder et al., 2005b], les auteurs précisent que le seuil d'acceptation du coefficient α de Cronbach dépend du type de questionnaire. Dans le cas d'un questionnaire qui mesure un niveau de connaissance, le seuil d'acceptation est fixé à 0.75 et dans le cas d'une mesure de préférences ou d'attitude à 0.50.

La version Française du questionnaire ILS de 44 questions présentée à la population cible a obtenu un coefficient α de Cronbach = 0.66. Ce qui indique un taux de fiabilité convenable pour une mesure du style d'apprentissage. Mais les résultats de l'analyse factorielle des réponses obtenues aux 44 questions, pour identifier les variables (les questions) expliquant correctement les 4 axes du modèle de Felder, se sont avérés mitigés. En effet, seulement 12 questions sur 44 expliquaient correctement les 4 axes (Tableau 2-3). Les autres questions obtiennent soit un taux faible (< 0.10), soit n'expliquent pas le bon facteur. Ceci ayant donc une conséquence sur la classification des élèves en fonction de leur style d'apprentissage.

Facteur 1		Facteur 2		Facteur 3		Facteur 4	
D1 (inductif-déductif)		D2 (visuel-verbal)		D3 (actif-réfléchi)		D4 (séquentiel-global)	
Q18	0.61	Q31	0.48	Q25	0.38	Q36	0.16
Q6	0.68	Q7	0.80	Q1	0.50	Q12	0.21
Q38	0.68	Q11	0.87	Q29	0.57	Q24	0.29

Tableau 2-3 : Facteurs expliqués par la première version du questionnaire ILS

Réduction du questionnaire

Suite aux résultats de l'analyse factorielle, qui a mis en évidence que seules 12 questions expliquaient convenablement les 4 facteurs, nous avons réduit le questionnaire (Annexe 3) à uniquement ces 12 questions. Le faible nombre de questions permet de diminuer le temps de passation. Ceci ayant obligatoirement une conséquence sur la fiabilité de la mesure. Le niveau de concentration des élèves faiblissant au-delà d'un nombre trop important de questions.

Ce nouveau questionnaire a été nommé RILS. Il est à la base de notre deuxième étude empirique.

2.5. Indicateurs observables pour inférer le style d'apprentissage d'un élève

Le modèle conceptuel des Figure 2-2 et Figure 2-3 montre les dépendances des concepts utilisés pour décrire le processus de profilage d'un élève. Dans ce paragraphe, nous présentons une étude empirique qui nous permet de vérifier la validité des dépendances entre indicateurs observables du comportement et le style d'apprentissage de l'élève pour une certaine activité pédagogique.

L'hypothèse que nous cherchons à vérifier est :

H_0 : Ajuster la démarche pédagogique au style d'apprentissage de l'élève influence son comportement lors de la séquence d'enseignement.

2.5.1. Echantillon de la population pour l'étude

Nous avons mené cette deuxième étude au Collège Malraux de Ramonville 31300 en avril 2005, dans quatre classes de cinquième. Seuls les élèves volontaires qui avaient l'autorisation de leurs parents ont participé à l'étude (Annexe 4). L'échantillon de la population a été composé au final de 117 élèves dont 56 filles et 61 garçons. Les élèves étaient tous âgés de 12 à 14 ans.

Le sujet principal de cette étude était de montrer que le comportement de l'élève pouvait être influencé par l'ajustement de la séquence pédagogique à son style d'apprentissage.

2.5.2. Mesures de l'étude

Dans cette thèse, nous adoptons trois sortes de mesures de comportement de l'élève. Nous mesurons le *style* d'apprentissage à l'aide de quatre variables représentant les dimensions de l'ILS. La *performance* est mesurée par un QCM administré en fin de séquence. De plus, nous proposons *trois nouvelles mesures* : l'énergie d'apprentissage que l'élève a fourni pendant la tâche, le degré d'attention représentant l'effort, et le niveau d'adaptation de la séquence. Nous montrons dans cette section leur mode de calcul. L'étude que nous avons menée a mis en œuvre l'ensemble de ces mesures.

Style d'apprentissage

Nous avons utilisé la version Française du questionnaire RILS réduite à 12 questions et adaptée aux élèves de cinquième (Annexe 3). Le nombre réduit de question impose un changement de l'échelle de codage [-3,-1,1,3] sur les quatre axes (Figure 2-9).

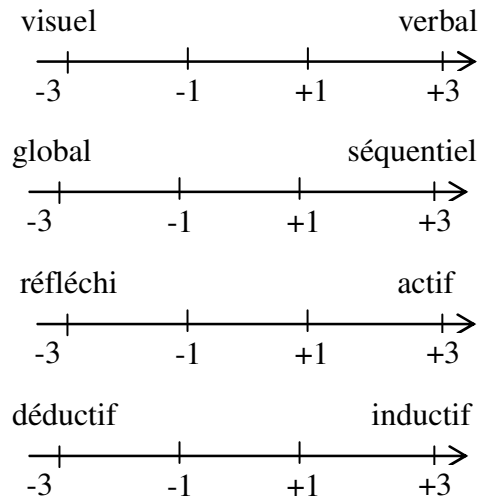


Figure 2-9 : dimensions adaptées du modèle

Performance

La mesure de la performance d'un élève est obtenue par une évaluation sommative. Après une séquence d'enseignement, l'élève répond à un questionnaire sous forme de QCM. Il obtient une note comprise entre [0,20].

Energie d'apprentissage

Afin d'évaluer l'effort déployé par un élève pour suivre une séquence pédagogique, nous proposons une mesure originale dite énergie d'apprentissage. Elle s'exprime de la manière suivante :

$$W = Bloom * (R * i^2 * t) \quad (1)$$

où :

- t est la durée d'apprentissage. Elle exprime le temps que l'élève a mis pour suivre la séquence pédagogique. Elle se mesure en nombre de minutes.
- i est l'intensité d'apprentissage. Elle exprime l'intensité du travail de l'élève pendant la séquence pédagogique. Elle se mesure par le nombre de pages visitées par unité de temps lors de la séquence pédagogique. L'unité utilisée est le nombre de pages visitées par minute (page.mn^{-1}).

$$i = \frac{\text{nbPagesVisitées}}{t}$$

Ainsi, lorsqu'un élève termine la séquence pédagogique en n'ayant visité aucune page de contenu, nous concluons qu'il aura fourni une intensité d'apprentissage nulle.

- R est la résistance à l'apprentissage. Elle exprime la difficulté du matériel pédagogique pour l'élève. Elle se mesure par le nombre moyen de visites d'une page de la séquence par le carré du temps moyen mis pour consulter les pages. L'unité utilisée est la minute au carré (mn^2).

$$R = \frac{\sum_{k \in [1, \text{nbPagesVisitées}]} \text{nbfoisvisitee}(k)}{\text{nbTotalPage}} * t\text{MoyenPage}^2$$

Ainsi quand un élève termine la séquence pédagogique en n'ayant visité aucune page de contenu, nous concluons qu'il aura éprouvé une résistance nulle à l'apprentissage.

$$R = \frac{0}{\text{nbTotalPage}} * t\text{MoyenPage}^2$$

- **Bloom** est un coefficient représentant le niveau de complexité de la tâche. Il permet de pondérer l'énergie à déployer en fonction de la complexité de la tâche. Son intervalle de définition est [1..6]. Il exprime l'ensemble des valeurs admises dans la taxinomie de Bloom (connaissance, compréhension, application, analyse, synthèse, évaluation) [Bloom, 1956].

L'énergie d'apprentissage s'exprime en nombres de pages au carré par unité de temps en minute. Cette approche impose la contrainte du temps optimal d'apprentissage théorique de 1 mn par page de contenu. Cette contrainte est en accord avec les études empiriques qui montrent que la quantité d'information pouvant être traitée par unité de temps est limitée [Mayer, 2001]. Parmi les conseils donnés, on trouve par exemple qu'il est recommandé de faire tenir l'information présentée à un élève sur un seul écran (sans défilement). Que les écrans ne doivent pas être trop chargés, en tenant compte du temps moyen qui montre qu'un élève qui consulte un contenu web passe en moyenne ¼ de seconde par mot et qu'il faut mettre un maximum de 20 phrases de 12 mots (soit 1mn de consultation par page). Ces recommandations trouvent leur origine dans des études cherchant à prouver les limites de l'humain à traiter de l'information par les canaux sensoriels. Elles s'appuient sur la théorie de la charge cognitive [Sweller, 1994] qui est basée sur les travaux de Miller [Miller, 1956].

En plus de cette mesure, nous définissons trois autres grandeurs qui nous servent lors de l'adaptation des contenus :

- $W_{theorique}$ est l'énergie d'apprentissage théorique d'une séquence pédagogique. Il s'agit de l'énergie à déployer par un élève pour suivre la séquence dans les meilleures conditions. Cette énergie se calcule en posant comme hypothèse que l'élève visite chaque page de la séquence une seule fois ($\sum_{k \in [1, nbPagesVisitées]} nbfoisvisitee(k) = nbTotalPage = nbPagesVisitées$) en un temps de consultation de 1 mn ($tMoyenPage = 1$). On obtient alors :

$$W_{theorique} = Bloom * nbTotalPage$$

- **TAL**³⁶ est le niveau d'adaptation de la tâche. Etant donné l'ensemble des notations posées ci-dessus, nous proposons le coefficient TAL pour mesurer le niveau d'adaptation de la tâche au style d'apprentissage de l'élève. C'est le rapport de l'énergie théorique $W_{theorique}$ à l'énergie réelle W :

$$TAL = \frac{W_{theorique}}{W} = \frac{nbTotalPage}{R * t^2 * t}$$

³⁶ TAL : Task Adaptation Level

- U est la tension représentant le degré d'attention. Elle exprime l'effort fourni par l'élève pendant la séquence pédagogique. L'unité utilisée est la page minute (page.mn). Elle se mesure par l'expression suivante :

$$U = R * i = \frac{\sum_{k \in [1, nbPagesVisitées]} nbfoisvisitee(k)}{nbTotalPage} * tMoyenPage^2 * \frac{nbPagesVisitées}{t}$$

Un élève ayant peu d'intérêt pour l'apprentissage, visitera peu de pages ($nbPagesVisitées \approx 0$) ou passera furtivement sur le contenu ($tMoyenPage \approx 0$). Ceci donnerait un effort faible ($U \approx 0$). Ainsi cette mesure nous permettrait d'identifier les élèves qui n'ont pas été sérieux lors de la séquence.

Nous posons l'hypothèse raisonnable que dans des conditions normales $W \geq W_{theorique}$. Ceci implique qu'un TAL valide appartient à l'intervalle $[0,1]$. La situation où $W < W_{theorique}$ indique que l'élève ne s'est pas comporté normalement dans la séquence. Par conséquent, les résultats obtenus ne sont pas significatifs. Ce cas se rencontre uniquement lorsque l'élève obtient moins d'une minute pour la durée moyenne de consultation d'une page ; ou bien lorsqu'il ne visite pas au moins une fois chaque page dans la séquence pédagogique. Cela montrerait que l'élève n'a pas fait sérieusement les activités proposées. De ce fait, il n'est pas utile de considérer ce cas comme une situation d'apprentissage à analyser pour vérifier l'adéquation du matériel pédagogique avec son style d'apprentissage.

Pour un $W_{theorique}$ donné, la Figure 2-10 montre la variation hyperbolique de TAL par rapport à W . Un W nettement plus grand que $W_{theorique}$ indique que la séquence n'était pas adaptée au style d'apprentissage.

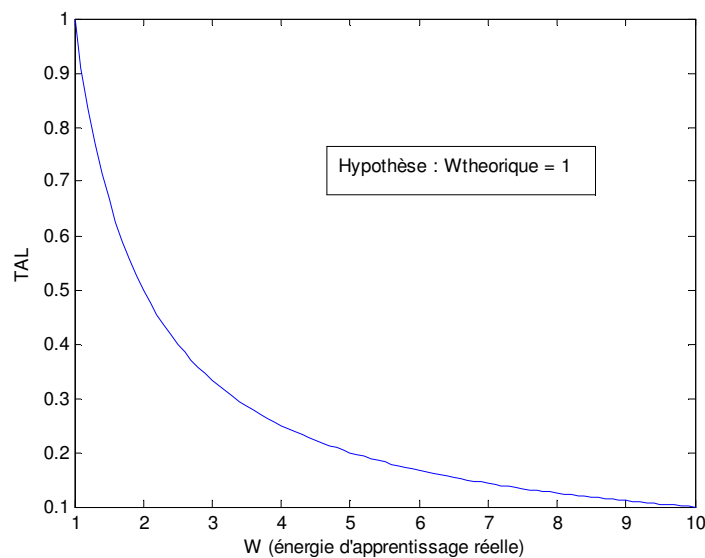


Figure 2-10 : variation du coefficient TAL

2.5.3. Conception de l'étude

Une étude exhaustive des 13 variables (2 pour la tâche, 4 pour le style et 7 pour le matériel) exige une population d'au moins 360000 élèves ! Ceci nécessiterait une organisation gigantesque qui n'est pas à notre portée.

Afin de faire une étude réaliste, nous avons opté pour un nombre restreint de variables. Les critères de choix des variables à retenir ont été les suivants :

- Représenter chacune des trois catégories cibles de l'étude : pédagogie, matériel et style d'apprentissage.
- Privilégier les dimensions des styles qui ont le plus de potentiel de varier significativement le matériel pédagogique.
- Opter pour les caractéristiques pédagogiques qui soient réalistes pour un niveau de cinquième.
- Choisir les caractéristiques permettant une création dans un temps raisonnable du matériel pédagogique.

Le premier critère se fonde sur l'hypothèse que si nous trouvons une causalité significative entre d'un côté le style et de l'autre côté quelques caractéristiques pédagogiques et du matériel, cette causalité restera valide si on étend le nombre de caractéristiques et de dimensions. Cette hypothèse ne remet pas en cause notre recherche car elle n'intervient qu'au moment de la conception du système. Ce même système étant adaptatif apprendra tout au long de son utilisation de nouveaux faits qui renforcent ses croyances.

En appliquant ces critères, nous avons retenu les variables du Tableau 2-4. La tâche pédagogique est décrite par 2 variables indépendantes : niveau de Bloom T_1 et type de l'activité T_2 . Le matériel pédagogique est décrit par les 2 variables contrôlées indépendantes : média M_1 et navigation M_2 . Le style est décrit par 2 variables indépendantes : axe sensoriel D_1 et axe de progression D_2 . Finalement, le comportement lui est décrit par les variables dépendantes performance P , TAL et U.

catégorie	Variable	Modalités
pédagogique	T_1	{Connaissance, Compréhension, Application}
	T_2	{Expositive, Active, Interrogative}
matériel	M_1	{Image, Texte}
	M_2	{Libre, Contraint}
Style	D_1	{Visuel, Verbal}
	D_2	{Séquentiel, Global}
comportement	P	[0,20]
	TAL	[0,1]
	U	[0,∞[

Tableau 2-4 : variables pour l'étude

Les élèves répondront une semaine avant au questionnaire RILS (Annexe 3). Sur la base des réponses obtenues, la population sera divisée en groupes couvrant les différents scénarios pédagogiques. Les élèves seront pris aléatoirement dans les 5 classes de cinquième.

Sur la base des variables de matériel et de style, nous utiliserons 4 groupes : 1 groupe expérimental et 3 groupes de contrôle (Tableau 2-5). Le groupe expérimental suivra une séquence pédagogique adaptée aux styles d'apprentissage. Les deux premiers groupes de contrôle suivront une séquence pédagogique uniquement adaptée à un des deux axes : visuel-verbal ou global-séquentiel. Quant au dernier groupe de contrôle, il suivra une séquence pédagogique inadaptée aux deux axes : visuel-verbal et global-séquentiel

Groupes	Visuel/Global	Visuel/Séquentiel	Verbal/Global	Verbal/Séquentiel
Expérimental	image/libre	image /contraint	texte/libre	texte/contraint
Contrôle 1	texte/libre	texte/contraint	image /libre	image /contraint
Contrôle 2	image /contraint	image /libre	texte /contraint	texte /libre
Contrôle 3	texte /contraint	texte /libre	image /contraint	image /libre

Tableau 2-5 : matériel pédagogique utilisé par groupe

En se basant sur les variables pédagogiques et de matériel, nous avons conçu un cours expérimental sur les nombres relatifs. Ce sujet a été choisi car il n'avait pas encore été étudié vu le programme et la période de l'année où intervenait l'étude. Nous minimisons ainsi l'effet de la connaissance à priori. La conception a suivi notre processus d'ingénierie d'objets d'apprentissage proposé dans la section 4.3.

La première étape fut de définir un ensemble d'objectifs d'apprentissage classés selon la taxinomie de Bloom (Figure 2-11).

Objectifs d'apprentissage	
1 - Connaissance	Identifier un nombre relatif Définir la valeur absolue d'un nombre relatif Définir l'opposé d'un nombre relatif
2 - Compréhension	Interpréter une expression arithmétique Permuter les termes d'une expression arithmétique Expliquer la valeur absolue d'un nombre relatif Expliquer l'opposé d'un nombre relatif
3 - Application	Simplifier la notation d'une expression arithmétique Calculer la valeur absolue d'un nombre relatif Calculer l'opposé d'un nombre relatif Additionner deux nombres relatifs Soustraire deux nombres relatifs Multiplier deux nombres relatifs Diviser deux nombres relatifs

Figure 2-11 : exemple d'objectifs d'apprentissage

Pour chaque objectif d'apprentissage, nous avons conçu différentes activités pédagogiques et avons créé le matériel pédagogique couvrant les différentes modalités des variables pédagogiques considérées. Les Figure 2-12 et Figure 2-13 montrent des exemples de matériel principalement graphique. La première correspond à un matériel graphique contraint ; la deuxième représente un matériel graphique libre.

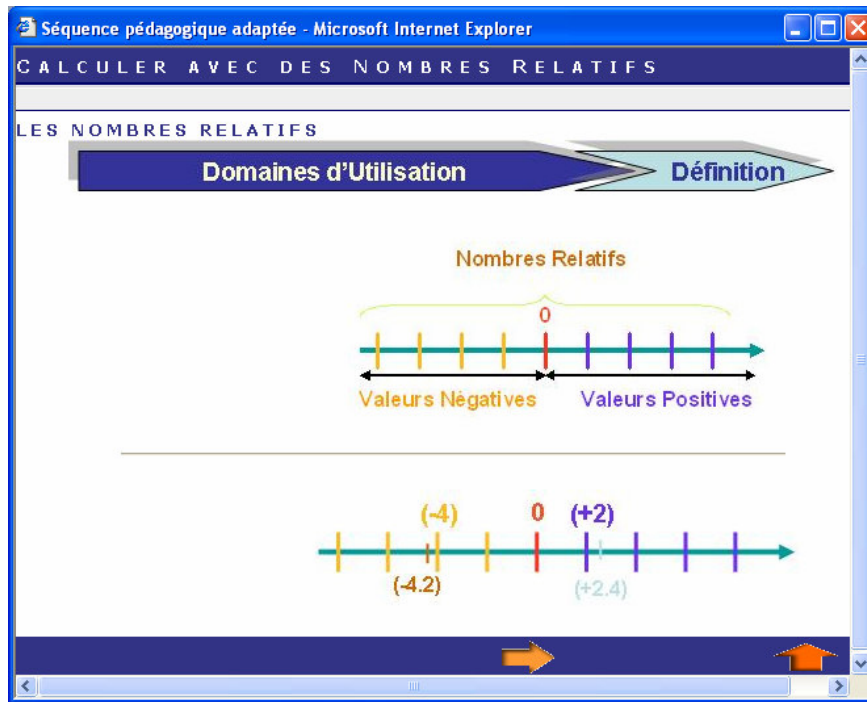


Figure 2-12 : matériel pédagogique graphique - contraint



Figure 2-13 : matériel pédagogique graphique - libre

Par ailleurs, les Figure 2-14 et Figure 2-15 montrent des exemples de matériels principalement textuel, respectivement contraint et libre.

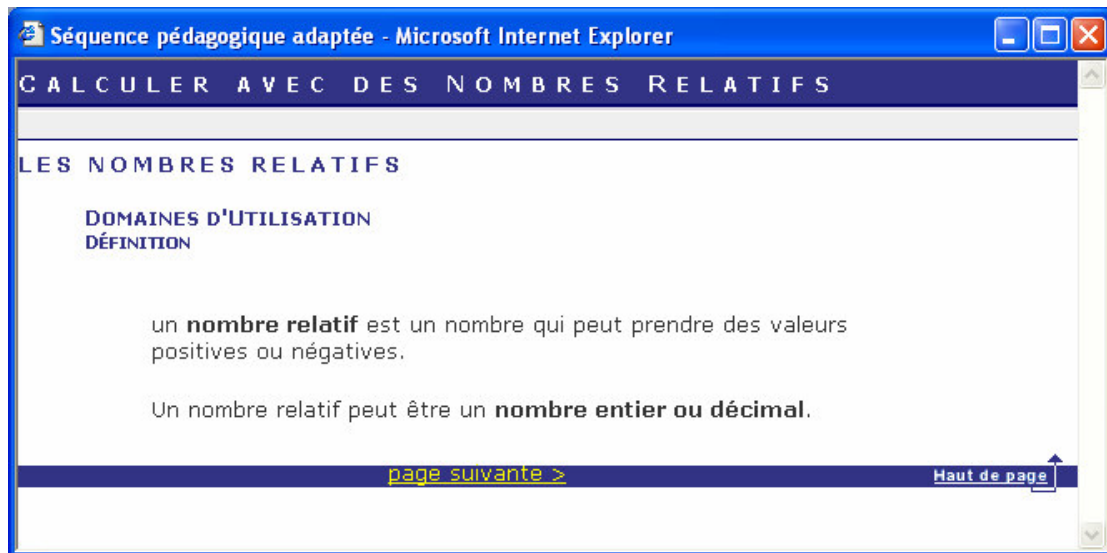


Figure 2-14 : matériel pédagogique textuel - contraint

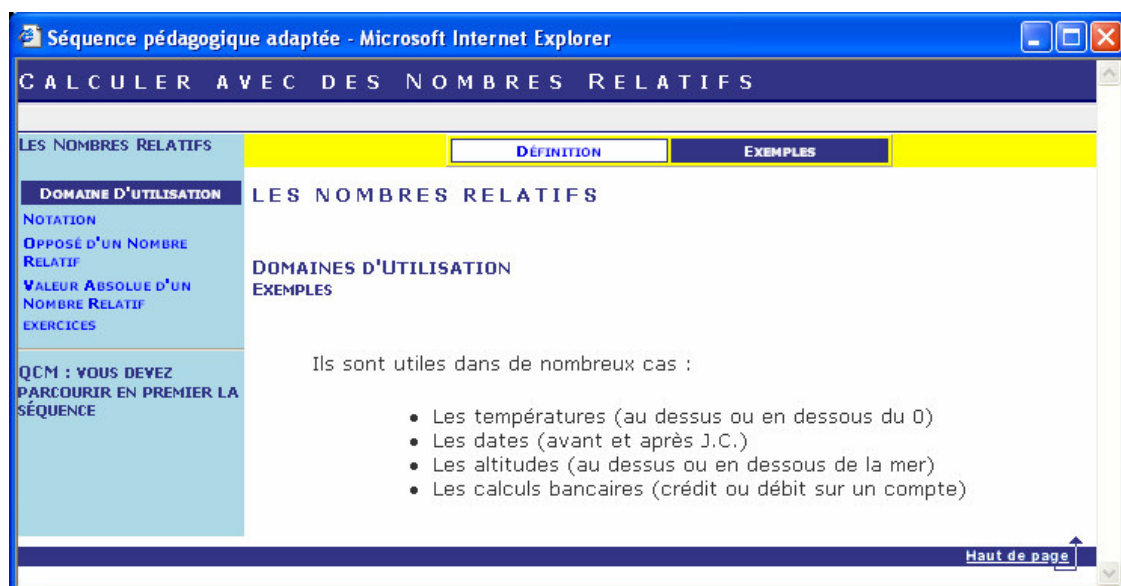


Figure 2-15 : matériel pédagogique textuel - libre

Évidemment, ceci a été fait en prenant en compte la combinaison des différentes dimensions du modèle de Felder (Tableau 2-5). Les activités pédagogiques ont toutes été conçues pour débiter avec un pré-test (Figure 2-16) et se terminer avec une évaluation sommative (Figure 2-17).

Séquence pédagogique adaptée - Microsoft Internet Explorer

CALCULER AVEC DES NOMBRES RELATIFS

QCM

PRETEST SUR LES NOMBRES RELATIFS

CONSIGNES
Cochez la réponse "a" ou "b" ou "c" pour chacune des questions.

Votre Sexe :
 (a) Féminin
 (b) Masculin

1 Le résultat d'une addition est <input type="radio"/> (a) une somme <input type="radio"/> (b) une différence <input type="radio"/> (c) un produit	2 Parmi les affirmations suivantes, laquelle est vraie ? <input type="radio"/> (a) -12.5 est un nombre entier relatif positif <input type="radio"/> (b) -4 est un nombre entier relatif négatif <input type="radio"/> (c) -12.5 est un nombre entier relatif négatif
3 L'opposé du nombre relatif -3 est <input type="radio"/> (a) -3 <input type="radio"/> (b) 1:3 <input type="radio"/> (c) +3	4 La somme (-2) + (-5) peut s'écrire <input type="radio"/> (a) 2 + 5 <input type="radio"/> (b) -2 + 5 <input type="radio"/> (c) -2 - 5
5 La différence $-(-9) - (+3)$ peut s'écrire <input type="radio"/> (a) $+9 + 3$ <input type="radio"/> (b) $9 - 3$ <input type="radio"/> (c) $-9 - 3$	6 La distance à zéro du nombre -34.25 est <input type="radio"/> (a) -34 <input type="radio"/> (b) 34.25 <input type="radio"/> (c) 0.25
7 $-17 + 21$ est égal à <input type="radio"/> (a) -38 <input type="radio"/> (b) 4 <input type="radio"/> (c) -4	8 $16 - 26$ <input type="radio"/> (a) est égal à -10 <input type="radio"/> (b) n'existe pas <input type="radio"/> (c) est égal à -42
9 $2.6 - (-3.6)$ est égal à <input type="radio"/> (a) -1 <input type="radio"/> (b) 5.12 <input type="radio"/> (c) 6.2	10 $-(-9) - 3$ est égal à <input type="radio"/> (a) -12 <input type="radio"/> (b) 6 <input type="radio"/> (c) -6

Suite de la séquence

Haut de page

Figure 2-16 : pré-test en début de séquence

Séquence pédagogique adaptée - Microsoft Internet Explorer

CALCULER AVEC DES NOMBRES RELATIFS

QCM

TESTEZ VOTRE CONNAISSANCE SUR LES NOMBRES RELATIFS

QUESTION 1
Parmi les affirmations suivantes, laquelle est vraie ?

-4 est un nombre entier relatif positif
 -4 est un nombre entier relatif négatif
 +4.2 est un nombre entier relatif positif

Question > 2/5

Haut de page

Figure 2-17 : évaluation sommative

2.5.4. Mise en œuvre de l'étude

L'étude s'est déroulée en deux phases principales. Nous avons commencé par mesurer le style d'apprentissage de chaque élève afin de former le groupe expérimental et les groupes de contrôle. Ensuite, chaque groupe a suivi la séquence pédagogique pendant laquelle nous avons effectué les mesures de performance et d'énergie d'apprentissage. A la fin de la séance, une tombola a été organisée. Il y avait en jeu 50 places de cinéma à gagner.

Mesure du style d'apprentissage

Les élèves ont répondu au questionnaire RILS quelques jours avant l'expérimentation. Le professeur de Mathématiques de chaque classe de cinquième, a distribué aux élèves une version papier du questionnaire. Ceci a permis la classification des élèves selon les quatre dimensions proposées par le modèle de Felder. Basé sur cette classification, la population a été divisée en groupes couvrant les différents scénarios pédagogiques. Chaque élève a reçu aléatoirement un numéro unique qui par la suite nous a permis de les identifier de façon anonyme. Ce nombre était celui tiré au sort lors de la tombola pour gagner les places de cinéma.

La séquence pédagogique

Le collège Malraux de Ramonville avait modifié l'emploi du temps des cinq classes de cinquième. L'expérimentation s'est déroulée le même jour et à la même heure, entre 9h00 et 10h00, dans cinq classes différentes disposées sur le même étage. Chaque élève disposait d'un ordinateur personnel pour jouer la séquence pédagogique à son rythme, en moyenne une trentaine de minutes. L'interface robuste de l'outil bloque certaines commandes. Par exemple, les commandes de retour et d'avance dans l'historique des actions sont indisponibles. De ce fait, un élève ne pourra se déplacer dans le contenu qu'en fonction des choix faits pour son groupe. Un enseignant était présent dans chacune des salles pour faire en sorte que l'expérimentation se déroule correctement.

En début de séance pendant le premier quart d'heure, chaque enseignant a expliqué à l'ensemble de la classe, comment utiliser l'outil qui présente le contenu pédagogique. Ensuite, l'enseignant a vérifié que chaque élève s'identifiait bien sur l'outil avec le bon numéro d'identification. Une fois la session ouverte, l'élève n'a plus reçu aucune aide.

En début de séquence avec l'outil, chaque élève a répondu à un pré-test pour évaluer son niveau de départ. Ensuite, pendant le déroulement de la séquence pédagogique, différentes mesures ont été faites (ex : le temps par page, le chemin suivi dans le contenu...). En fin de séquence chaque élève a répondu à un post-test qui reprend le contenu du pré-test et le complète par des questions liées aux nouvelles connaissances abordées pendant la séquence. L'ensemble des informations et des actions faites par l'élève a été stocké dans des fichiers de log. Ces données ont été filtrées avant d'être exploitées pour produire les résultats de l'étude. Un premier niveau de filtrage a été fait par l'analyse du contenu des fichiers de log qui contenaient des erreurs techniques d'enregistrement, donc inexploitable. Un second niveau de filtrage a été fait en utilisant le critère $W \geq W_{theorique}$. Nous gardons ainsi que les élèves qui ont eu un comportement cohérent lors de l'expérimentation.

2.5.5. Résultats de l'étude

Cette étude nous a permis d'obtenir différents résultats. Le premier résultat nous permet de montrer le taux de fiabilité des instruments de mesures utilisés dans l'expérimentation : le questionnaire ILS adapté, le pré-test et le post-test. Le second résultat permet d'identifier parmi les variables mesurées, celles qui expliquent correctement la variation de certains facteurs comme par exemple l'attention, le niveau de connaissance et le style d'apprentissage. Enfin, le troisième résultat montre que le comportement d'un élève varie bien en fonction du niveau d'adaptation de la séquence à son style d'apprentissage. Ces trois résultats sont détaillés dans les sous-sections suivantes.

Fiabilité et validité des questionnaires

L'analyse statistique plate des réponses au questionnaire RILS, nous donne la répartition suivante de la population (Tableau 2-6) :

Style	Nb élève	% population
Visuel	114	88%
Verbal	16	12%
Global	31	24%
Séquentiel	99	76%
Déductif	61	47%
Inductif	69	53%
Actif	85	65%
Réfléchi	45	35%

Tableau 2-6 : Répartition des Styles dans la population étudiée

On observe que les élèves visuels, séquentiels et actifs sont largement majoritaires dans cette population. En revanche la population est équitablement répartie entre les élèves inductifs et déductifs.

Sur la base de ces résultats, nous avons donc constitué le groupe expérimental de 33 élèves et les groupes de contrôle 1, 2 et 3 respectivement de 33, 33 et 31 élèves.

La nouvelle version simplifiée du questionnaire a obtenu un coefficient α de Cronbach=0.65 satisfaisant car supérieur à 0.50 [Felder et al., 2005b]. L'amélioration se situe essentiellement sur l'analyse factorielle (Tableau 2-7). On obtient un taux équivalent ou supérieur à la première version du questionnaire et chaque triplet explique le bon facteur. On augmente ainsi nos chances de placer les élèves dans les bons groupes en fonction de leur style d'apprentissage.

Facteur 1		Facteur 2		Facteur 3		Facteur 4	
D1 (inductif-déductif)		D2 (visuel-verbal)		D3 (actif-réfléchi)		D4 (séquentiel-global)	
Q6	0.54	Q7	0.53	Q9	0.22	Q12	0.15
Q10	0.65	Q3	0.73	Q5	0.45	Q4	0.41
Q2	0.70	Q11	0.85	Q1	0.99	Q8	0.65

Tableau 2-7 : Facteurs expliqués par le questionnaire RILS

Le pré-test a obtenu un coefficient α de Cronbach = -0.06. Ceci signifie qu'il ne peut être utilisé dans l'étude. Ce mauvais résultat est certainement dû à plusieurs facteurs : mauvais choix des questions, réponses aléatoires des élèves... Alors que le post-test a obtenu un coefficient α de Cronbach = 0.79 satisfaisant puisque au dessus du seuil fixé pour les mesures de connaissance ($\alpha = 0.75$)[Felder et al., 2005b].

Ces résultats nous permettent donc de penser que les dépendances que nous observerons entre le style d'apprentissage et la performance au post-test des élèves seront significatives.

Explication de certains facteurs

Lors de la séquence pédagogique, nous avons pour chaque élève, observé un certains nombres de variables (Tableau 2-8).

Variable	description
TAL	Adaptation de la tâche au style d'apprentissage
U	degré d'attention de l'élève lors de la séquence
T1	Niveau de Bloom (connaissance, compréhension, application)
M1	Type de support de l'information (image, texte)
M2	Mode de parcours des contenus (libre, contraint)
P	Note obtenue au post-test
Ng	Numéro de groupe
id	Numéro d'identification de l'élève

Tableau 2-8 : Variables observées lors de la séquence pédagogique

Une analyse factorielle (Tableau 2-9) de ces variables montre que le facteur 1, qui peut être considéré comme le niveau d'attention, est expliqué à 90% par la variable U. Le Facteur 3 que l'on peut voir comme le niveau de connaissance est expliqué à 99% par la variable P et les facteurs 2 (l'axe D1) et 4 (l'axe D2) expliquent respectivement les variables M1 à 98% et M2 à 88%.

	Facteur 1	Facteur 2	Facteur 3	Facteur 4
Variable	Attention	D1	Connaissance	D2
TAL	-0.8854	-0.0627	0.0097	0.0849
U	0.9048	0.0998	-0.1145	0.3916
T1	-0.0187	-0.0097	-0.3036	-0.0424
M1	0.0958	0.9843	-0.0188	-0.1290
M2	0.2095	-0.1660	-0.1054	0.8819
P	-0.0776	-0.0323	0.9899	-0.0891
Ng	0.1061	0.4688	-0.1665	0.3564
id	0.1545	-0.0107	-0.0317	-0.0207

Tableau 2-9 : Variation des variables mesurées lors de l'expérimentation

Il est à noter que la variable TAL s'oppose à la variable U. En effet, un TAL faible (séquence inadaptée) correspond à un degré d'attention (U) élevé. C'est-à-dire que l'élève doit fournir plus d'effort pour accomplir la tâche.

La Figure 2-18 montre la variation de la motivation à l'apprentissage (U) déployée par un élève pendant une séquence pédagogique en fonction du coefficient TAL. On observe que la fonction $U = f(TAL)$ est une fonction du type $1/x$. Cela conforte largement notre prédiction théorique.

La première remarque que nous pouvons faire sur ce graphique est que l'ensemble des mesures se trouve au dessus de $U=1$. Cela est dû au fait que le filtrage par le critère $W \geq W_{theorique}$ a éliminé l'ensemble des sujets considérés comme non représentatifs. Ces derniers ayant obtenu une valeur U inférieure à 1. Ensuite, à partir de ce constat, on observe que la valeur de U, grandit très rapidement quand le coefficient TAL s'approche de 0. Cela indique qu'un élève à qui on propose une séquence non adaptée à son style d'apprentissage devra faire preuve d'un plus grand effort pour suivre, à l'inverse d'un élève à qui on propose une séquence adaptée.

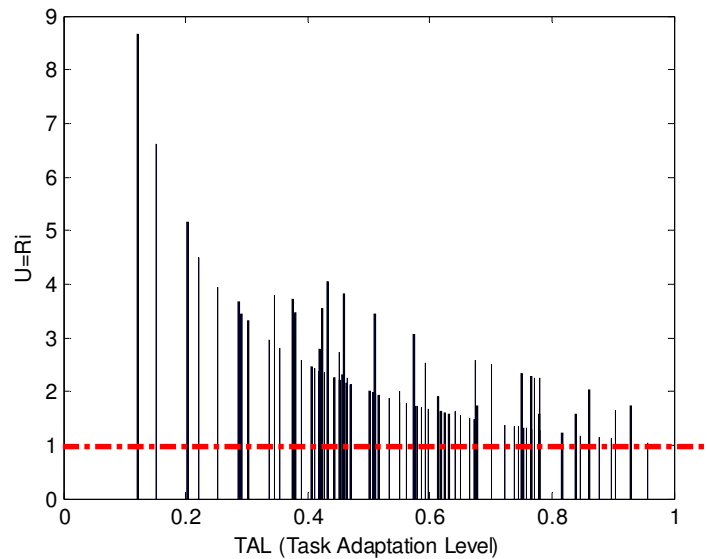


Figure 2-18 : Relation entre U et TAL

Dans la suite de cette thèse, nous exploiterons ce résultat pour évaluer si une séquence proposée à un élève était adaptée à son style d'apprentissage. Cela se fait par l'analyse continue du comportement de l'élève. Les évaluations faites, nous permettront d'affiner (renforcer) notre croyance quant au style d'apprentissage. Ce qui itérativement nous permet d'adapter les séquences pédagogiques.

Variation de la performance

Lors de la séquence pédagogique, nous avons mesuré pour chaque élève la performance obtenue au post-test. La Figure 2-19 montre la variation de la moyenne des notes pour chaque groupe ayant participé à l'expérimentation. Le groupe expérimental (Exp) a suivi une séquence adaptée aux styles d'apprentissage des élèves qui le composent. Le groupe de contrôle 3 (Ctrl3) a suivi une séquence dont le contenu pédagogique était inadapté sur les axes sensoriel (Visuel-Verbal) et de progression (Global-Séquentiel). Le groupe de contrôle 1 (Ctrl1) avait une séquence inadaptée uniquement sur l'axe sensoriel, alors que le groupe de contrôle 2 (Ctrl2) avait une séquence inadaptée sur l'axe de progression.

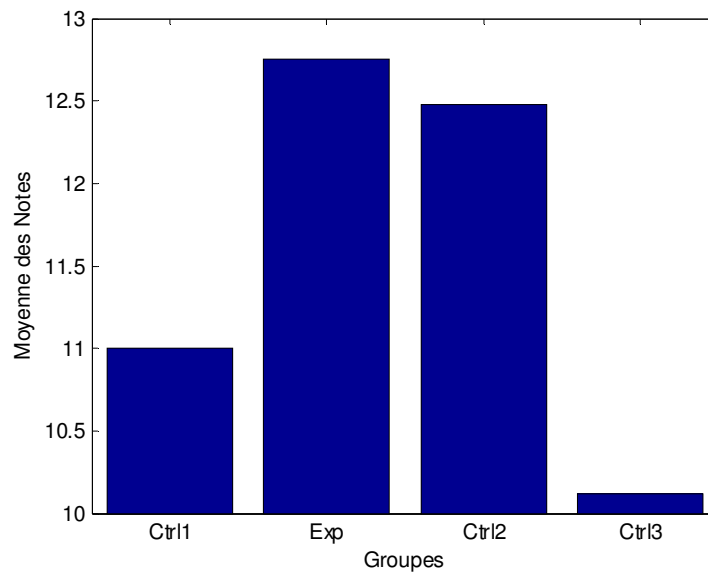


Figure 2-19 : Variation de la note du post-test par groupe

Le résultat obtenu montre que le groupe Exp a le meilleur score (note moyenne) et le groupe Ctrl3 le moins bon. Les moyennes des groupes Ctrl1 et Ctrl2 se trouvent dans l'intervalle des moyennes définies par les groupes Exp et Ctrl3. Le groupe Ctrl1 faisant moins bien que le groupe Ctrl2. Ce résultat relatif entre les groupes valide clairement l'hypothèse que l'adaptation influence la performance. En effet, les élèves ayant obtenu les meilleurs résultats en moyenne ont suivi une séquence pédagogique adaptée à leur style d'apprentissage. Les élèves ayant le moins bien réussi, ont suivi une séquence inadaptée. Les élèves ayant obtenu des notes intermédiaires ont travaillé avec une séquence inadaptée sur un axe uniquement. Les élèves dont la séquence était inadaptée sur l'axe sensoriel ont moins bien réussi que ceux pour qui l'axe inadapté était celui de la progression. Néanmoins, un fait curieux est à remarquer. L'ensemble des groupes a une moyenne au dessus de 10/20.

L'analyse des fichiers de log créés à chaque session utilisateur nous apprend que la robustesse de l'outil a été mise à mal par certains élèves. Certaines sessions ont été fermées puis rouvertes plusieurs fois. Des traces d'événements montrent une navigation dans l'historique des actions du navigateur pourtant interdites. Nous complétons cette analyse par l'étude du critère de l'énergie d'apprentissage $W \geq W_{theorique}$. Chaque élève qui ne répond pas au critère de l'énergie d'apprentissage dénote une activité anormale pendant la séquence. Nous décidons d'éliminer les résultats de ces élèves des données utilisées pour étudier le comportement des groupes d'élèves. Le Tableau 2-10 présente le pourcentage de résultats ignorés par groupe.

	Exp	Ctrl1	Ctrl2	Ctrl3
Anomalies fichiers de logs	0%	0%	8%	20%
$W \geq W_{theorique}$	2%	10.7%	8%	60%
Total	2%	10.7%	16%	80%

Tableau 2-10 : Pourcentage de résultats éliminés par groupe

Les pourcentages que nous observons montrent clairement que le groupe de contrôle Ctrl3 a le plus fort taux de données erronées. On en déduit que les élèves de ce groupe ont eu un comportement anormal à cause de la mauvaise association de leur style d'apprentissage avec le contenu de la séquence. On imagine que voulant malgré tout réussir le test, ils ont été rudes avec l'outil mettant à mal sa robustesse et qu'ils ont été pour certains poussés à copier sur le voisin les réponses au post-test. Leurs résultats faussant donc les différents calculs : énergie d'apprentissage, performance, etc. Ce fait est un autre moyen (indirect) pour conforter l'hypothèse H_0 .

Considérant que l'étape précédente réalise une sorte de filtrage des données, il paraît intéressant d'observer à nouveau l'évolution des moyennes des notes obtenues au post-test par les différents groupes (Figure 2-20).

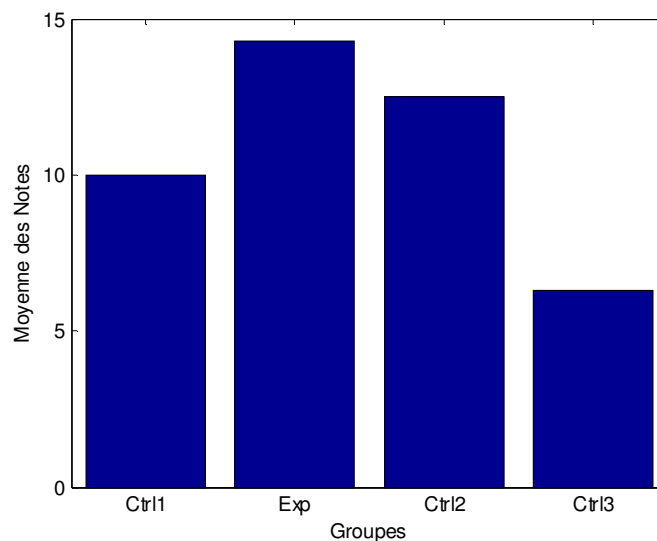


Figure 2-20 : Variation de la note du post-test par groupe après filtrage

On note clairement, cette fois-ci, que la moyenne (14.3/20) du groupe Exp est supérieure à l'ensemble des groupes. Le groupe Ctrl3 obtenant comme attendu, la moins bonne performance (6.3/20) et maintenant largement en dessous du 10/20. On peut déduire aussi de cette courbe que les élèves qui ont eu une séquence inadaptée sur l'axe sensoriel (Visuel-Verbal) ont moins bien réussie (10/20) que ceux qui avaient une séquence inadaptée sur l'axe de progression (Global-Séquentiel) (12.5/20). On peut l'expliquer par le fait qu'un élève qualifié de séquentiel sera peu gêné par un contenu qui offre une progression libre. En effet, l'élève peut dans ce cas suivre une progression séquentielle en développant successivement les branches de l'arbre présentant le matériel pédagogique. On note une forte variation entre les moyennes du groupe Exp et Ctrl3 : 8 points.

Cette mesure a été affinée en observant la variation des moyennes par groupe sur les trois premiers niveaux de Bloom (Figure 2-21) : connaissance, compréhension et application.

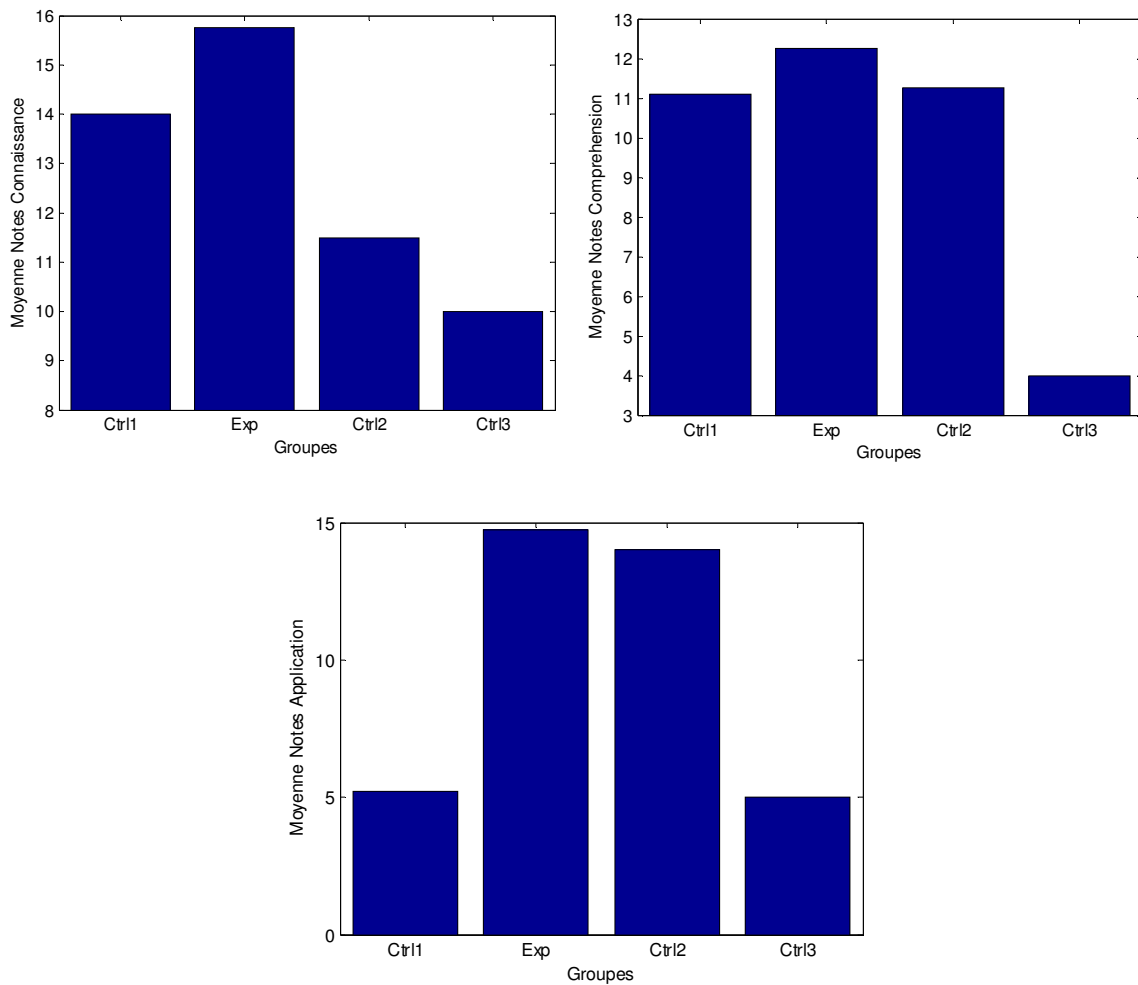


Figure 2-21 : Variation de la note du post-test par groupe et niveaux de Bloom

On montre clairement que le groupe Exp obtient le meilleur résultat sur les trois niveaux de Bloom. On peut de plus ajouter que le manque d'adaptation est moins pénalisant sur le niveau connaissance puisque les trois groupes de contrôle ont une moyenne au dessus de 10/20. Le niveau applicatif pouvant être considéré comme le plus sensible à l'adaptation de la séquence puisque la moyenne des groupes Ctrl1 et Ctrl3 se situe 5 points en dessous de 10/20 et avec une variation de 10 points par rapport aux résultats du groupe Exp.

L'analyse de la variation de la variable TAL (niveau d'adaptation de la tâche au style d'apprentissage) montre bien que le groupe Ctrl3 avait une séquence inadaptée, alors que le groupe Exp obtient le meilleur score (Figure 2-22). On peut noter aussi que le groupe Ctrl1 a un TAL inférieur au Ctrl2. Ceci exprime que le TAL est plus sensible au manque d'adaptation sur l'axe sensoriel que sur l'axe de progression. Ceci est en parfaite adéquation avec le résultat sur les performances. Il s'agit de deux variables qui varient selon le même mode sous l'influence du style.

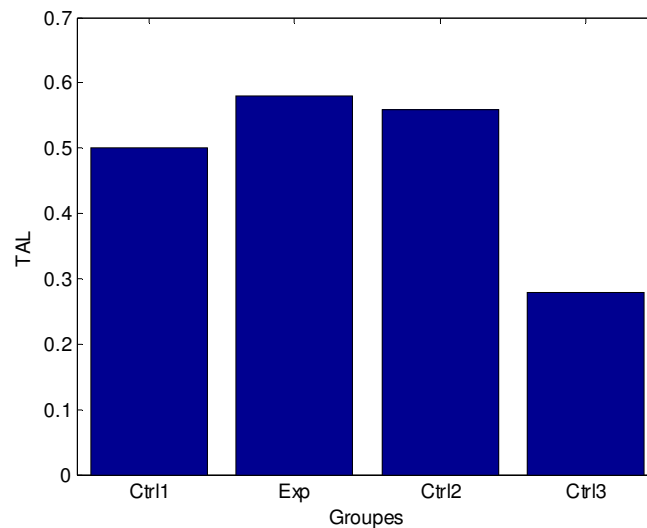


Figure 2-22 : Variation du TAL par groupe

L'analyse de la variation de la variable U (effort d'attention) par groupe montre parfaitement que les élèves du groupe Ctrl3 ont dû fournir un plus gros effort d'attention pendant la séquence que les élèves du groupe Exp (Figure 2-23). Ceci est un résultat attendu puisque le groupe Ctrl3 avait une séquence inadaptée aux styles des élèves. On observe de plus que les élèves du groupe Ctrl1 ont déployé un effort supérieur à ceux des groupes Exp et Ctrl2. Cela indique que l'inadaptation sur l'axe sensoriel perturbe plus que l'inadaptation sur l'axe de progression. Il est aussi à remarquer la faible différence entre les groupes Exp et Ctrl2. Cela confirme que l'inadaptation sur l'axe de progression produit peu d'effet négatif.

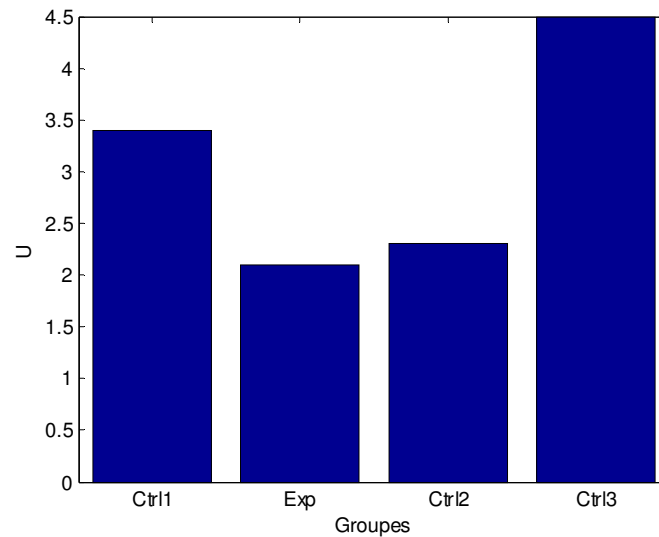


Figure 2-23 : Variation de l'effort d'un élève par groupe

Les résultats que nous avons obtenus prouvent bien que l'ajustement de la démarche pédagogique au style d'apprentissage de l'élève, influence son comportement lors de la séquence d'enseignement. Les indicateurs que nous avons utilisés dans les mesures peuvent donc être pris sérieusement en considération dans le modèle du processus de profilage du style d'apprentissage d'un élève.

2.6. Variables de profilage du style d'apprentissage

En se basant sur les résultats de notre étude empirique, nous présentons dans cette section une instance concrète du profilage du style d'apprentissage. Les variables observables et non observables découvertes et validées dans notre étude empirique sont complétées par un ensemble d'indicateurs apportant une description plus fine du matériel et de l'activité pédagogique.

2.6.1. Variables non observables : Le style d'apprentissage de l'élève

Cette théorie suppose que le style d'apprentissage de l'élève peut être estimé par une mesure de croyance projetée sur quatre dimensions [Felder et al, .2005b]. En se basant sur notre étude présentée dans le paragraphe précédent, l'ensemble des variables décrivant le style d'apprentissage de l'élève se compose de quatre éléments :

$$\mathcal{D} = \{D_1, D_2, D_3, D_4\}$$

$D_i, i \in [1,4]$, sont respectivement les dimensions sensorielle, progression, réflexion et raisonnement. Afin de garder une cohérence avec le codage de notre instrument de mesure du style d'apprentissage (RILS), chacune des variables $\{D_i\}_{i \in [1,4]}$ prend une valeur entière comprise entre -3 et 3.

$$D_i \in [-3, 3] \quad \forall i \in \{1, 2, 3, 4\}$$

2.6.2. Variables observables

Les variables observables caractérisent le matériel pédagogique utilisé dans la séquence d'apprentissage, l'activité pédagogique et le comportement de l'élève. Ces trois catégories de variables sont présentées dans les sous-sections suivantes. Nous utilisons les résultats de notre étude pour définir le comportement de l'élève sur la base de la mesure de performance, de la mesure de l'effort d'apprentissage ainsi que du niveau d'adaptation de la tâche (TAL). Nous considérons ces données comme observables car il est possible de capturer par exemple le type de parcours que fait l'élève dans une séquence (séquentiel : suivi des chapitres dans l'ordre, ou bien global : parcours aléatoire des chapitres). De même, il est aussi possible de connaître le niveau de difficulté de la tâche à accomplir au sens de Bloom ainsi que le type de connaissances abordées (concept, fait, processus, etc.).

Le matériel pédagogique

Nous adoptons sept métadonnées pour décrire le matériel pédagogique [Piombo & al, 2005].

$$\mathcal{M} = \{M_1, M_2, M_3, M_4, M_5, M_6, M_7\}$$

Ces sept variables ont la signification suivante :

- M_1 représente les médias principaux utilisés dans le matériel pédagogique. Trois possibilités sont considérées : majoritairement texte probablement illustré avec des images (T), majoritairement graphiques et images accompagnés avec des explications textuelles (I), voix prédominante avec des graphiques d'illustration (V).

$$M_1 \in \{T, I, V\}$$

- M_2 représente la structure externe du contenu. Elle prend les valeurs : F pour la navigation libre ou C pour la navigation contrainte.

$$M_2 \in \{C, F\}$$

- M_3 est une variable booléenne qui montre si le matériel contient des résumés et des aides de navigation.

$$M_3 \in \{vrai, faux\}$$

- M_4 est également une variable booléenne qui est vraie quand le matériel contient des définitions et de la théorie. Elle est fausse quand le matériel consiste seulement en travaux pratiques, exemples, exercices, etc.

$$M_4 \in \{vrai, faux\}$$

- M_5 est une variable booléenne qui montre si le matériel contient des exemples.

$$M_5 \in \{\text{vrai}, \text{faux}\}$$

- M_6 représente le type d'exercices pratiques proposé dans le matériel : questions vrai ou faux (TF), questions à choix multiples avec seulement une réponse possible (QCM1), questions à choix multiples avec plusieurs réponses possibles (QCMN), déplacer et déposer (DD), texte à trous (FS), pas d'exercice (N).

$$M_6 \in \{TF, QCM1, QCMN, DD, FS, N\}$$

- M_7 est semblable à M_6 mais ajoute la notion d'évaluation. Il décrit le type de questions utilisées pour examiner le niveau de connaissances de l'élève.

$$M_7 \in \{TF, QCM1, QCMN, DD, FS, N\}$$

L'activité pédagogique

L'activité pédagogique est décrite par l'objectif d'apprentissage (T_1) cible, le type d'activité (T_2) et la nature de la connaissance (T_3) à acquérir [Piombo et al, 2005].

$$\mathcal{T} = \{T_1, T_2, T_3\}$$

Pour représenter l'objectif d'apprentissage, nous adoptons la taxinomie proposée par Bloom [Bloom, 1956]. Ce modèle déclare qu'une activité pédagogique peut viser un des six niveaux d'objectif d'apprentissage : la connaissance (1), la compréhension (2), l'application (3), l'analyse (4), la synthèse (5), et l'évaluation (6). Basé sur cette approche, notre variable T_1 peut prendre une valeur entière entre 1 et 6.

$$T_1 \in \{1, 2, 3, 4, 5, 6\}$$

T_2 indique le type de l'activité pédagogique proposée à l'élève. Trois types différents sont à considérer : expositive (1), active (2) et interrogative (3).

$$T_2 \in \{1, 2, 3\}$$

T_3 indique la nature de la connaissance traitée dans l'activité. Cinq types différents sont à considérer : concept (1), fait (2), processus (3), principe (4), et procédé (5).

$$T_3 \in \{1, 2, 3, 4, 5\}$$

Un concept est une idée ou une représentation de l'esprit exprimée par un terme (ex : mammifère, ordinateur). Un fait est représenté par des données observables (ex : la température de l'eau). Un processus désigne un ensemble de phases ou d'états d'une opération (ex : la corrosion). Un principe définit une règle ou une loi (ex : conservation de l'énergie). Un procédé est une activité humaine de fabrication d'un produit (ex : vinification)

Le comportement de l'élève

Nous considérons que le comportement d'un élève ne peut être réduit uniquement à la mesure d'une performance. En effet, un élève au maximum de sa forme, produirait une performance optimale quand le contexte d'apprentissage lui est adapté (niveau d'objectif, matériel pédagogique utilisé,.....). Or, dans une situation réelle un élève peut réussir moins bien que ce qu'il était permis d'espérer. Ceci peut, évidemment, être causé par des perturbations externes comme la motivation, la santé, l'environnement... Pour mesurer ces efforts, nous adoptons le triplet de variables performance, TAL et U. Cette approche est pleinement justifiée par les résultats de notre étude empirique (présentée plus haut). Ces variables ont en effet été prouvées pertinentes pour expliquer le comportement.

$P^m \in [0, 20]$, mesure la performance réelle obtenue par l'élève à une évaluation sommative en fin de séquence.

$TAL \in [0, 1]$, mesure le niveau d'adaptation de la tâche au style d'apprentissage de l'élève.

$U \in [0, \infty[$, mesure le degré d'attention d'un élève pendant la séquence pédagogique.

Le modèle de dépendance ainsi établi sera le fondement de notre modèle probabiliste des préférences de l'élève.

2.7. Conclusion

Dans ce chapitre nous avons proposé un méta-modèle conceptuel du processus de profilage. Nous l'avons en effet défini comme étant la transformation de données observables en variables non observables. Nous nous sommes servis de ce méta-modèle pour concevoir un modèle de dépendance du style d'apprentissage ILS de l'élève.

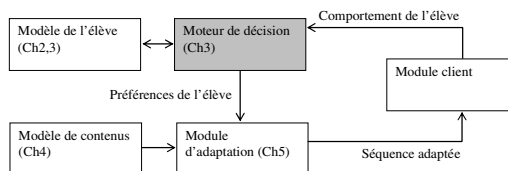
Le modèle ainsi construit se base sur deux hypothèses psycho-cognitives que nous avons émises. La première stipule que la complexité de la tâche d'apprentissage ainsi que le style d'apprentissage de l'élève déterminent la séquence pédagogique permettant d'optimiser la performance. La deuxième suppose que la complexité de la tâche, le style de l'élève, ainsi que les caractéristiques de la séquence pédagogique conjointement influencent le comportement de l'élève lors de l'apprentissage.

Afin de valider ces hypothèses, et donc notre modèle, nous avons mis en œuvre une étude empirique. Celle-ci s'est composée de deux parties. La première a consisté à valider notre questionnaire de mesure de l'ILS. Des analyses statistiques notamment le coefficient α de Cronbach et factorielles, ont été réalisées. Celles-ci ont amené à une révision du questionnaire le rendant plus court et plus cohérent. La deuxième partie de l'étude a consisté à valider les dépendances entre les dimensions de l'ILS d'un côté, et de l'autre côté les variables observables caractérisant la séquence pédagogique et le comportement de l'élève. Pour cette étude, en plus de la performance, nous avons proposé trois grandeurs originales pour mesurer le comportement de l'élève : l'énergie d'apprentissage, le degré d'attention et le niveau d'adaptation de la tâche. Nous avons établi leurs modes de calcul. Nous avons aussi proposé un ensemble de métadonnées, pédagogiquement pertinent, pour caractériser une séquence pédagogique. Celles-ci incluent le niveau de Bloom, la nature de l'activité, la structure interne des documents, l'organisation des ressources, ainsi que les médias utilisés. Par ailleurs, quatre variables représentant les dimensions de l'ILS ont été adoptées pour représenter le style de l'élève. Les analyses statistiques effectuées ont clairement établi que :

- Les grandeurs proposées pour mesurer le comportement de l'élève sont valides et pertinentes.
- L'adaptation de la méthode et du contenu pédagogique à l'ILS influence le comportement de l'élève lors de la séquence pédagogique.
- Et qu'il y a une dépendance directe entre le style et la séquence optimale.

Partant de ces résultats, nous avons généralisé les métadonnées décrivant la séquence pédagogique. Ces nouvelles variables seront utilisées dans notre système mais restent à valider statistiquement.

Nous avons aussi défini les variables mesurant le style et le comportement, et posé leurs notations. Cet ensemble de variables nous servira dans la conception de notre système adaptatif. Cela va notamment servir à créer le modèle probabiliste bayésien pour estimer le style et le moteur de décision (chapitre 3), l'indexation des contenus pédagogiques (chapitre 4), ainsi que la sélection des séquences adaptées (chapitre 5).



CHAPITRE

3. Apprentissage supervisé du modèle de l'élève

Résumé :

La majorité des travaux de recherche sur les systèmes d'enseignement adaptatifs se sont concentrés sur une modélisation des élèves fondée sur leur niveau de connaissance et leurs préférences.

Dans ce chapitre nous présentons un modèle de préférences qui est utilisé pour adapter le contenu au style d'apprentissage de l'élève. Ce modèle est une application du modèle conceptuel du processus de profilage que nous avons proposé au chapitre 2.

Pour implanter le modèle, nous proposons deux solutions inspirées de la théorie statistique de l'apprentissage. La première est organisée autour de l'utilisation d'un réseau bayésien alors que la seconde exploite les machines à vecteurs de support.

3.1. Introduction

La majorité des travaux de recherche sur les systèmes d'enseignement adaptatifs se sont concentrés sur une modélisation des élèves fondée sur leur niveau de connaissances et leurs préférences [Ross et al, 1987 ; Zhou et al, 1999]. En parallèle, plusieurs études expérimentales dans la psychologie différentielle ont montré que les différences individuelles jouent un rôle important dans l'enseignement [Zhang, 2002 ; Jonassen et al, 1993 ; Felder, 1993 ; Grigorenko et al, 1995]. Les élèves acquièrent et traitent les informations pédagogiques de différentes manières; ils apprennent à des rythmes différents suivant le contexte d'apprentissage [Barbe et al, 1981 ; Corno et al, 1986]. Ces résultats laissent suggérer que l'adaptation des séquences pédagogiques aux styles d'apprentissage permet à l'élève de retenir l'information plus longtemps et de l'appliquer plus efficacement [Dunn, 1999 ; Riding et al, 1999].

Quelques chercheurs ont envisagé d'incorporer les styles individuels dans la conception de systèmes d'enseignement [Chen et al, 2002 ; Dufresne et al, 1997 ; MacGregor, 1999 ; McLoughlin, 1999 ; Paredes et al, 2002 ; Souto, 2002 ; Moallem, 2002]. Dans [Paredes et al, 2002], les auteurs considèrent la dimension raisonnement (inductif/déductif) du style d'apprentissage pour adapter l'enchaînement des définitions et exemples dans le système TANGOW. Dans [Moallem, 2002], l'auteur a utilisé les styles d'apprentissage comme fondement de son système d'enseignement. Elle utilise les dimensions du style d'apprentissage (ILS) de Felder-Silverman [Felder et al, 1988] comme base pour concevoir et structurer les contenus pédagogiques. Chacun des huit styles d'apprentissage est pris en considération. Le contenu du cours incorpore des supports conçus pour chaque dimension. Avant la présentation, les élèves répondent au questionnaire de Felder-Soloman pour déterminer leur style d'apprentissage. Des supports pertinents leur sont alors présentés. Dans [Souto, 2002], les auteurs utilisent le test de Ross pour évaluer le style cognitif d'apprentissage (CLS). Ils analysent la navigation dans un cours expérimental sur le Web pour identifier le comportement des élèves dans la phase d'apprentissage. Les données des deux expériences sont corrélées pour établir cinq classes de CLS; chacune caractérisée par un certain comportement pendant la navigation.

Dans ce chapitre, nous présentons une approche qui utilise un modèle de l'élève scindé en un modèle de connaissances et un modèle de préférences. Nous commençons par présenter l'architecture du système adaptatif utilisant le modèle de l'élève. Ensuite, après avoir introduit le modèle de connaissances, nous étudions plus particulièrement le modèle de préférences qui est utilisé pour adapter le contenu au style d'apprentissage. Ce modèle de préférences est une application du modèle conceptuel du processus de profilage que nous avons proposé (cf. paragraphe 2.2). Nous proposons ensuite deux solutions inspirées de la théorie statistique de l'apprentissage pour la mise en œuvre de notre modèle. La première est organisée autour de l'utilisation d'un réseau bayésien alors que la seconde exploite des machines à vecteurs de support.

3.2. Système adaptatif proposé

Habituellement, les systèmes d'enseignement intelligents et adaptatifs se composent de quatre constituants principaux [Kay, 2001] : modèle de contenus, modèle de l'élève, modèle d'adaptation, et module d'interface. On peut classer leurs architectures en deux catégories : multicouches et multi-agents [Kobsa et al, 2001]. Les systèmes multicouches sont centralisés; leur caractéristique principale est la structuration des contenus pédagogiques [Cristea et al, 2002b]; les systèmes multi-agents, quant à eux, sont généralement distribués et utilisent des agents locaux pour la modélisation des élèves et l'adaptation [Razek et al, 2002]. Cependant, de nombreux systèmes combinent ces deux approches. Ils emploient comme architecture globale celle des systèmes multi-agents et pour les différents modèles des architectures multicouches.

Le système proposé se base sur l'apprentissage statistique du modèle de l'élève [Batatia et al., 2006]. Il a une architecture mixte multi-agents et multicouches (Figure 3-1). Il se compose du modèle de l'élève, du modèle de contenus, d'un moteur de décision, d'un module d'adaptation, d'un module de présentation ainsi que d'un agent collecteur.

Le module de présentation est responsable de jouer le contenu adapté et d'interagir avec l'élève. L'agent collecteur surveille les interactions et mesure le comportement de l'élève : sa performance (P^m), son degré d'attention (U) et le niveau d'adaptation de la tâche (TAL)

Le modèle de l'élève maintient les connaissances et les préférences. Le moteur de décision est composé de deux étages : un séquenceur d'objectifs et un estimateur de style d'apprentissage. Le premier utilise le modèle de connaissances et la performance (P^m) obtenue par l'élève. Il détermine l'objectif pédagogique suivant à enseigner. Il prédit aussi la performance théorique (P^T) que l'élève peut réaliser si la séquence pédagogique est optimalement adaptée.

La complexité de l'objectif fixé (T_1) ainsi que les indicateurs de comportement de l'élève (TAL, U, P^m) sont les données en entrée du deuxième étage. Il utilise aussi le modèle de préférences. Il renforce l'estimation du style d'apprentissage et infère les métadonnées de la séquence idéale à proposer à l'élève. Ces métadonnées sont alors fournies au module d'adaptation (cf. chapitre 5) qui utilise le modèle de contenus (cf. chapitre 4) pour sélectionner la séquence pédagogique optimale et adapte le contenu.

Dans cette thèse, nous ne traitons pas dans le détail le modèle de connaissances et le séquenceur d'objectif. Nous nous concentrons sur le modèle de préférences, l'estimateur du style, le modèle de contenus et le module d'adaptation. Ce chapitre décrit le moteur de décision et le modèle de l'élève.

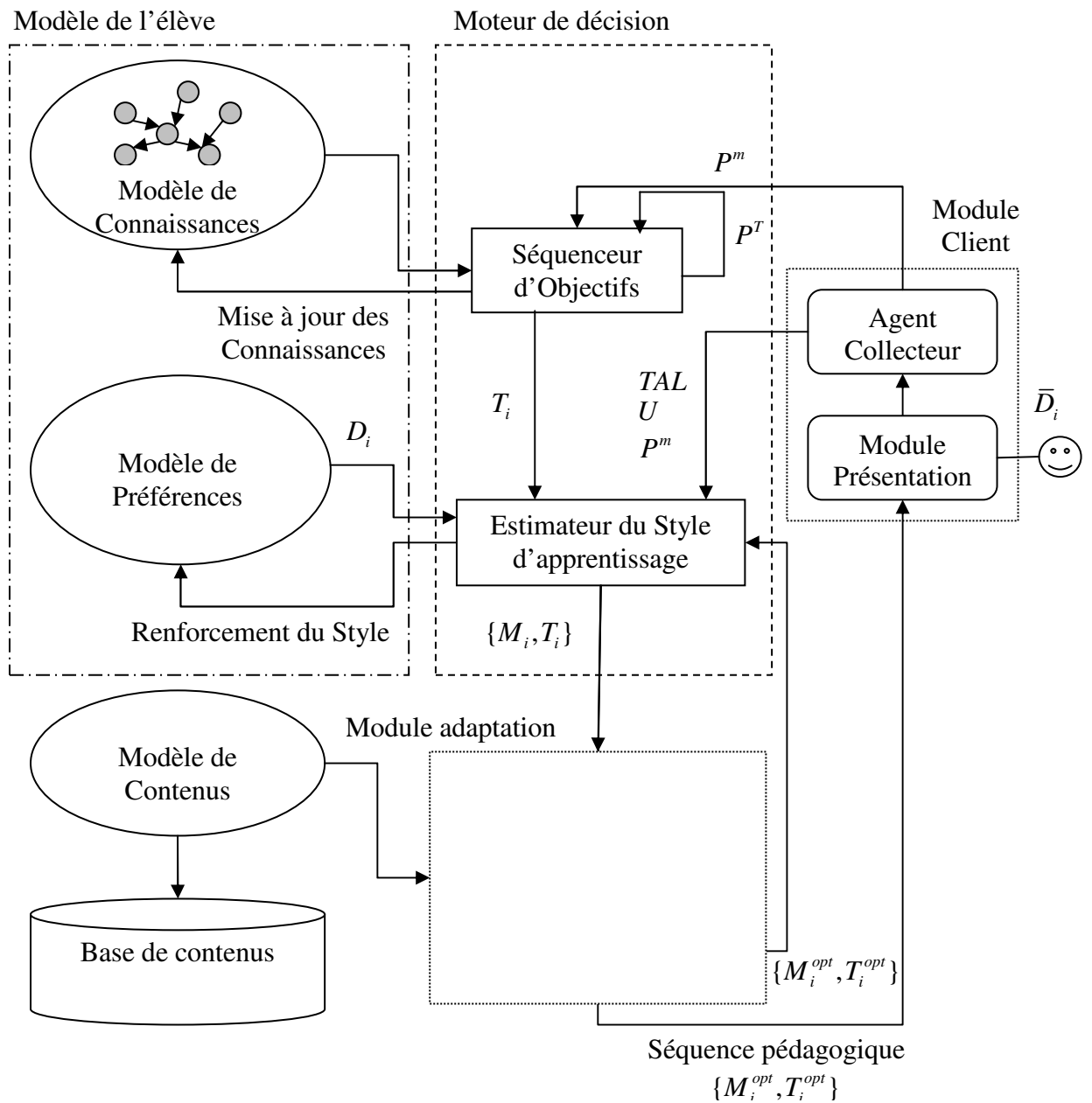


Figure 3-1 : architecture du système proposé

3.3. Modèle de l'élève

Les modèles d'élèves sont habituellement utilisés pour représenter leurs connaissances et leurs caractéristiques comportementales et cognitives. En dépit de ce but commun, les modèles varient considérablement dans leurs contenus. La plupart d'entre eux se focalisent sur les connaissances et les compétences, bien que quelques chercheurs aient plaidé pour que d'autres caractéristiques soient également traitées [Vassileva, 1990]. Dans les systèmes existants, nous pouvons distinguer trois finalités à l'utilisation de modèles d'élèves : l'identification des chemins empruntés par un élève dans la résolution de problèmes [Conati et al, 1997]; l'évaluation des performances et compétences de l'élève [Katz et al, 1993]; le diagnostic des erreurs et des violations de contraintes par l'élève [Mayo et al, 2000]. Dans notre cas, le modèle de l'élève est utilisé pour ajuster la démarche pédagogique au niveau de connaissance et au style d'apprentissage.

Le modèle de l'élève que nous proposons, se compose d'un modèle de connaissances et d'un modèle de préférences. Le premier est utilisé pour prédire, le prochain objectif pédagogique à étudier (T_i), en fonction du niveau de connaissances actuel. Le deuxième spécifie les contenus pédagogiques (M_i) adaptés au style d'apprentissage de l'élève (D_i) pour l'objectif visé (T_i).

Les sections 3.4 et 3.5 présentent respectivement ces deux modèles. Dans les sections 3.6 et 3.7 nous proposons deux implémentations différentes de l'estimateur de style d'apprentissage qui exploite le modèle de préférences. La première utilise un réseau bayésien. La seconde met en œuvre les machines à vecteurs de support.

3.4. Modèle de connaissances

La littérature regorge de travaux de recherche sur la modélisation des connaissances de l'élève. Dans [Gamboa et al, 2001 ; Carberry, 2001; Horvitz et al, 1998 ; Conati et al, 1997], les auteurs utilisent des réseaux bayésiens pour analyser le niveau de connaissance de l'élève et pour prédire les actions futures.

Inspiré de ces travaux, le modèle de connaissance que nous proposons, est un modèle overlay sur le modèle de contenus (cf. chapitre 4). Cela signifie que l'ensemble de la structure du modèle de contenus est représentée dans le modèle de connaissances et décrite sous la forme d'objectifs pédagogiques à atteindre. Il indique ainsi, pour chaque objectif pédagogique, le niveau de connaissance de l'élève. En effet, on associe à chaque objectif pédagogique déjà étudié la performance mesurée P^m , que l'élève a obtenue à l'évaluation sommative de la dernière séquence pédagogique correspondante. Nous disposons par ailleurs d'un modèle de prédiction nous permettant d'estimer la performance théorique P^T envisageable pour un élève dans un objectif donné en fonction de ses connaissances actuelles dans le domaine. L'objectif sera recommandé uniquement quand sa P^T est supérieure à un seuil s fixé. Dans notre cas, s est pris égal à 10.

Notre modèle de connaissance est un modèle probabiliste qui utilise un réseau bayésien. Il nous est donc possible d'inférer le prochain objectif pédagogique à proposer à l'élève en ayant observé un certain nombre de P^m dans des objectifs dépendants. On définit la mesure de niveau d'acquisition d'un objectif pédagogique comme une fonction des performances théorique P^T et réelle P^m . Cette fonction prend la forme suivante :

$$\Upsilon = h(P^T, P^m) = \begin{cases} 1 & \text{si } P^m \geq P^T \\ \frac{P^m}{P^T} & \text{sin on} \end{cases} \quad \text{avec } \Upsilon \in [0,1] \text{ niveau d'acquisition d'un objectif, } P^m \in [0,20] \text{ la performance réelle mesurée et } P^T \in]s,20] \text{ la meilleure performance théorique estimée.}$$

La Figure 3-2 montre les variations de Υ . On considère ainsi que les performances de l'élève peuvent s'améliorer tant qu'il est en dessous de sa performance théorique. Au-delà, nous considérons que l'objectif est acquis.

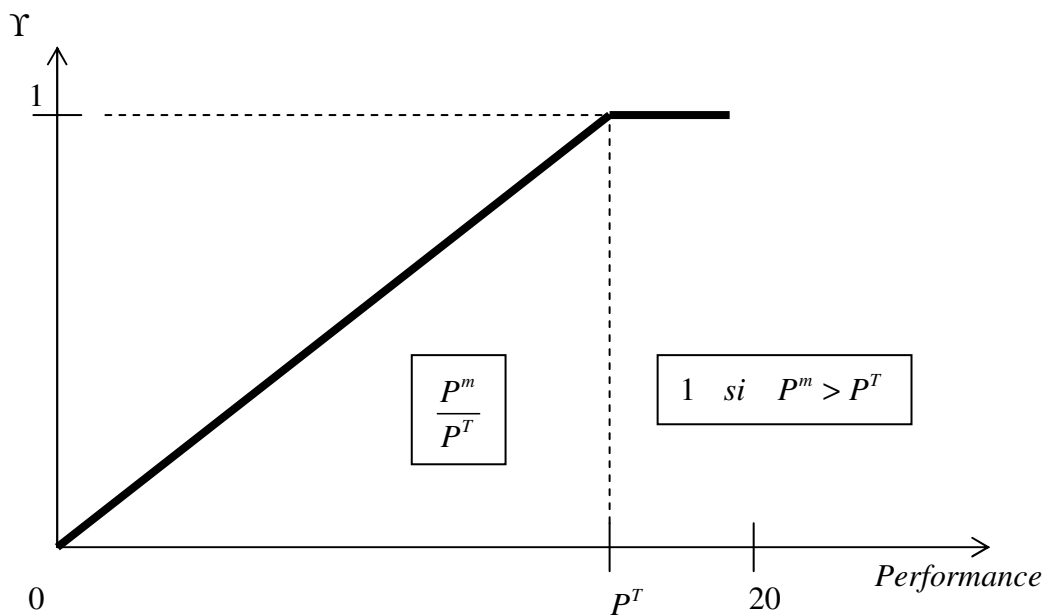


Figure 3-2 : variation du niveau d'acquisition

La Figure 3-3 montre un aperçu de l'algorithme utilisé par le séquenceur d'objectifs. Tant que $\Upsilon < 1$, le système suggère de continuer à étudier le même objectif en proposant une nouvelle séquence pédagogique adaptée. Dès que $\Upsilon \geq 1$, le séquenceur fait une inférence dans le modèle probabiliste des connaissances et suggère le passage à un nouvel objectif. Il est à noter que notre système émet des recommandations. L'élève peut dans tous les cas opter pour un choix qui semble lui convenir.

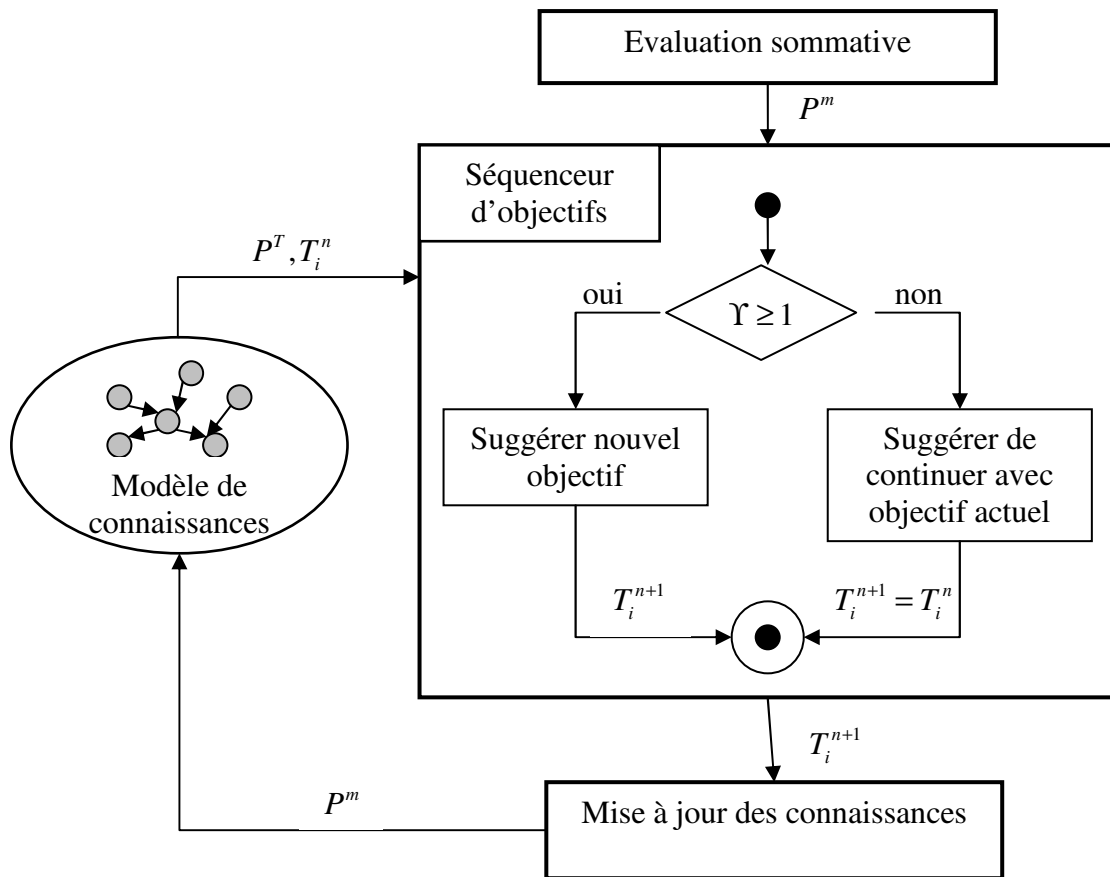


Figure 3-3 : Inférence du prochain objectif

3.4.1. Les variables

Les différentes variables qui composent le réseau bayésien du modèle de connaissances sont directement liées à la structure du modèle de contenus décrite au chapitre 4. Elles permettent d'avoir une vue globale du niveau de connaissance de l'élève pour le domaine considéré.

- Un *objectif final* correspond à l'objectif pédagogique à atteindre par l'élève. Il est considéré dans le réseau comme étant l'élément atomique de connaissance à manipuler. Nous considérons qu'il ne peut pas être décomposé en sous-objectifs plus élémentaires. Pour représenter un objectif final, nous utilisons une variable aléatoire F avec une distribution de Bernoulli. Cette variable peut prendre deux états : 1 si l'élève a atteint l'objectif, 0 sinon. La loi de probabilité de la variable F est :

$$P(F = x) = p^x (1 - p)^{1-x}$$

Avec p la probabilité que l'élève a atteint l'objectif F et x peut prendre la valeur 0 ou 1.

- Un *objectif global* correspond à une compétence métier qui doit être acquise par l'élève. Un objectif global est un couple (F, w) . F est un ensemble d'objectifs finaux qui sont mutuellement indépendants $F = \{F_1, \dots, F_n\}$. $w = \{w_1, \dots, w_n\}$ est un vecteur de poids qui mesure la relative importance de chaque objectif final F_i pour l'objectif global. On considère que $\sum_{i=1}^n w_i = 1$. Pour mesurer le niveau de connaissance d'un élève pour un objectif global, nous utilisons la variable aléatoire G telle que :

$$P(G = x) = \sum_{j=1}^n w_j P(F_j = x)$$

- Une *tâche* est une activité métier que l'élève doit être capable de réaliser. Une tâche est un couple (G, α) . G est un ensemble d'objectifs globaux qui sont mutuellement indépendants $G = \{G_1, \dots, G_s\}$. $\alpha = \{\alpha_1, \dots, \alpha_s\}$ est un vecteur de poids qui mesure la relative importance de chaque objectif global G_i pour la tâche. On considère que $\sum_{i=1}^s \alpha_i = 1$. Pour chaque $i = 1, \dots, s$ l'objectif global G_i est composé d'un ensemble d'objectifs finaux $F_i = \{F_{i1}, \dots, F_{in_i}\}$ et d'un vecteur de poids $w_i = \{w_{i1}, \dots, w_{in_i}\}$. Pour représenter le niveau de maîtrise d'un élève dans une tâche, nous utilisons une variable aléatoire T telle que :

$$P(T = x) = \sum_{i=1}^s \alpha_i P(G_i = x) \text{ avec } P(G_i = x) = \sum_{j=1}^{n_i} w_{ij} P(F_{ij} = x)$$

3.4.2. Relations causales : arcs et paramètres du réseau

Le modèle de connaissances que nous proposons, est un modèle probabiliste (Figure 3-4). Les nœuds représentent les tâches métier (T_i), les compétences métier (G_i) et les objectifs pédagogiques (F_i) associés. Les arcs entre ces nœuds expriment une relation de causalité.

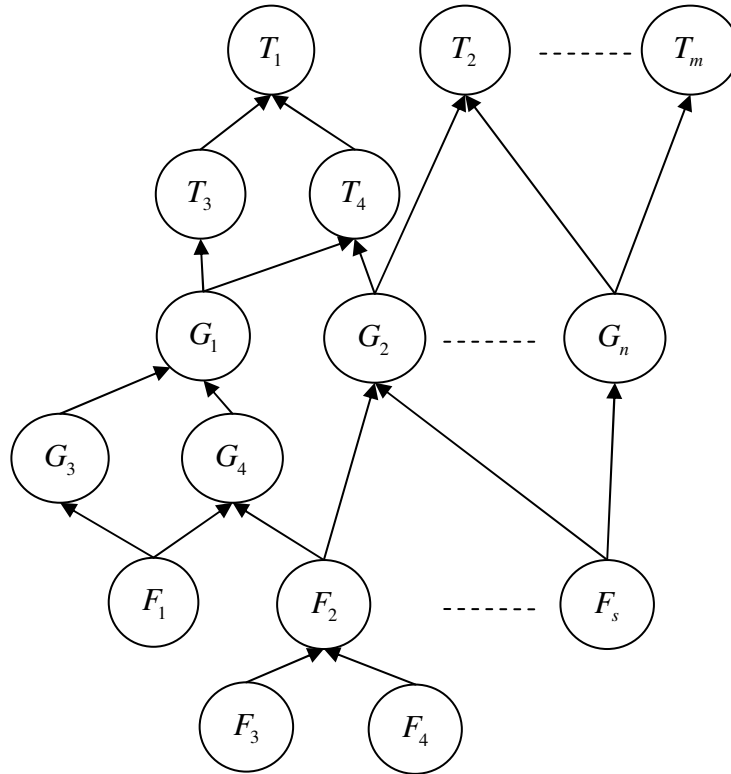


Figure 3-4 : modèle de connaissance

Un niveau de connaissance suffisant pour un objectif pédagogique F_i à une influence causale sur la maîtrise d'une compétence métier G_i . Similairement un niveau de maîtrise d'une compétence métier G_i a une influence causale sur le niveau de maîtrise de la tâche T_i . Des objectifs pédagogiques (F_i) peuvent avoir une relation causale avec un autre objectif (F_2). Cela précise que le niveau de connaissance de chaque objectif (F_i) influence le niveau de connaissance de l'objectif (F_2). Le raisonnement est le même pour les compétences et les tâches métier.

Nous considérons, par cette structure, que la maîtrise d'un élément de connaissances spécifique, a une influence causale sur la maîtrise des éléments de connaissances plus généraux.

Nous adoptons donc une approche ascendante (bottom-up). Ce choix est justifié par le fait que l'approche alternative (Figure 3-5) n'est pas réaliste pédagogiquement. En effet, celle-ci stipule que la connaissance de G'_4 influence la connaissance de F'_1 et F'_2 . Ainsi, la mise en évidence de G'_4 change les probabilités de F'_1 et F'_2 , mais bloque la communication entre eux. Nous ne pourrions pas alors opérer le choix de l'objectif suivant à étudier.

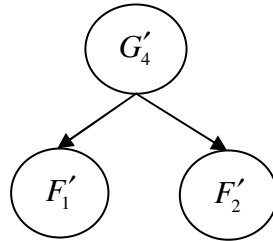


Figure 3-5 : approche top-down

Ainsi, l'évidence de la maîtrise de F_1 ou F_2 change la probabilité de maîtriser G_4 . Inversement, l'évidence de maîtriser G_4 change la probabilité de F_1 et F_2 , tout en ouvrant la communication entre les parents de G_4 . Dans ce cas, une évidence sur F_1 changera notre croyance sur F_2 .

Les paramètres à spécifier dans le réseau de la Figure 3-4 sont les probabilités a priori de maîtrise de chaque objectifs pédagogiques sans parent, et la distribution conditionnelle des nœuds associés aux compétences et tâches métier ainsi qu'aux objectifs ayant des nœuds parents.

- $P(F_i) \quad i = 1, \dots, s$
- $P(T_j | G_i) \quad \forall G_i \in pa(T_j), j = 1, \dots, m$
- $P(G_j | F_i) \quad \forall F_i \in pa(G_j), j = 1, \dots, n$
- $P(T_j | T_i) \quad \forall T_i \in pa(T_j), j = 1, \dots, m$
- $P(G_j | G_i) \quad \forall G_i \in pa(G_j), j = 1, \dots, n$
- $P(F_j | F_i) \quad \forall F_i \in pa(F_j), j = 1, \dots, s$

La structure de dépendance entre les nœuds de ce réseau se compose comme suit :

- T_j indépendant des G_i $\forall G_i \notin pa(T_j), j = 1, \dots, m$
- G_j indépendant des F_i $\forall F_i \notin pa(G_j), j = 1, \dots, n$
- Etant donné T_k et $G_i \in pa(T_k)$; G_j dépend de G_i $\forall G_j \in pa(T_k); j \neq i$
- Etant donné G_k et $F_i \in pa(G_k)$; F_j dépend de F_i $\forall F_j \in pa(G_k); j \neq i$

Dans cette thèse, nous considérons qu'à tout instant, par inférence dans le réseau de la Figure 3-4, nous sommes capables d'identifier l'objectif pédagogique F^{pert} à proposer à l'élève en fonction de son niveau actuel de connaissances.

Soit Ω , l'ensemble des F_i déjà traités par l'élève dont on connaît l'état $x \in \{0, 1\}$. L'objectif pertinent F^{pert} à proposer à l'élève est celui qui possède la probabilité de réussite la plus élevée.

$$F^{pert} = \arg \max_{F_i \notin \Omega} \{P(F_i = 1 | F_j, G_k, T_l); j \neq i\}$$

et

$$i = 1, \dots, s; j = 1, \dots, s$$

$$k = 1, \dots, n$$

$$l = 1, \dots, m$$

Les métadonnées T_i décrivant F^{pert} sont alors retournées par la fonction *obji* et la performance théorique est calculée par la fonction *perfThéo* tel que :

$$T_i = obji(F^{pert})$$

$$P^T = perfTheo(F^{pert})$$

Encore une fois, dans cette thèse nous n'étudions pas la fonction *perfTheo*. En effet, celle-ci peut être réalisée par une méthode de programmation dynamique. Cependant, à ce stade, nous optons pour une constante $perfTheo(F) = 10$. Ainsi, la performance théorique prédite pour tout élève sera toujours 10.

3.5. Modèle de préférences

De nombreux systèmes ont tenté d'utiliser des modèles de préférences de l'utilisateur. Le besoin s'est alors fait sentir d'avoir des modèles de l'utilisateur génériques pour répondre au mieux à divers cas de figure. Dans [Kobsa et al, 2001], les auteurs passent en revue le développement de systèmes de modélisation générique d'utilisateurs et présentent leurs applications dans le cadre de systèmes adaptatifs. Notre système adopte un modèle de préférences original pour recommander une séquence pédagogique adaptée. En effet, il permet de choisir les caractéristiques pédagogiques (M_i) de la séquence à présenter à l'élève étant donné son style d'apprentissage (D_i) et l'objectif (T_i) suggéré par le modèle de connaissances.

3.5.1. Conception

Le modèle proposé est une instance du méta-modèle conceptuel présenté dans la section 2.2. Notre approche est générique dans le sens qu'elle permet de manipuler diverses données observables et non observables, ainsi que divers outils mathématiques pour inférer les données non observables. Dans notre cas, les indicateurs observables et les paramètres non observables deviennent des variables aléatoires, et les relations d'influence deviennent des dépendances statistiques. On a donc d'une part un ensemble de métadonnées (indicateurs observables) décrivant l'activité pédagogique, le matériel pédagogique et le comportement de l'élève, et d'autre part le style d'apprentissage (paramètres non observables). Notre modèle est construit en établissant des dépendances statistiques entre ces métadonnées et le style d'apprentissage de l'élève. Cette solution offre un cadre d'expérimentation pour tester différentes approches pédagogiques ainsi que divers outils mathématiques. En effet, des investigations essayent d'identifier des règles pédagogiques, dérivées de la pédagogie différentielle, pour choisir le matériel pédagogique approprié pour optimiser la performance de l'élève dans la phase d'apprentissage. En étudiant les travaux de recherche liés à ce sujet, on ne trouve aucun consensus sur de telles règles. Dans ce domaine, il y a deux approches contradictoires. Certains pédagogues préconisent que les méthodes d'enseignement doivent être compatibles avec le style d'apprentissage de l'élève. D'autres suggèrent que l'utilisation de méthodes d'enseignements non adaptées au style de l'élève lui permettent de développer des stratégies personnelles pour faire face aux difficultés [Felder et al, 1988 ; Felder et al, 2001 ; Felder et al, 2002]. Pour notre part, nous optons pour la première règle.

Comme mentionné précédemment, le modèle de préférences est construit sur la base du modèle conceptuel du processus de profilage présenté dans les Figure 2-2 et Figure 2-3. Il se compose donc de deux types de variables qui sont des instances des classes *Observable* et *Non Observable*:

- **variables observables** : Elles sont des instances de la classe Observable. Elles décrivent l'activité pédagogique (T_i), le matériel pédagogique (M_i) et le comportement de l'élève (P_i).
- **variables non-observables** : Elles sont des instances de la classe Non Observable. Elles représentent les dimensions du style d'apprentissage (D_i).

Notre modèle conceptuel précise un certain nombre de dépendances entre ces variables. Nous allons nous intéresser plus particulièrement à la dépendance qui existe entre les variables (T_i, D_i) et (M_i) dans notre modèle de préférences de l'élève.

$$\{T_i, D_i\} \xrightarrow{\text{influence}} M_i$$

L'algorithme de la Figure 3-6 et le schéma bloc de la Figure 3-7 proposent une solution pour calculer cette dépendance qui consiste à la décomposer en deux fonctions : f et g .

La fonction g réalise une inférence sur le style d'apprentissage de l'élève sur la base des observations faites en cours de séquence pédagogique, alors que la fonction f réalise une inférence en utilisant l'estimation du style d'apprentissage obtenue par g tout en cherchant à maximiser le comportement de l'élève ($P_i^{opt} = \{TAL^{opt}, U^{opt}, P^{opt}\}$).

3.5.2. Apprentissage

Le modèle de dépendances statistiques dont la structure a ainsi été établie nécessite l'apprentissage de ses paramètres. En effet, pour le rendre opérationnel, nous devons apprendre les fonctions f et g . Il est donc nécessaire de choisir un formalisme mathématique pour représenter ces fonctions. Il existe deux approches principales : paramétrique ou non paramétrique. L'approche paramétrique nécessite l'établissement d'un modèle analytique pour chacune. Les paramètres de ces fonctions seraient alors appris par des méthodes de régression. Nous avons testé cette approche sur des modèles f et g linéaires en utilisant les méthodes ODR³⁷ et TLS³⁸. Ceci a donné de mauvais résultats que nous ne rapportons pas dans cette thèse.

L'approche non paramétrique s'adapte mieux à notre cas. En effet, nos variables décrivent des comportements humains. Postuler un modèle analytique dans ce domaine est hasardeux. Apprendre les paramètres d'un classificateur non paramétrique peut se faire de quatre manières différentes. Le choix de la méthode dépend de la disponibilité et du timing du feedback. Le Tableau 3-1 montre ces méthodes.

³⁷ Orthogonal Distance Regression

³⁸ Total Least Square

Méthodes	Conditions d'utilisation
Apprentissage supervisé	Des couples observation-classes sont donnés. Dans notre cas, des associations métadonnées-style.
Apprentissage partiellement supervisé	Un mélange de couples observations-classe et d'observations non classées sont donnés. Dans notre cas, nous aurions en plus des couples métadonnées-classe, des observations où nous n'avons que les métadonnées
Apprentissage non supervisé	On ne dispose que d'observations non classées
Apprentissage par renforcement	On dispose d'observations liées à une valeur d'utilité globale. On parle de fonction de valeur. Dans notre exemple, on aurait des métadonnées associées à un comportement (positif ou négatif) de l'élève.

Tableau 3-1 : méthodes d'apprentissage

Pour établir notre modèle de préférences nous avons opté pour une approche non paramétrique (de f et g). Nous avons testé deux formalismes mathématiques : les réseaux bayésiens et les machines à vecteurs de support. Dans le premier cas, les paramètres sont les probabilités conditionnelles. Dans le second, il s'agit des paramètres des hyperplans qui séparent les classes. Dans les deux cas, nous avons adopté une méthode d'estimation de paramètres en deux étapes.

Afin d'établir le modèle initial, nous avons utilisé un apprentissage supervisé. En effet, les données de notre étude empirique nous permettent de construire des couples métadonnées-style. Ainsi, on dispose d'un étiquetage des comportements et des séquences associées par les styles des élèves. Cependant, ces données ne sont pas représentatives de la population entière. Elles sont incomplètes et imprécises. Pour cela, nous avons conçu une deuxième étape d'apprentissage par renforcement. Les nouvelles observations, pour lesquelles on dispose d'indices fiables quant au comportement, sont utilisées en continu pour affiner les paramètres (Figure 3-7). En pratique, seules celles qui vérifient le critère $W \geq W_{theorique}$ serviront dans ce processus de renforcement.

Dans le restant de ce chapitre, nous décrivons les deux solutions proposées et présentons l'apprentissage de leur paramètres. Nous réalisons une comparaison de leurs performances.

3.5.3. Inférence

L'algorithme permettant d'exploiter le modèle de préférences est présenté en Figure 3-6 et Figure 3-7. Cet algorithme est le fondement du composant estimateur de style de notre système.

Soit

D_i^0 le style initial le plus probable

T_i^0 l'activité d'apprentissage à réaliser

Calcul des caractéristiques les plus probables du matériel pédagogique

$$M_i^0 = f(D_i^0, T_i^0, TAL^{opt}, U^{opt}, P^{opt})$$

Composer le nouveau matériel pédagogique et le soumettre à l'élève

Observer le comportement de l'élève et calculer les indicateurs de comportement

$$P^{m_0}, TAL^0, U^0$$

Soit e le numéro de l'itération courante (initialement 0)

tantque $P_i^{m_e}$ n'est pas satisfaisant

Réaliser une nouvelle estimation du style d'apprentissage

$$D_i^{e+1} = g(TAL^e, U^e, T_i^0, M_i^e, P^{m_e})$$

Réaliser une inférence directe pour calculer les caractéristiques les plus probables du matériel pédagogique

$$M_i^{e+1} = f(D_i^{e+1}, T_i^0, TAL^{opt}, U^{opt}, P^{opt})$$

Composer le nouveau matériel pédagogique et le soumettre à l'élève

Observer le comportement de l'élève et calculer les indicateurs de comportement

$$P^{m_{e+1}}, TAL^{e+1}, U^{e+1}$$

fantantque

Figure 3-6 : inférence des métadonnées M_i en fonction de $\{T_i, D_i\}$

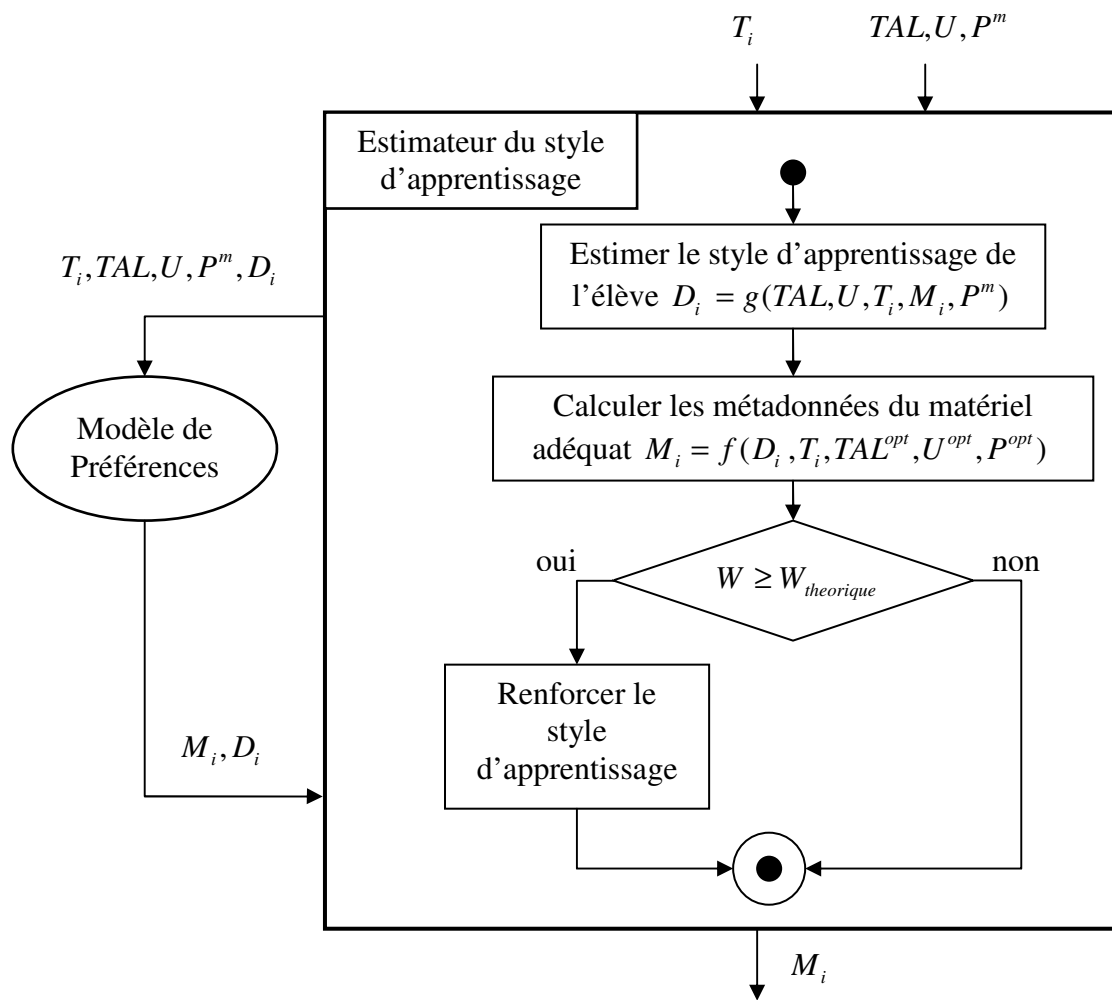


Figure 3-7 : Inférence du style de l'élève

3.6. Un réseau bayésien pour inférer les préférences d'un élève

Dans ce paragraphe, nous concevons une instance de notre modèle de préférences en adoptant le formalisme de réseau bayésien. Nous présentons dans un premier temps, un aperçu théorique sur les réseaux bayésiens. Puis en second lieu, nous décrivons les variables et la structure du réseau. Nous présentons ensuite les résultats de validation de ce modèle.

3.6.1. Aperçu théorique

Un réseau bayésien [Pearl, 1988] est un graphe acyclique orienté dans lequel les nœuds représentent des variables et les arcs représentent des dépendances probabilistes entre les variables. Les paramètres utilisés pour représenter l'incertain sont les probabilités conditionnelles de chaque nœud en fonction des différents états de ses nœuds parents. Si $\{X_i, i = 1, \dots, n\}$ sont les variables du réseau et $pa(X_i)$ représente l'ensemble des parents de X_i , pour chaque $i = 1, \dots, n$, alors les paramètres du réseau sont $\{P(X_i | pa(X_i)); i = 1, \dots, n\}$. Ces paramètres forment l'ensemble des distributions de probabilités conditionnelles discrètes de chaque variable en fonction de ses parents. Cet ensemble de probabilités définit la distribution de probabilité jointe pour l'ensemble du réseau tel que :

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | pa(X_i))$$

Ainsi pour définir un réseau bayésien, nous devons spécifier :

- L'ensemble des variables X_1, X_2, \dots, X_n
- L'ensemble des arcs entre ces variables. Ces arcs représentent les influences causales entre les variables. Le réseau formé de ces variables et de ces arcs doit être un graphe acyclique orienté.
- Pour chaque variable X_i , les tables de probabilités conditionnelles en fonction de ses parents $P(X_i | pa(X_i)); i = 1, \dots, n$.

Dans le cadre de la modélisation d'un élève, les variables peuvent représenter différentes choses en fonction du domaine : des concepts, des compétences, des objectifs pédagogiques, des activités... Les arcs entre ces variables peuvent définir des relations de différents types : est composé de, est un pré-requis de, influence... Une fois les arcs et les variables définis, il faut spécifier les probabilités conditionnelles pour chaque variable possédant au moins un parent.

3.6.2. Structure du réseau

Les différentes variables qui composent le réseau bayésien du modèle de préférences sont directement définies par le modèle conceptuel du profilage du style d'apprentissage d'un élève décrit au paragraphe 2.6. On y retrouve les variables non observables et les variables observables du modèle conceptuel.

Variables non observables : Le style d'apprentissage de l'élève

Le style d'apprentissage d'un élève se compose d'un ensemble de quatre variables aléatoires discrètes $\mathcal{D} = \{D_1, D_2, D_3, D_4\}$. Chaque variable représente la projection de l'élève sur l'un des quatre axes du modèle de Felder (Visuel-Verbal, Global-Sequentiel, Réfléchi-Actif, Déductif-Inductif). Chacune des variables $\{D_i\}_{i \in [1,4]}$ prend une valeur entière comprise entre -3 et 3. $D_i \in [-3, -1, 1, 3]; \forall i \in \{1, 2, 3, 4\}$.

Variables observables : Le matériel pédagogique

Le matériel pédagogique décrivant une séquence pédagogique se compose d'un ensemble de sept variables aléatoires discrètes $\mathcal{M} = \{M_1, M_2, M_3, M_4, M_5, M_6, M_7\}$.

- La variable M_1 représente les médias principalement utilisés dans le matériel pédagogique. Elle peut prendre trois valeurs $M_1 \in \{T, I, V\}$: (T) principalement texte probablement illustré avec des images, (I) principalement graphiques et images accompagnés avec des explications textuelles, (V) principalement voix avec des graphiques d'illustration.
- La variable M_2 représente la structure externe du contenu. Elle peut prendre deux valeurs $M_2 \in \{C, F\}$. (F) pour la navigation libre ou (C) pour la navigation contrainte.
- La variable M_3 est une variable booléenne $M_3 \in \{vrai, faux\}$. Elle montre si le matériel contient des résumés et des aides de navigation.
- La variable M_4 est également une variable booléenne $M_4 \in \{vrai, faux\}$. Elle est vraie quand le matériel contient des définitions et de la théorie. Elle est fausse quand le matériel consiste seulement en travaux pratiques, exemples, exercices, etc.
- La variable M_5 est une variable booléenne $M_5 \in \{vrai, faux\}$. Elle montre si le matériel contient des exemples.
- La variable M_6 représente le type d'exercices pratiques proposé dans le matériel. Elle peut prendre 6 valeurs $M_6 \in \{TF, QCM1, QCMN, DD, FS, N\}$. (TF) questions vrai ou faux, ($QCM1$) questions à choix multiples avec seulement une réponse possible, ($QCMN$) questions à choix multiples avec plusieurs réponses possibles, (DD) déplacer et déposer, (FS) texte à trous, (N) pas d'exercice.
- La variable M_7 est semblable à la variable M_6 mais ajoute la notion d'évaluation. Elle décrit le type de questions utilisées pour examiner le niveau de connaissances de l'élève. Elle peut prendre 6 valeurs $M_7 \in \{TF, QCM1, QCMN, DD, FS, N\}$

Variables observables : L'activité pédagogique

L'activité pédagogique est décrite par un ensemble de trois variables aléatoires discrètes $\mathcal{T} = \{T_1, T_2, T_3\}$.

- La variable T_1 représente l'objectif d'apprentissage. En se basant sur la taxinomie proposée par Bloom [Bloom, 1956], cette variable peut prendre 6 valeurs $T_1 \in \{1, 2, 3, 4, 5, 6\}$. Cette variable décrit qu'une activité pédagogique peut viser un des six niveaux d'objectifs d'apprentissage : la connaissance (1), la compréhension (2), l'application (3), l'analyse (4), la synthèse (5), et l'évaluation (6).
- La variable T_2 indique le type de l'activité. Elle peut prendre 3 valeurs $T_2 \in \{1, 2, 3\}$: expositive (1), active (2) et interrogative (3).
- La variable T_3 indique la nature de la connaissance à acquérir dans l'activité. Elle peut prendre 5 valeurs $T_3 \in \{1, 2, 3, 4, 5\}$: concept (1), fait (2), processus (3), principe (4), et procédé (5).

Variables observables : Le comportement de l'élève

Nous considérons que le comportement d'un élève est représenté par l'ensemble des trois variables aléatoires continues $\mathcal{P} = \{P^m, TAL, U\}$. Rappelons que :

- La variable P^m mesure la performance réelle obtenue par l'élève à une évaluation sommative en fin de séquence. Elle est continue sur l'intervalle $P^m \in [0, 20]$.
- La variable TAL mesure le niveau d'adaptation de la tâche au style d'apprentissage de l'élève. Elle est continue sur l'intervalle $TAL \in [0, 1]$.
- La variable U mesure le degré d'attention d'un élève pendant la séquence pédagogique. Elle est continue sur l'intervalle $U \in [0, \infty[$.

Toutes les variables aléatoires booléennes du modèle suivent une distribution de Bernoulli. Ces variables peuvent prendre deux états. Par exemple la variable M_3 prendra la valeur 1 si la séquence pédagogique contient un résumé, 0 sinon. La loi de probabilité de la variable M_3 est :

$$P(M_3 = x) = p^x(1-p)^{1-x}$$

avec p la probabilité que la séquence contienne un résumé et x peut prendre la valeur 0 ou 1.

Relations causales

Pour définir complètement la structure du réseau bayésien, nous devons établir les dépendances (ou les arcs) entre les nœuds. Il y a deux types de dépendances :

- dépendances entre les paramètres non observables et les indicateurs observables (**NObs-Obs**) : Elles permettent d'exprimer que le style d'apprentissage influence le choix du matériel pédagogique et le comportement de l'élève (Figure 3-8).

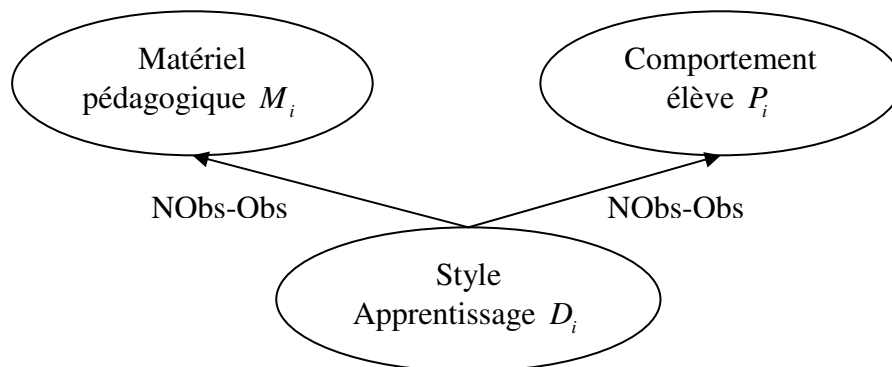


Figure 3-8 : dépendances entre M_i, P_i et D_i

Ainsi, l'observation du comportement de l'élève P_i change les probabilités sur le style d'apprentissage D_i ; qui à leur tour changent les probabilités sur le matériel pédagogique M_i . Dans le cas où nous mettons en évidence une croyance sur le style d'apprentissage, les probabilités sur le matériel pédagogique et le comportement de l'élève changent.

Les paramètres à définir pour cette structure sont les probabilités a priori de chaque $D_i \{P(D_i), i = 1, 2, 3, 4\}$ et les distributions des probabilités conditionnelles de chaque $M_j \{P(M_j | D_i); j = 1, \dots, 7; i = 1, 2, 3, 4\}$ et $P_j \{P(P_j | D_i); j = 1, 2, 3; i = 1, 2, 3, 4\}$.

- dépendances entre indicateurs observables (**Obs-Obs**) : Elles permettent d'exprimer que l'activité pédagogique influence le choix du matériel pédagogique. Elles montrent aussi que l'association d'une activité pédagogique à un matériel pédagogique influence le comportement de l'élève (Figure 3-9).

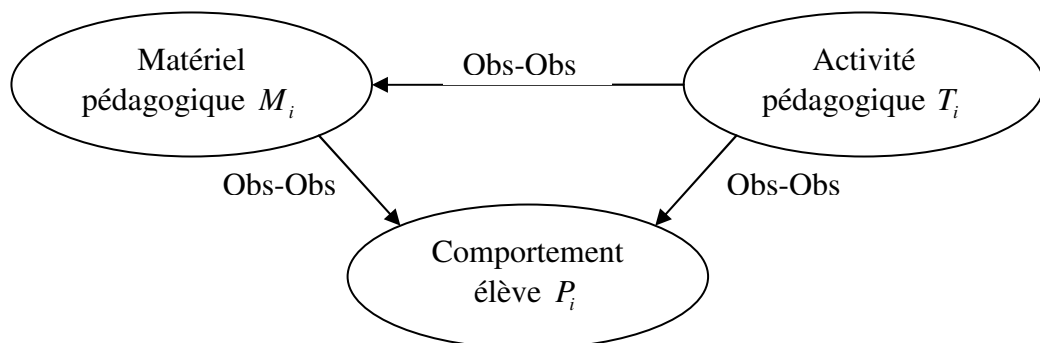


Figure 3-9 : dépendances entre M_i, P_i et T_i

Ainsi, l'observation du comportement de l'élève P_i change les probabilités de l'activité pédagogique T_i ; Celles-ci changent à leur tour les probabilités sur le matériel pédagogique M_i . La mise en évidence de l'activité pédagogique T_i change les probabilités sur le matériel pédagogique M_i et le comportement de l'élève P_i . L'observation d'un certain comportement de l'élève P_i change les probabilités sur le matériel pédagogique M_i , l'activité pédagogique T_i , mais aussi sur le style d'apprentissage D_i . Dans ce cas, la *communication* entre D_i, M_i et T_i est ouverte. Cela signifie que toute nouvelle croyance sur l'un des trois ensembles de variables change les probabilités sur les deux autres. Il est alors rendu possible, en fixant l'activité pédagogique T_i et le comportement de l'élève P_i , d'inférer le style d'apprentissage D_i et le matériel pédagogique M_i .

Les paramètres à définir pour cette structure sont les probabilités a priori de chaque $T_i \{P(T_i), i = 1, 2, 3\}$ et les distributions des probabilités conditionnelles de chaque $M_j \{P(M_j | T_i); j = 1, \dots, 7; i = 1, 2, 3\}$ et $P_j \{P(P_j | T_i, M_k); j = 1, 2, 3; i = 1, 2, 3; k = 1, \dots, 7\}$.

On obtient donc la vue globale de notre réseau bayésien (Figure 3-10).

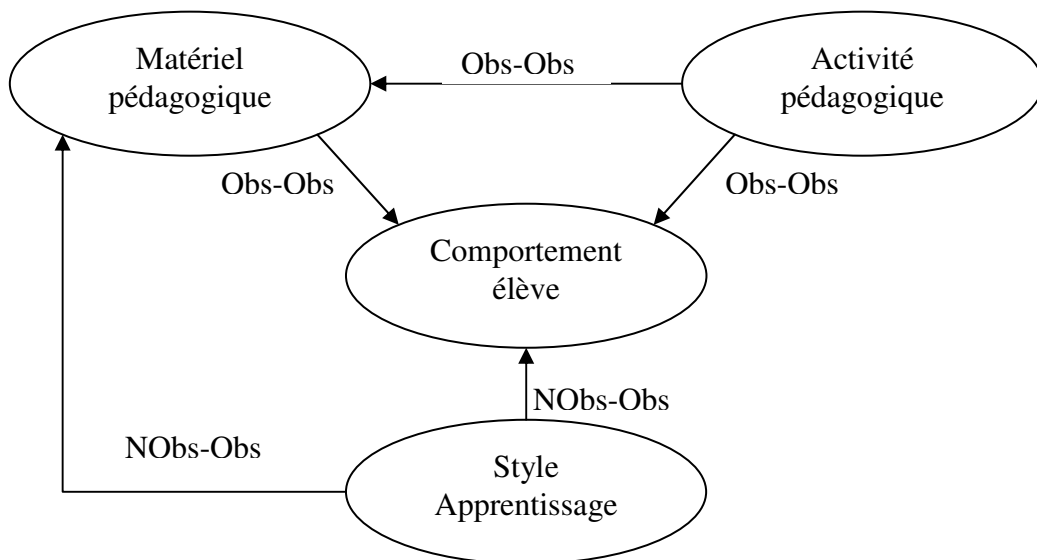


Figure 3-10 : schéma réseau bayésien

La structure détaillée du réseau bayésien est composée de quatre couches (Figure 3-11). Les trois premières couches comportent les variables associées aux données observables. La première contient les nœuds T1, T2 et T3 qui décrivent l'activité pédagogique. La seconde contient les nœuds M1 à M7 qui caractérisent le matériel pédagogique. La troisième couche contient les nœuds P^m , TAL et U qui mesurent le comportement de l'élève.

La dernière couche contient les variables associées aux données non observables qui sont dans notre cas celles qui expriment le style d'apprentissage de l'élève (section 2.6.1).

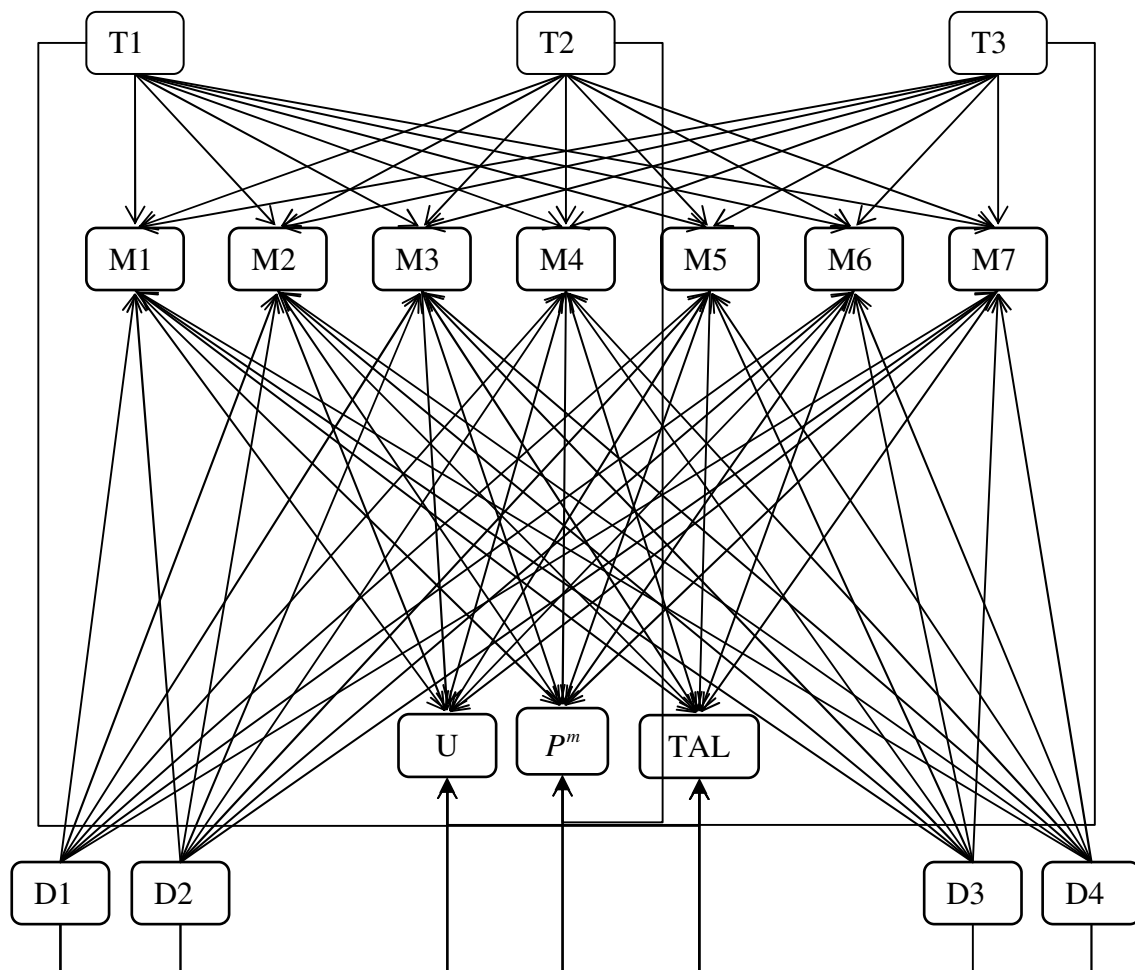


Figure 3-11 : structure du réseau bayésien

3.6.3. Entraînement du réseau

Afin de compléter la conception du réseau bayésien, nous devons établir les probabilités conditionnelles à chaque nœud. Nous appelons ce processus l'entraînement du réseau.

Deux méthodes sont habituellement appliquées à cette fin :

- Expertise,
- Observation statistique.

La première approche se fonde sur la connaissance d'un expert ou des résultats de recherches antérieures pour remplir les tables de probabilités conditionnelles. La seconde approche, consiste à entreprendre des études empiriques à partir desquelles nous extrayons des dépendances statistiques.

Vu la complexité du réseau de la Figure 3-11, un entraînement par observation statistique est irréalisable. En effet, cela nécessiterait une très large étude pour collecter les données. D'un autre côté, la méthode par expertise n'est pas fiable. Tenant à utiliser des données terrain, nous avons opté pour la méthode d'observation partielle. Ainsi, nous avons simplifié le réseau en ignorant les nœuds $T_2, T_3, M_3, M_4, M_5, M_6, M_7$. Nous avons alors, utilisé les résultats de notre étude empirique pour l'entraînement.

L'étude empirique

L'étude menée au Collège Malraux nous a permis de recueillir des données afin de calculer les probabilités conditionnelles du réseau bayésien. Cependant, ces données présentent deux limitations majeures.

La première concerne la validité de l'évaluation. Puisque toutes les évaluations ont été réalisées en ligne, il se peut que quelques élèves passent rapidement la séquence pédagogique et répondent aux questions d'évaluation aléatoirement. Les réponses doivent donc être considérées comme des conjectures aléatoires et ne doivent pas être apprises par le réseau. Les données correspondantes sont considérées comme bruitées et seront éliminées. Nous avons ainsi mis en œuvre une méthode de filtrage fondée sur notre mesure originale : l'énergie d'apprentissage W . Ainsi, tout résultat obtenu après une séquence où $W < W_{theorique}$ est éliminé.

La deuxième limitation est liée à la portée des données. Comme expliqué ci-dessus, le style d'apprentissage est mesuré en utilisant notre questionnaire RILS. Cependant, les styles ne sont pas distribués équitablement sur les quatre dimensions. La population cible pourrait avoir principalement des élèves visuels et très peu d'élèves verbaux. Ceci rend les données trop faibles pour former définitivement le réseau bayésien. Afin de remédier à ce problème, nous adoptons une méthode de renforcement continu du réseau bayésien. Les nouveaux cas observés, filtrés selon la méthode précédente, seront utilisés pour renforcer les paramètres du réseau.

Les données manquantes

Le traitement de filtrage sur les données obtenues par l'étude empirique ainsi que le manque d'exhaustivité de la représentation des styles d'apprentissage dans l'échantillon de la population font que nous obtenons une matrice similaire à celle présentée dans le Tableau 3-2. Rappelons que nous travaillons ici sur un ensemble réduit de variables. C'est ce qui explique que n'apparaissent que T_1 , M_1 et M_2 (cf. chapitre 2).

T_1	M_1	M_2	D_1	D_2	D_3	D_4	TAL	P	U
1	T	C	3	-1	-1	-3	0.70	16	1.2
1	I	C	3	-3	1	3	0.20	8	2.5
1	I	C	3	-1	-1	3	0.22	?	0.1
2	I	F	-3	1	3	1	?	?	?

Tableau 3-2 : matrice de données après filtrage

Tous les élèves ne répondant pas au critère $W \geq W_{theorique}$ ont eu leur performance ignorée. Cela génère des lignes sur lesquelles seule la performance P n'est pas connue. Pour tous les styles qui n'ont pas été rencontrés, nous obtenons des lignes avec des valeurs inconnues pour les colonnes TAL, P et U (Tableau 3-2). Cet état de fait nous situe dans le contexte de l'apprentissage de distributions de probabilités conditionnelles à partir d'un ensemble de données incomplet. On trouve dans la littérature de nombreux travaux sur ce domaine [Hartley, 1958 ; Dempster et al., 1977 ; Tanner, 1996 ; Neal et Hinton, 1998 ; Minka, 1998 ; Ramoni et Sebastiani, 2001 ; Leray et François, 2005].

Nous rappelons que notre étude présentée au chapitre 2, a permis de faire apparaître le nombre de combinaisons supprimées par groupe (Tableau 3-3).

	Exp (33 élèves)	Ctrl1 (33 élèves)	Ctrl2 (33 élèves)	Ctrl3 (31 élèves)
Anomalies fichiers de logs	0%	0%	8%	20%
$W \geq W_{theorique}$	2%	10.7%	8%	60%
Total	2%	10.7%	16%	80%

Tableau 3-3 : Pourcentage de résultats éliminés par groupe

Nous avons adopté l'algorithme EM [McLachlan et Krishnan, 1997] pour résoudre ce problème. Cet algorithme est une méthode d'optimisation itérative pour estimer des paramètres inconnus Θ , en fonction de données mesurées K et de variables cachées J contenues dans un espace S . Le principe consiste à maximiser les probabilités a posteriori des paramètres Θ sachant K tout en marginalisant J ,

$$\Theta^* = \arg \max_{\Theta} \sum_{J \in S} P(\Theta, J | K)$$

L'ensemble de données ainsi obtenues permet de calculer les paramètres pour chaque nœud du réseau bayésien. Le réseau initial résultant est utilisé pour inférer les préférences de l'élève.

3.6.4. Inférences et renforcement du réseau

L'inférence des préférences de l'élève se décompose en deux parties : estimation du style d'apprentissage D_i , puis ensuite estimation de la séquence pédagogique M_i .

Des systèmes existants basés sur le style d'apprentissage sont habituellement employés pour adapter l'enseignement au profil de l'élève. De tels systèmes commencent par évaluer le style d'apprentissage de l'élève en utilisant un questionnaire par individu [Papanikolaou et al, 2003].

Le résultat est alors employé pour calculer le meilleur matériel pédagogique. Dans notre cas, le style d'apprentissage initial est le plus probable dans la population des élèves. Etant donné une activité pédagogique, les variables $\{T_i\}_{i \in [1, k]}$ et les variables du style d'apprentissage $\{D_i\}_{i \in [1, q]}$ sont initialisées aux valeurs correspondantes.

Plus précisément, $T_i, i \in [1, k]$, sont initialisées avec les valeurs $\{t_{target}^i\}_{i \in [1, k]}$ décrivant l'activité cible.

$$P(T_i = t_j^i) = \begin{cases} 1 & \text{si } t_j^i = t_{target}^i \\ 0 & \text{autrement} \end{cases} \quad \forall j \in [1, \text{card}(T_i)] .$$

$\text{card}(T_i)$ est le nombre de modalités de la variable T_i

Les variables du style d'apprentissage $D_i, i \in [1, q]$ sont placées aux valeurs initiales pour la première fois où l'élève utilise le système. Soit $d_0^i, \forall i \in [1, q]$ les valeurs correspondant aux probabilités les plus hautes du réseau initial,

$$d_0^i = \underset{d_j^i; j \in [1, \text{card}(D_i)]}{\text{argmax}} \left(P(D_i = d_j^i) \right) \forall i \in [1, q] .$$

Nous plaçons les probabilités des dimensions du style d'apprentissage de la façon suivante :

$$P(D_i = d_j^i) = \begin{cases} 1 & \text{si } d_j^i = d_0^i \\ 0 & \text{autrement} \end{cases} \quad \forall j \in [1, \text{card}(D_i)].$$

$\text{card}(D_i)$ est le nombre de modalités de la variable D_i

Pour les utilisations suivantes du système, $D_i, i \in [1, q]$ sont ajustées au style d'apprentissage estimé de l'élève dans l'itération précédente. Si $d_e^i, \forall i \in [1, q]$ sont les valeurs estimées des dimensions du style après l'itération e , nous plaçons l'évidence des variables du style d'apprentissage comme suit :

$$P(D_i = d_j^i) = \begin{cases} 1 & \text{si } d_j^i = d_e^i \\ 0 & \text{autrement} \end{cases} \quad \forall j \in [1, \text{card}(D_i)]$$

Une inférence probabiliste directe est ensuite exécutée dans le réseau pour calculer les probabilités conditionnelles a posteriori des variables du matériel pédagogique $\{M_i\}_{i \in [1, m]}$. L'algorithme utilisé est « Message Passing » appliqué aux arbres de jonction des cliques maximales [Spiegelhalter et al., 1993 ; Jensen, 1996 ; Neapolitan, 1990]. Les valeurs $\{m_e^i\}_{i \in [1, m]}$ correspondant aux probabilités les plus élevées sont sélectionnées,

$$m_e^i = \underset{m_j^i \in [1, \text{card}(M_i)]}{\text{argmax}} \left(P(M_i = m_j^i / \{T_i = t_{\text{target}}^i\}_{i \in [1, k]}, \{D_i = d_e^i\}_{i \in [1, q]}) \right).$$

Ces métadonnées choisies sont employées pour composer une activité pédagogique qui est proposée à l'élève. Surveiller le comportement de l'élève pendant la séquence pédagogique mène à mesurer les variables $\{P_i\}_{i \in [1, m]}$. Soit $\{p_e^i\}_{i \in [1, m]}$ les valeurs mesurées à l'itération e . Si le critère $W \geq W_{\text{theorique}}$ est respecté, ces valeurs sont placées comme évidence dans le réseau et une inférence en arrière est exécutée pour calculer les nouvelles valeurs des dimensions du style d'apprentissage $\{D_i\}_{i \in [1, q]}$. Soit $\{d_{e+1}^i\}_{i \in [1, q]}$ les valeurs estimées :

$$d_{e+1}^i = \underset{d_j^i \in [1, \text{card}(D_i)]}{\text{argmax}} \left(P(D_i = d_j^i / \{T_i = t_{\text{target}}^i\}_{i \in [1, k]}, \{P_i = p_e^i\}_{i \in [1, m]}) \right)$$

Dans le cas où le critère $W \geq W_{\text{theorique}}$ n'est pas respecté, aucune inférence n'est faite sur le style d'apprentissage. On garde l'ancienne estimation.

Ces valeurs estimées constitueront le nouveau style de l'élève. Après un certain nombre d'itérations, ce processus de raffinement converge avec précision à une évaluation du style caché d'apprentissage de l'élève.

De nouvelles fréquences sont calculées et les probabilités conditionnelles sont mises à jour pour prendre en compte ces nouvelles données. Le renforcement du réseau est ainsi obtenu par la découverte ou la mesure de nouveau cas.

3.6.5. Validation du modèle

Afin de valider notre modèle prédictif, nous adoptons une approche expérimentale basée sur les données recueillies lors de notre étude empirique. 80% des mesures ont été utilisées pour entraîner le réseau bayésien (Figure 3-12) et 20% pour le valider.

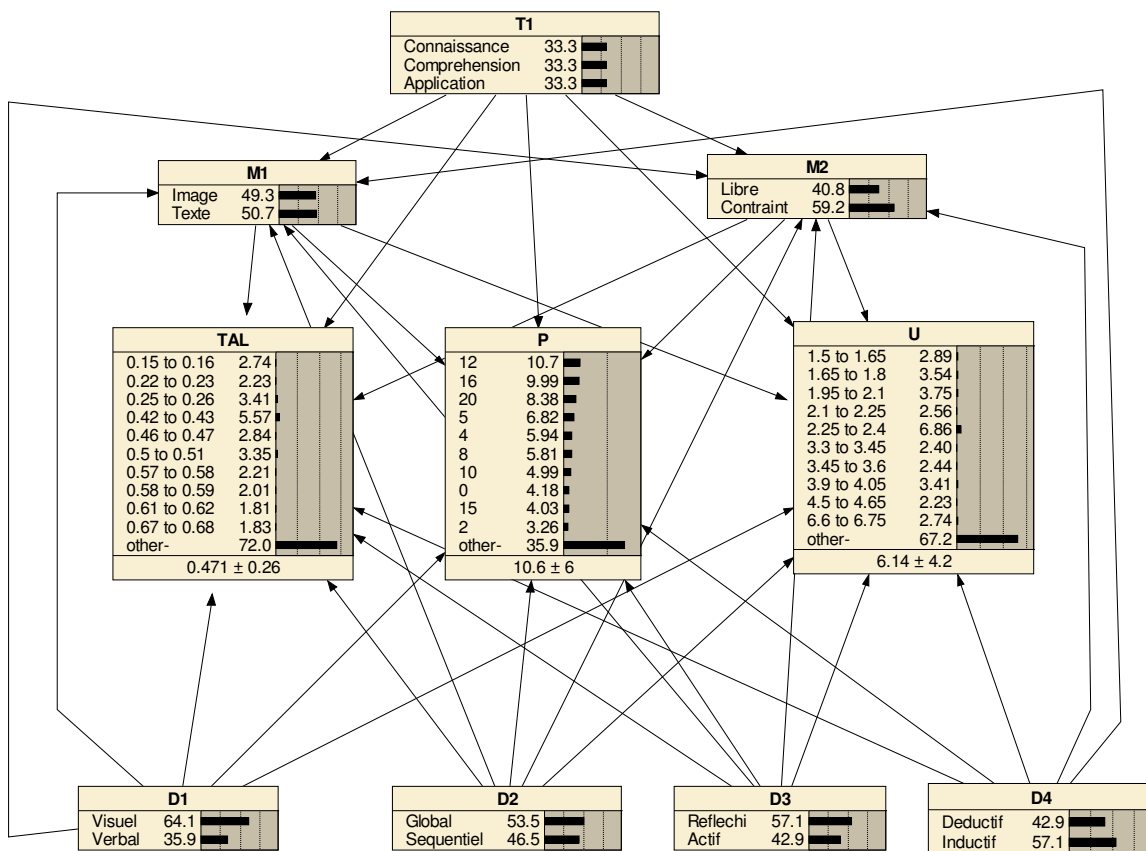


Figure 3-12 : réseau bayésien

La structure du réseau est constituée des variables manipulées dans notre étude empirique (cf. chapitre 2). C'est pourquoi nous n'avons à nouveau qu'un nombre limité de nœuds dans le réseau. Le réseau se compose de trois couches. La couche des données observables est composée des nœuds T1, M1 et M2 alors que la couche des données à estimer est composée des nœuds D1, D2, D3 et D4. Les nœuds TAL, P^m et U représentent le comportement que nous allons analyser afin d'émettre une croyance sur la valeur des nœuds à estimer en fonction de celle des données observées.

La première étape consiste à valider l'estimation du style. La Figure 3-13 montre le processus que nous avons suivi pour valider le réseau bayésien (RB) à partir de 20% des données mesurées lors de l'expérimentation en classe de cinquième.

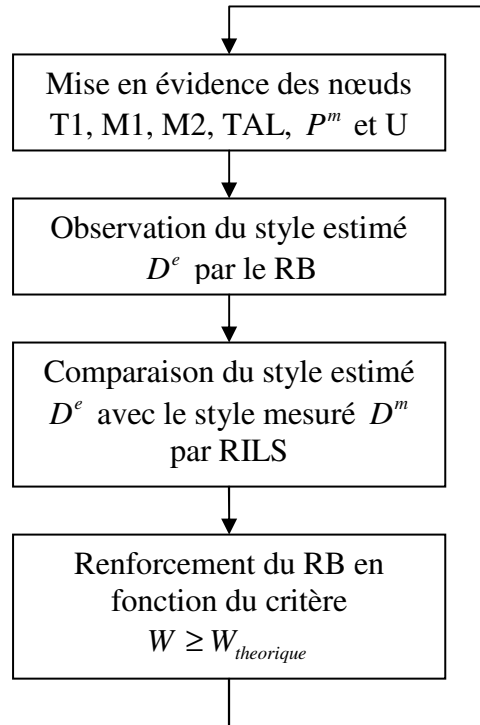


Figure 3-13 : processus de validation du réseau bayésien (RB)

Après avoir entraîné le RB avec 80% des données, nous cherchons à mesurer le nombre de fois que le RB estimera correctement le style de l'élève dans les 20% de données restantes. A chaque nouvelle recherche nous renforçons les croyances du RB en fonction du critère $W \geq W_{\text{théorique}}$. Ainsi, les nouveaux cas découverts seront ajoutés au fur et à mesure du processus de validation. La Figure 3-14 présente les résultats obtenus.

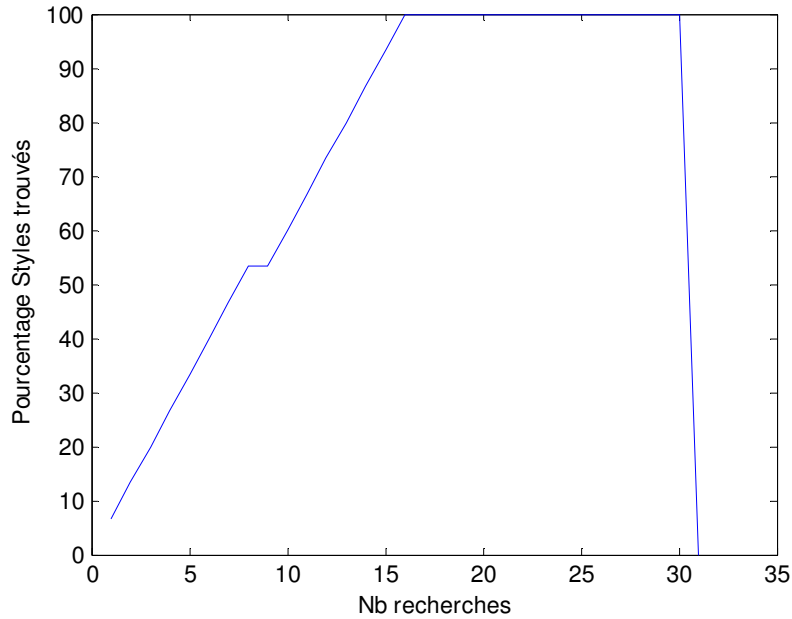


Figure 3-14 : pourcentage de style correctement estimé par le RB

On observe qu'à la première passe, seulement 7% des styles estimés correspondent à ceux mesurés avec le questionnaire ILS adapté. Ensuite, grâce au processus de renforcement du réseau, on remarque que le nombre de bonnes réponses augmente régulièrement jusqu'à obtenir 100% de résultats corrects. Ceci s'explique par le fait que le RB est capable de retrouver sans erreur des cas déjà rencontrés, mais commet de nombreuses erreurs sur l'estimation de cas jamais observés. Ainsi, lors de son utilisation, il ne pourra estimer correctement que les cas déjà rencontrés. Ceci limite sa capacité d'apprentissage par renforcement.

La seconde étape consiste à valider l'estimation des métadonnées. La Figure 3-15 montre le processus que nous avons suivi pour valider le réseau bayésien (RB) à partir de 20% des données mesurées lors de l'expérimentation en classe de cinquième.

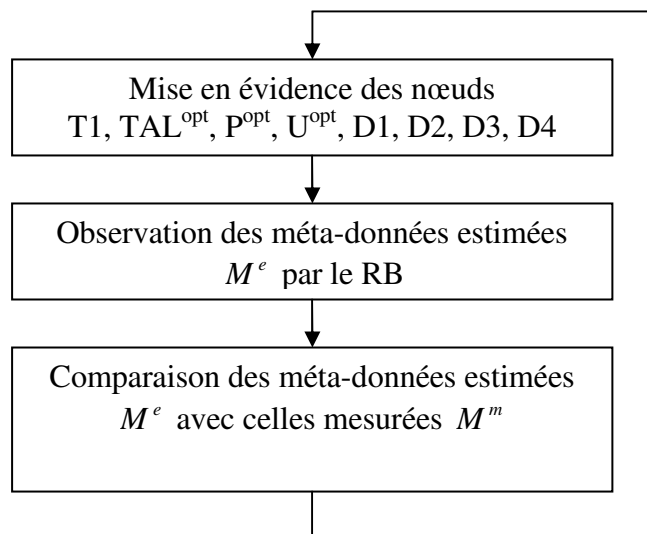


Figure 3-15 : processus de validation du réseau bayésien (RB)

Après avoir entraîné le RB avec 80% des données, nous cherchons à mesurer le nombre de fois que le RB estimera correctement les métadonnées dans les 20% de données restantes. La Figure 3-16 présente les résultats obtenus.

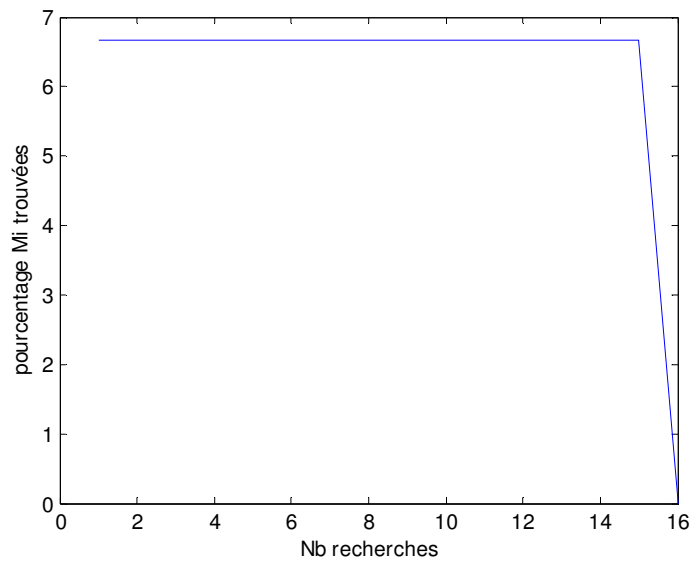


Figure 3-16 : pourcentage de M_i correctement estimées par le RB

On observe que seulement 7% des métadonnées estimées correspondent à celles réellement utilisées lors des mesures. Ceci a la même explication que pour le style. En effet, nous cherchons à maximiser les variables TAL, U et P. Les valeurs de ces trois variables, correspondant au comportement optimal, n'ont peut être jamais été rencontrées. Dans cette phase, on ne fait pas de renforcement du réseau avec les cas traités. C'est ce qui explique que la performance du réseau reste inchangée.

3.7. Machine à Vecteurs de Support pour inférer les préférences d'un élève

Dans cette section, nous montrons l'utilisation des machines à vecteurs de support (SVM) [Boser et al., 1992], appelée aussi séparateur à vaste marge [Cornuéjols et al., 2002], comme méthode à noyaux pour faire de la classification par apprentissage supervisé [Lauer et Bloch, 2006 ; Scholkopf et Smola, 2002 ; Shawe-Taylor et Cristianini, 2000]. Le principe de SVM est de trouver un classificateur, appelé hyperplan, qui sépare un ensemble de points en plusieurs classes. L'hyperplan tente de maximiser la distance entre les classes. L'apprentissage supervisé par induction permet d'arriver à des conclusions par l'examen d'exemples particuliers représentés par un ensemble de couples de données d'entrée/sortie. Le but est d'apprendre une fonction qui correspond aux exemples vus et qui prédit les sorties pour des entrées qui n'ont pas encore été vues. Les entrées peuvent être des descriptions d'objets et les sorties les classes des objets données en entrée.

Dans un premier temps, nous présentons un aperçu théorique des SVM en abordant les classificateurs bi-classes et multi-classes. Ensuite, nous posons notre problème d'inférence des préférences d'un élève comme étant celui de concevoir un estimateur à deux étages mettant en œuvre deux classificateurs multi-classes. Nous finissons par présenter les résultats obtenus avec les données recueillies lors d'une expérimentation empirique.

3.7.1. Aperçu théorique

On se place dans le contexte des systèmes discriminants et de la théorie statistique de l'apprentissage [Guermeur, 2004]. Les modèles de discrimination abordent principalement la classification d'éléments appartenant à un espace de description X suivant deux problématiques distinctes : la discrimination bi-classes ou multi-classes. Ces deux types de discrimination sont couverts par les SVM [Guermeur et al., 2005 ; Guermeur et al., 2000].

Discrimination bi-classes

Le principe théorique des SVM comporte deux étapes fondamentales :

- Premièrement, la recherche d'une transformation non linéaire (Φ) des exemples de l'espace d'entrée (X) vers un espace (F) dit de redescription. F est généralement de grande dimension, et muni d'un produit scalaire (espace de Hilbert) (Figure 3-17),

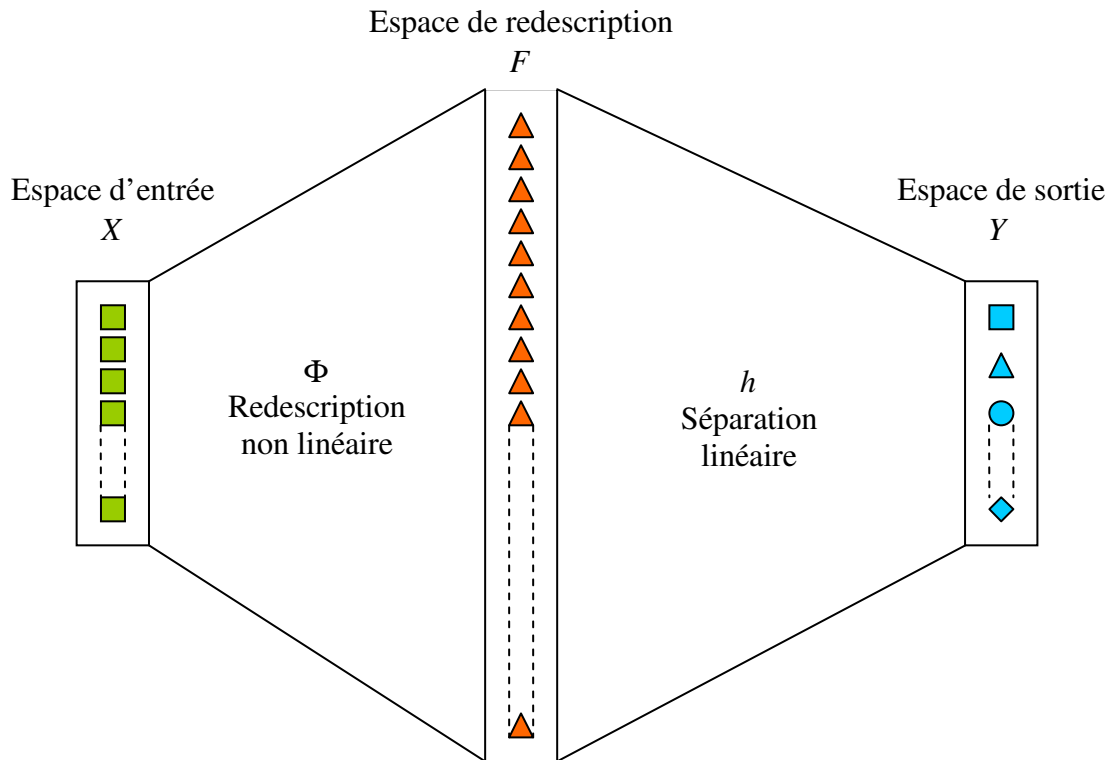


Figure 3-17 : la transformation non linéaire des données offre une séparation linéaire des exemples dans un nouvel espace

Concrètement, soient les échantillons $\{(X_i, Y_i), i = 1, \dots, n\}$ d'un problème de classification telle que $X_i \in X \subset \mathfrak{R}^d$ et $Y_i \in \{-1, 1\}$. On cherche une transformation $\Phi: X \rightarrow F$ tel que les échantillons transformés $\{(\Phi(X_i), Y_i), i = 1, \dots, n\}$ soient linéairement séparables dans F .

- Deuxièmement, la détermination d'un hyperplan permettant une séparation linéaire optimale dans cet espace (F) de grande dimension (Figure 3-18).

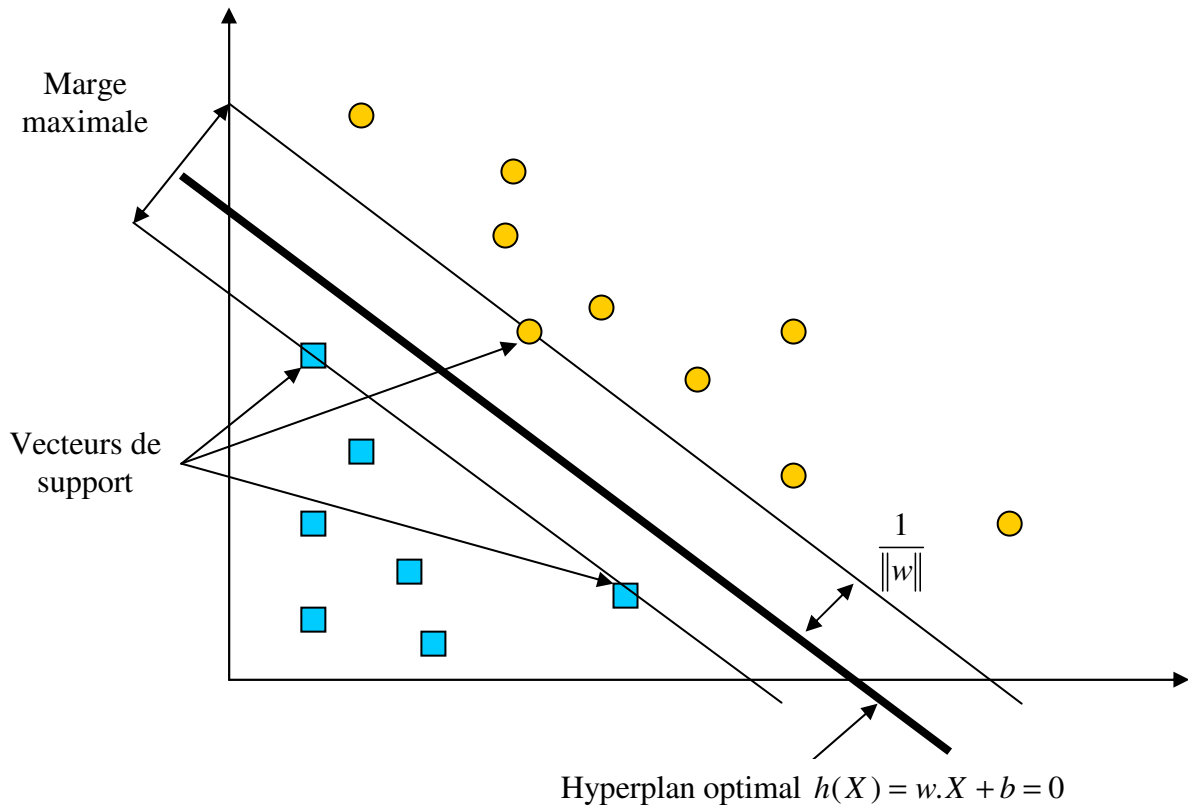


Figure 3-18 : L'hyperplan séparateur optimal est celui qui maximise la marge dans l'espace de redescription

L'intérêt est que des échantillons non séparables linéairement dans l'espace d'entrée peuvent le devenir dans l'espace de redescription. En effet, intuitivement, plus la dimension de l'espace de redescription est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les données est élevée (Figure 3-19).

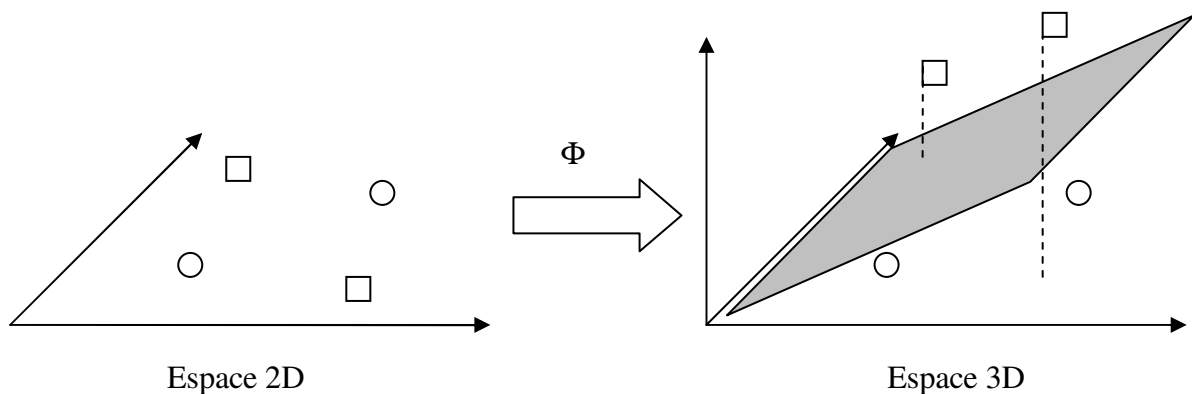


Figure 3-19 : changement d'espace pour transformer une séparation non linéaire en linéaire

Dans l'exemple de la Figure 3-19, les échantillons de l'espace d'entrée se situent tous dans un plan, et ne sont pas séparables par une ligne droite. Dans l'espace de redescription, l'ajout d'une troisième dimension permet de séparer les données par un plan.

La distance d'un point p à l'hyperplan est $d(p) = \frac{|w \cdot p + b|}{\|w\|}$. On cherche un hyperplan tel que sa distance aux points les plus proches soit maximale. Cela revient donc à minimiser $\|w\|$ sous la contrainte que l'hyperplan reste séparateur :

$$\begin{cases} \min \|w\|^2 \\ \forall i, Y_i(w \cdot X_i + b) \geq 0 \end{cases}$$

La résolution de l'optimisation sous contrainte se fait le plus souvent par la méthode des multiplicateurs de Lagrange

$$\begin{cases} L(w, b, \alpha) = \|w\|^2 - \sum_{i=1}^n \alpha_i \{Y_i(w \cdot X_i + b)\} \\ \alpha_i \geq 0, (1 \leq i \leq n) \end{cases}$$

qui ramène au problème dual, dont la résolution est en $O(n)$.

$$\begin{cases} \max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j Y_i Y_j (X_i \cdot X_j) \\ \alpha_i \geq 0, (1 \leq i \leq n) \\ \sum_{i=1}^n \alpha_i Y_i = 0 \end{cases}$$

Pour les problèmes non linéairement séparables, la transformation non linéaire (Φ) est réalisée via une fonction noyau k (ou noyau de Hilbert-Schmidt). Ainsi, tout ce dont nous avons besoin est une façon de calculer le produit scalaire entre les échantillons transformés.

$$\Phi(X_i) \cdot \Phi(X_j) = k(X_i, X_j)$$

La séparation linéaire optimale dans F est alors l'hyperplan optimal d'équation $h(X) = \sum_{i=1}^n \alpha_i Y_i k(X_i, X) + b$ qui maximise la marge. La marge étant vue comme la distance euclidienne entre $\Phi(X_i)$ et la frontière de décision $h(X) = 0$. Cette solution est obtenue en résolvant le problème de programmation quadratique suivant :

$$\left\{ \begin{array}{l} \max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j Y_i Y_j k(X_i, X_j) \\ 0 \leq \alpha_i \leq C, (1 \leq i \leq n) \\ \sum_{i=1}^n \alpha_i Y_i = 0 \end{array} \right.$$

avec C un hyper-paramètre appelé la constante de marge douce, qui permet de travailler à risque empirique non nul. La valeur b se déduisant à l'optimum des conditions de Kuhn-Tucker.

Ainsi, le cas non linéaire est une généralisation du cas linéaire pour lequel la fonction noyau serait le produit scalaire.

$$\Phi(X_i) \cdot \Phi(X_j) = k(X_i, X_j) = X_i \cdot X_j$$

Plusieurs fonctions peuvent être utilisées. Elles doivent vérifier les conditions de Mercer [Mercer, 1909] qui se résument à vérifier que la fonction k corresponde bien à un produit scalaire définissant une norme dans un espace étendu F .

Les principaux noyaux utilisés sont :

- Noyau polynomial : qui est le produit scalaire élevé à une puissance ; la puissance entière L , est un paramètre de l'algorithme d'apprentissage qui détermine les propriétés de la solution.

$$k(X_i, X_j) = (1 + X_i \cdot X_j)^L$$

- Noyau RBF³⁹ appelée fonction de base radiale. C'est une exponentielle du carré de la distance entre les vecteurs, divisé par sigma au carré. C'est la formule d'une cloche gaussienne dans l'espace d'entrée, où sigma est la déviation standard de la gaussienne. Sigma est un paramètre de l'apprentissage qui détermine les propriétés de la solution.

$$k(X_i, X_j) = e^{-\frac{\|X_i - X_j\|^2}{2\sigma^2}} = e^{-\frac{(X_i - X_j) \cdot (X_i - X_j)}{2\sigma^2}}$$

³⁹ RBF : Radial Basis Function

- Noyau sigmoïdal

$$k(X_i, X_j) = \tanh(aX_i \cdot X_j - b)$$

Discrimination multi-classes

Dans le cas de la discrimination multi-classe, nous avons pour un espace de sortie de N classes un $Y_i \in \{c_1, \dots, c_N\}$. Une SVM dans ce contexte réalise des fonctions de séparations linéaires optimales $H = (h_p)$; $1 \leq p \leq N$ définies par :

$$h_p(X) = \sum_{i=1}^n \alpha_{ip} Y_i k(X_i, X) + b_p ; n \text{ étant le nombre d'observations}$$

Un élément de l'espace d'entrée est associé à la classe de l'espace de sortie pour laquelle $h_p(X)$ est maximale. L'ensemble des fonctions $h_p(X)$ est obtenu à nouveau par la résolution d'un problème de programmation quadratique [Darcy et al., 2006].

Dans cette thèse, nous utilisons cette approche pour concevoir un estimateur des préférences d'un élève en mettant en œuvre une discrimination multi-classes par SVM.

Ainsi pour définir le système, nous avons spécifié :

- L'ensemble des variables qui constituent l'espace d'entrée X_1, X_2, \dots, X_E
- L'ensemble des classes qui constituent l'espace de sortie Y_1, Y_2, \dots, Y_S
- La fonction noyau k à utiliser pour transformer les échantillons de l'espace d'entrée X dans l'espace de redescription F .

3.7.2. SVM multi-classes pour estimer les préférences d'un élève

Dans le cas du profilage du style d'apprentissage, nous proposons d'utiliser les variables observables pour définir les éléments de l'espace d'entrée, puis d'utiliser les variables non observables pour décrire l'espace de sortie.

L'espace d'entrée

L'espace d'entrée X est représenté par les variables observables définies par le modèle conceptuel du profilage du style d'apprentissage d'un élève décrit dans la section 2.6. Soient les ensembles \mathcal{T} , \mathcal{M} et \mathcal{P} . Nous les rappelons brièvement ci-dessous :

Le matériel pédagogique support d'une séquence pédagogique se compose d'un ensemble de sept variables aléatoires discrètes $\mathcal{M} = \{M_1, M_2, M_3, M_4, M_5, M_6, M_7\}$ (cf. section 3.6.2).

L'activité pédagogique est décrite par un ensemble de deux variables aléatoires discrètes $\mathcal{T} = \{T_1, T_2, T_3\}$ (cf. section 3.6.2).

Nous considérons que le comportement d'un élève est représenté par un ensemble de trois variables aléatoires continues $\mathcal{P} = \{P^m, TAL, U\}$ (cf. section 3.6.2).

L'espace de sortie

L'espace de sortie Y est représenté par les combinaisons offertes par $\mathcal{D} = \{D_1, D_2, D_3, D_4\}$, l'ensemble de variables non observables représentant les dimensions du style d'apprentissage. Chaque variable représente la projection de l'élève sur l'un des quatre axes du modèle de Felder (Visuel-Verbal, Global-Sequential, Réfléchi-Actif, Déductif-Inductif). Rappelons que chacune des variables $\{D_i\}_{i \in [1,4]}$ prend une valeur entière comprise entre -3 et 3. $D_i \in [-3, -1, 1, 3]; \forall i \in \{1, 2, 3, 4\}$. En considérant que les valeurs -3 et -1 peuvent être assimilées à une extrémité de l'axe et que les valeurs 1 et 3 peuvent être assimilées à l'autre extrémité de l'axe, on obtient 16 classes en sortie.

Structure de l'estimateur de préférences

La structure de l'estimateur met en œuvre deux étages utilisant chacun une SVM différente (Figure 3-20).

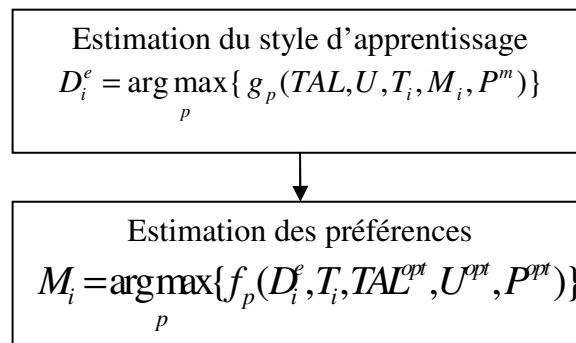


Figure 3-20 : structure de l'estimateur de préférences

Nous posons le problème de l'estimateur du style d'apprentissage d'un élève comme étant un problème de discrimination multi-classe. Dans ce cas, nous avons un espace de sortie de $ng = 16$ classes de style : $Y_i \in \{c_1, \dots, c_{ng}\}$. Une SVM dans ce contexte réalise des fonctions de séparations linéaires optimales $G = (g_p); 1 \leq p \leq ng$ définies par :

$$\arg \max_p (g_p(X) = \sum_{i=1}^n \alpha_{ip} Y_i k(X_i, X) + b_p)$$

Un élément de l'espace d'entrée est associé à la classe de l'espace de sortie pour laquelle $g_p(X)$ est maximale tel que $X = \{T_1, T_2, T_3, M_1, M_2, M_3, M_4, M_5, M_6, M_7, TAL, U, P\}$. Nous avons adopté le noyau polynomial $k(X_i, X_j) = (1 + X_i \cdot X_j)^L$.

Suivant le même principe nous avons défini l'estimateur des préférences d'un élève. Avec nf le nombre total de classes de métadonnées, soit 1152. La SVM utilisée réalise des fonctions de séparation linéaires optimales $F = (f_q); 1 \leq q \leq nf$ définies par :

$$\arg \max_q (f_q(X) = \sum_{i=1}^n \alpha_{iq} Y_i k(X_i, X) + b_q)$$

Un élément de l'espace d'entrée est associé à la classe de l'espace de sortie pour laquelle $f_q(X)$ est maximale telle que $X = \{T_1, T_3, TAL^{opt}, U^{opt}, P^{opt}, D_1, D_2, D_3, D_4\}$. Nous avons adopté la aussi un noyau polynomial $k(X_i, X_j) = (1 + X_i \cdot X_j)^L$.

3.7.3. Validation du modèle

Pour valider le modèle, nous avons implémenté SVM avec le pack matlab SpiderSVM⁴⁰. Les données recueillies lors de l'expérimentation en classe de cinquième du collège Malraux, nous ont permis de tester notre classificateur SVM pour estimer les préférences des élèves basées sur leur style d'apprentissage. Nous rappelons que les données utilisées sont réduites pour la description de l'objectif à T_1 et pour la description du matériel pédagogique à (M_1, M_2) codés par deux modalités chacune (cf. section 2.5.3).

- L'espace d'entrée est alors composé des variables pédagogiques, du matériel et du comportement :

$$X = \{T_1, M_1, M_2, TAL, U, P\}$$

- L'espace de sortie est une matrice carrée 16x16 régulière dont la diagonale est à 1 et le reste des composants à -1. Les lignes et les colonnes représentent les 16 classes. Les éléments de la matrice représentent les intersections des classes : 1 indique une égalité, -1 indique l'ensemble vide :

$$Y = \begin{pmatrix} \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & \mathbf{1} \end{pmatrix}$$

⁴⁰ <http://www.kyb.tuebingen.mpg.de/bs/people/spider/>

- Nous avons adopté le noyau polynomial :

$$k(X_i, X_j) = (1 + X_i \cdot X_j)^3$$

80% des données mesurées ont été utilisées pour entraîner le système et 20% pour le valider. La Figure 3-21 montre les résultats obtenus. On remarque que le système estime à plus de 86% correctement le style des élèves dès le départ de l'expérience. Ensuite, au bout de la quatrième recherche, après l'apprentissage de 4 nouveaux cas, le classifieur estime à 100% le bon style.

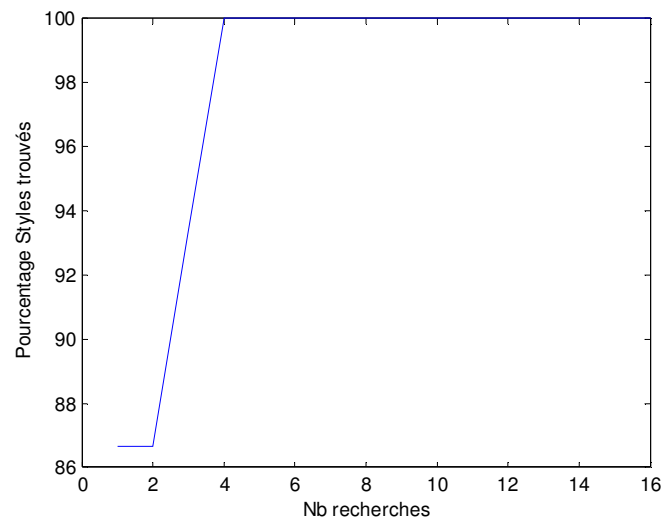


Figure 3-21 : pourcentage de style correctement estimé par SVM

Ensuite nous avons travaillé sur $M_i = f(D_i^e, T_i, TAL^{opt}, U^{opt}, P^{opt})$.

En considérant qu'en maximisant le comportement \mathcal{P} , les données mesurées dans l'espace d'entrée (\mathcal{T}, \mathcal{D}) peuvent être classées dans l'espace de sortie en 4 classes obtenues par l'ensemble $\mathcal{M} = \{M_1, M_2\}$, nous avons réalisé l'expérience suivante.

- L'espace d'entrée, cette fois-ci est composé des variables pédagogiques, du comportement optimal et du style d'apprentissage :

$$X = \{T_i, TAL^{opt}, U^{opt}, P^{opt}, D_1, D_2, D_3, D_4\}$$

- L'espace de sortie est une matrice 4x4 dont la diagonale est à 1 et le reste à -1 :

$$Y = \begin{Bmatrix} \mathbf{1} & -1 & -1 & -1 \\ -1 & \mathbf{1} & -1 & -1 \\ -1 & -1 & \mathbf{1} & -1 \\ -1 & -1 & -1 & \mathbf{1} \end{Bmatrix}$$

- Le noyau est aussi polynomial :

$$k(X_i, X_j) = (1 + X_i \cdot X_j)^3$$

La Figure 3-22 présente le pourcentage de \mathcal{M} correctement trouvé. On observe que 100% des \mathcal{M} sont correctement estimés à chaque recherche.

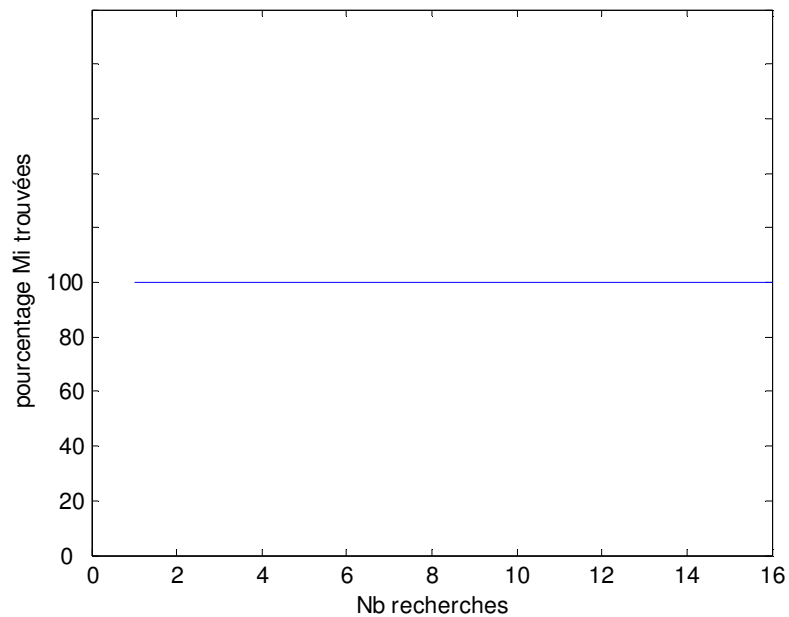


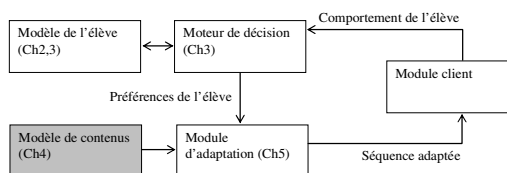
Figure 3-22 : pourcentage de \mathcal{M} correctement estimé par SVM

Il est facile d'observer que l'on obtient de meilleurs résultats avec SVM qu'avec le RB. Nous expliquons cela par le fait que SVM travaille uniquement avec les points de support pour classer une nouvelle entrée. La classification est donc faite par une mesure de proximité de la nouvelle entrée par rapport aux hyperplans définis par les points de support. Le fait que cette entrée ait déjà été observée ou pas n'a aucune influence sur sa classification.

3.8. Conclusion

Dans ce chapitre, nous avons présenté une approche qui utilise un modèle de l'élève scindé en un modèle de connaissances et un modèle de préférences. Le modèle de connaissances est utilisé pour séquencer les objectifs d'apprentissage alors que le modèle de préférences est utilisé pour estimer le style d'apprentissage de l'élève et inférer ses préférences pédagogiques. Le modèle de préférences a été conçu comme une instance du modèle conceptuel du processus de profilage que nous avons proposé (cf. paragraphe 2.2). En effet, il en hérite la structure des dépendances. Cette structure a été complétée pour établir un modèle probabiliste. Nous avons montré la méthode que nous employons pour apprendre ce modèle. Celle-ci se compose de deux étapes : un apprentissage supervisé pour entraîner le modèle initial ; puis un apprentissage par renforcement pour continuer à l'affiner tout au long de son utilisation. A la base de ce modèle, nous avons établi l'architecture de notre système adaptatif. Plus précisément, nous avons conçu le moteur de décision qui exploite le modèle de l'élève. Celui-ci consiste en deux étages : le séquenceur d'objectifs et l'estimateur de style. Le deuxième étage prend en entrée le modèle de préférences et le comportement observé de l'élève. Il procède en deux étapes. Premièrement, il renforce la croyance sur le style. Deuxièmement, il infère les préférences pédagogiques de l'élève. Ces préférences sont passées au module d'adaptation pour composer la séquence à suivre. Nous avons proposé deux implémentations différentes du modèle de préférences. La première est un réseau bayésien dont on a appris les paramètres initiaux à l'aide des données de notre étude empirique. La deuxième est un classificateur fondé sur le formalisme de machine à vecteurs de support. La comparaison des deux solutions a montré que SVM apporte de meilleurs résultats.

Le chapitre suivant réutilisera les variables définies dans le modèle de préférences pour indexer le contenu. Le chapitre 5 montrera comment les préférences sont mappées sur le modèle de contenus pour recommander la séquence pédagogique adaptée.



CHAPITRE

4. Contenu adaptatif : modèle, méthode et implémentation

Résumé :

Ce chapitre traite de la conception, la création et l'implémentation du modèle de contenus de notre système adaptatif. Nous y proposons l'utilisation d'une méta-ontologie composée de trois modèles conceptuels : métier, pédagogique et de ressources. Une méthode d'ingénierie est proposée pour créer les contenus en instanciant la méta-ontologie.

Nous proposons aussi SCORM CAM comme moyen technique pour structurer l'ontologie et les ressources. Nous améliorons les métadonnées LOM pour implémenter les informations nécessaires à l'adaptation. Nous montrons finalement la mise en œuvre de cette approche dans le projet européen UP2UML.

4.1. Introduction

De nos jours, la production de contenus pédagogiques vise d'une part à améliorer l'efficacité économique des investissements et d'autre part à améliorer l'efficacité pédagogique des objets d'apprentissage produits. L'ingénierie dans ce domaine est donc devenue une nécessité compte tenu des enjeux. Dans [Tchounikine, 2002], l'auteur apporte une définition de ce que peut être une ingénierie dans le domaine de l'apprentissage humain : « *Travaux visant à définir des concepts, méthodes et techniques reproductibles et/ou réutilisables facilitant la mise en place (conception, réalisation, expérimentation, évaluation et diffusion) d'environnements de formation ou d'apprentissage en permettant de dépasser le traitement ad hoc des problèmes* ».

Dans [Labat et al., 2006], les auteurs précisent que la conception pédagogique souffre d'une absence de cadre méthodologique pour envisager le développement complet et opérationnel de solutions de formation efficaces. Dans [Merrill, 2002], les auteurs ont proposé un processus largement inspiré de méthodes de génie logiciel : le modèle du « Caillou dans la mare ». C'est un processus incrémental décomposé en cinq phases (expression du problème, analyse, stratégie, conception et production), où chaque phase correspond à une « ondulation » croissante provoquée par le jet initial d'un projectile (le problème à résoudre). Une autre approche se fonde sur les pratiques mises en œuvre dans l'ingénierie pédagogique. Dans [Paquette et Tchounikine, 2002], les auteurs présentent l'ingénierie pédagogique comme une méthodologie soutenant l'analyse, la conception, la réalisation et la planification de l'utilisation des systèmes d'apprentissage. Dans ce domaine, on peut citer les travaux du laboratoire Licef à Québec [Paquette, 2002] : le système Adisa (atelier distribué d'ingénierie d'un système d'apprentissage), la méthode Misa (méthode d'ingénierie des systèmes d'apprentissage) et l'outil de modélisation graphique Mot (modélisation par objets typés).

Néanmoins, dans l'ensemble de ces méthodes, on déplore l'absence de prise en compte du paradigme des objets d'apprentissage [Wiley, 2000] qui est une question centrale dans le cadre de la production de contenus pédagogiques. Certains auteurs, mettent en avant les questions de production d'objets d'apprentissage, de granularisation, d'indexation et de réutilisation [Downes, 2000 ; Koper, 2001 ; Catteau, Vidal et al., 2007] qui doivent être prises en considération dans tout processus de production. Les travaux d'ingénierie visent à rationaliser de tels processus de production. Les normes développées actuellement comme Lom⁴¹ et Scorm⁴² offrent un cadre de représentation pour les objets pédagogiques impliqués, facilitant interopérabilité et réutilisation [Tchounikine, 2006].

Dans ce contexte d'ingénierie, les ontologies trouvent tout naturellement leur place. Dans [Desmoulins et Grandbastien, 2006], les auteurs considèrent que développer une ingénierie, consiste à produire un ensemble de connaissances (méthodes, techniques, outils) qui guideront le processus de conception et de réalisation. En accord avec Newell [Newell, 1981], ils présentent les ontologies comme un modèle de connaissances et de résolution de problèmes. Dans [Bourdieu, 1980], l'auteur ajoute que ce modèle de connaissance représente un consensus d'une communauté de pratique. Comme tout modèle [Baker, 2000], une ontologie peut avoir deux fonctions :

⁴¹ Lom : Learning Object Metadata

⁴² Scorm : Sharable Content Object Reference Model

- Proposer un cadre conceptuel permettant au concepteur de représenter le monde, en particulier en vue de faire des conjectures sur son comportement.
- Fournir des représentations, manipulables par l'humain ou par la machine, utiles au soutien de l'activité manuelle ou automatique.

Le projet IMAT [Desmoulins et Grandbastien, 2002] s'inscrit dans cette logique. Il utilise les ontologies comme guides pour le découpage, l'indexation et la recherche de fragments de manuels techniques dans le but de concevoir un document de formation. Ces ontologies modélisent plusieurs aspects. On trouve des ontologies relatives à la forme des documents, des ontologies relatives au domaine de travail et des ontologies relatives aux aspects pédagogiques. Les auteurs qualifient ces ontologies de semi-formelles. Ce sont des structurations de concepts à base principalement de relations hiérarchiques *est-un*, *est-partie-de*. On y trouve accessoirement d'autres relations comme *x décrit y* ou *x précise y*. Ces ontologies ne contiennent pas d'axiomatique précisant le sens des concepts. L'ontologie sur la forme des documents décrit de façon générique le concept de fragment. On y trouve par exemple les termes *corps de texte*, *index*, *glossaire*, *photo*, *figure*, *équation*,... L'ontologie du domaine, dans le cas de la carrosserie automobile, décrirait les composants d'une voiture (capot, aile, ...), les outils de réparation et les savoirs-faire de réparation. L'ontologie pédagogique contient la description du type de connaissance (*Metacognition*, *declarative knowledge*, *perceptual knowledge*, *procedural knowledge*, *performing skills*...) [Kabel et al., 2001] et la description des séquences pédagogiques en termes d'objectifs pédagogiques (*being able to maintain ...*, *being able to reason about ...*, *being able to perform ...*) et d'activités pédagogiques. L'objectif étant d'utiliser ces ontologies comme indexation des fragments de manuels techniques, sur des critères techniques, structurels ou pédagogiques.

Dans la suite de ce chapitre, nous proposons un cadre formel de production de contenus pédagogiques. Le produit de cette méthode est un ensemble d'objets d'apprentissage indexé par une ontologie du domaine et une ontologie pédagogique. Dans un premier temps, nous décrivons le modèle conceptuel de contenus que nous manipulons dans notre processus d'ingénierie d'objets d'apprentissage. Ensuite, nous détaillons le dit processus en décrivant les phases qui le composent. Nous poursuivons par l'étude de l'aspect structuration standardisée des objets d'apprentissage. Enfin nous terminons, par la validation et la mise en œuvre des contenus pédagogiques produits.

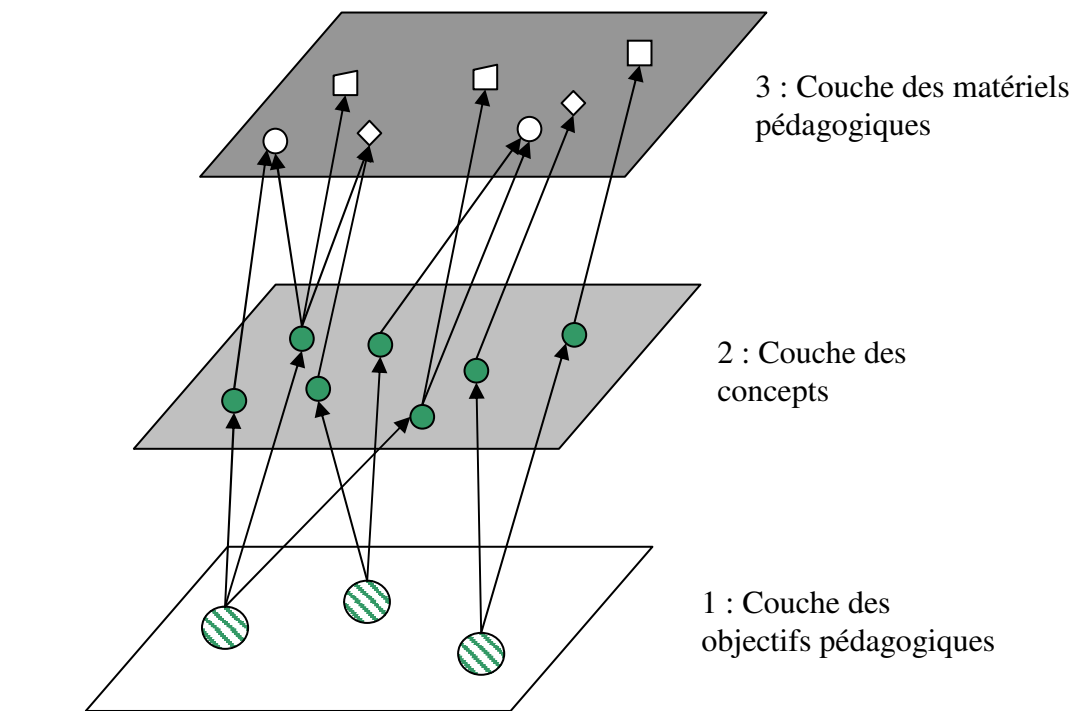
4.2. Modèle de contenus

Les modèles de contenus des systèmes adaptatifs d'enseignement décomposent le domaine à étudier en éléments de connaissances appelés de façon générique des concepts [Brusilovsky, 2003d]. Ces concepts peuvent être connectés entre eux pour faire apparaître des relations de différents types, tels que « est composé de » ou « est similaire à » [Trella et al, 2002] utilisées pour adapter la navigation dans les contenus [Cristea et al, 2004c].

Il est souvent admis qu'un ensemble de concepts doit être traité par l'élève pour considérer qu'un certain objectif d'apprentissage est atteint (ex : framework LAOS) [Cristea et al, 2003a]. Le niveau de connaissance étant alors estimé par une mesure de performance ou simplement qualifié par un état : visité/non visité [Cristea et al., 2004c]. Chaque concept est généralement associé à du matériel pédagogique que l'élève utilise dans l'activité d'apprentissage qui lui est proposée par le système (ex : système INSPIRE)[Papanikolaou et al, 2003].

La structure des modèles de contenus est communément composée de trois couches reliées entre elles (Figure 4-1) [Papanikolaou et al., 2003] :

1. La couche des objectifs d'apprentissage,
2. La couche des concepts,
3. La couche des matériels pédagogiques.



Légende :

- | | |
|--|--|
|  : Objectif pédagogique |  : présentation théorique |
|  : Concept |  : exercice |
|  : évaluation |  : exemple |

Figure 4-1 : structure des modèles de contenus

Cette structure comporte principalement deux limitations :

- Les objectifs pédagogiques ne sont pas directement liés à des activités métier. Cela rend particulièrement difficile l'indexation de contenus adaptés à l'apprentissage d'un métier. La réussite d'un apprentissage passe par l'étude et la modélisation de la nature particulière des contenus en jeu [Balacheff et al., 2006]. Par exemple, la didactique en milieu professionnel [Pastre, 2002],
- Les concepts sont reliés à un ensemble figé de matériels pédagogiques. La variation de séquences pédagogiques composées de différents matériels est difficilement réalisable.

Dans cette section, nous proposons un modèle de contenus (Figure 4-2) qui permet d'indexer les ressources pédagogiques par une ontologie du domaine et une ontologie pédagogique. Nous exploitons dans ce sens, l'idée avancée par Brousseau [Brousseau, 1998]. L'auteur soutient qu'une bonne connaissance du domaine associée à des principes pédagogiques généraux permet d'assurer les conditions d'un apprentissage efficace.

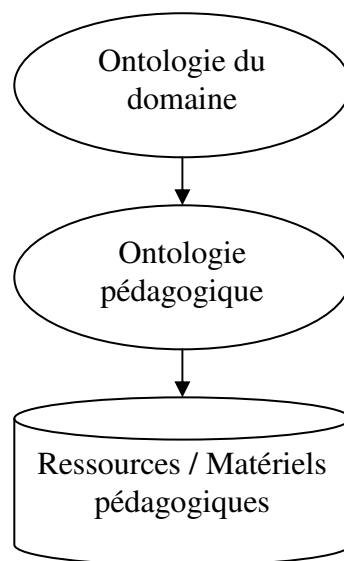


Figure 4-2 : Approche conceptuelle du modèle de contenus

Le modèle de contenus se compose de trois modèles conceptuels :

- Un modèle conceptuel du domaine qui est utilisé pour créer l'ontologie du domaine à enseigner.
- Un modèle conceptuel pédagogique qui est utilisé pour créer l'ontologie pédagogique.
- Un modèle conceptuel d'objets d'apprentissage qui sert à décrire les ressources pédagogiques.

4.2.1. Modèle conceptuel du domaine

Une approche pour l'acquisition de connaissances consiste à s'appuyer sur l'idée d'objectifs à atteindre [De Landsheere, 1989]. La notion d'objectif est dans ce contexte proche de la notion de but telle que conçue en intelligence artificielle. Dans [Peachey et McCalla, 1986], les auteurs proposent d'appliquer une planification descendante, qui consiste à décomposer un objectif en sous-objectifs.

En se basant sur cette approche, nous considérons que le modèle conceptuel du domaine doit offrir une description des objectifs d'apprentissage associée à une description des tâches métiers (Figure 4-3). Le domaine à enseigner est donc décomposé en tâches métier à réaliser par des acteurs du domaine. Une tâche métier est ensuite décomposée en un ensemble d'objectifs globaux. Chaque objectif global correspond à une compétence métier qui doit être acquise pour mener à bien la tâche métier correspondante. Ces objectifs globaux peuvent eux-mêmes être itérativement décomposés en sous-objectifs globaux afin de faciliter l'acquisition et la gestion des compétences.

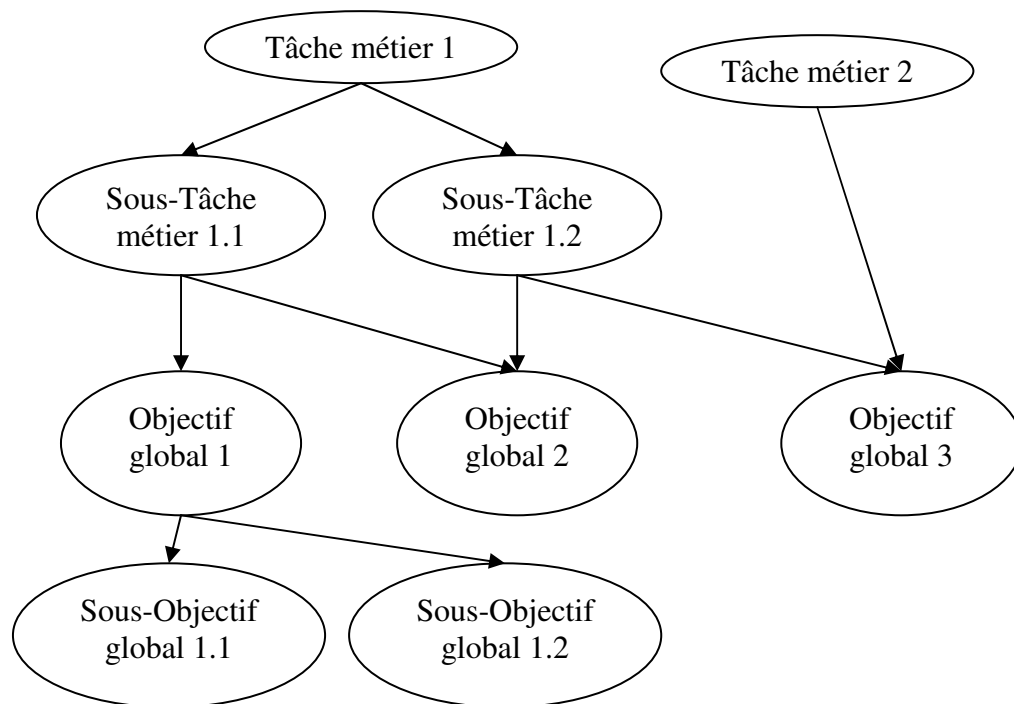


Figure 4-3 : exemple de décomposition du domaine

4.2.2. Modèle conceptuel pédagogique

La recherche d'efficacité dans le processus d'enseignement se base sur la mise en relation des objectifs de l'apprentissage et des méthodes d'enseignement [Gagné, 1985]. Dans [Pernin et Lejeune, 2004], les auteurs proposent de découper l'enseignement en unités élémentaires pour permettre un séquençement pertinent. Au moins trois niveaux sont à considérer :

- Niveau 1 : Activité élémentaire permettant à un élève de remplir une tâche précise (ex : lire un texte, faire un exercice...)
- Niveau 2 : Séquence d'activités élémentaires organisées temporellement en un ensemble homogène pour l'élève (ex : séquence pédagogique)
- Niveau 3 : unités de structuration permettant d'organiser les séquences pédagogiques en un ensemble structurellement pertinent (ex : unité d'enseignement, module...)

Dans [Tchounikine, 2006], l'auteur ajoute à l'identification des objectifs d'apprentissage, le besoin de spécifier et modéliser l'activité proposé à l'élève, ce qui nécessite la conception de la tâche à réaliser, l'identification des acteurs (apprenants, co-apprenants, enseignant, tuteurs, formateurs) ainsi que la spécification des outils à utiliser.

Dans notre modèle conceptuel pédagogique, nous considérons que chaque objectif global du modèle du domaine est décomposé en objectifs finaux (Figure 4-4). Un objectif final correspond à l'objectif pédagogique à atteindre pour considérer la compétence métier acquise. C'est un élément essentiel qui va permettre de définir et de mesurer les acquisitions. Il indique avec précision l'action attendue (un verbe qui représente un comportement observable qui respecte la taxinomie de Bloom), les conditions dans lesquelles la performance doit se réaliser (avec quel moyen, dans quel contexte, avec quelle aide, quelle référence l'élève pourra utiliser) et les critères de performance (rapidité, précision, qualité).

Nous caractérisons un objectif final par un niveau de complexité de la tâche exprimé selon la taxinomie de Bloom (connaissance, compréhension, application, analyse, synthèse, évaluation) et un type de connaissances défini selon la classification de Gagné (concept, fait, processus, principe, procédé). Chaque objectif final est associé à une ou plusieurs séquences pédagogiques. Le contenu de la séquence est corrélé au niveau de difficulté de la tâche à réaliser (Bloom) ainsi qu'au style d'apprentissage visé. Il existe donc plusieurs séquences pédagogiques pour un même objectif final.

Une séquence pédagogique définit si l'élève peut naviguer librement dans la séquence ou bien s'il est soumis à un ordonnancement particulier. Chaque séquence pédagogique est décomposée en activités pédagogiques.

Une activité pédagogique est décrite par un certains nombre de caractéristiques :

- *Méthode* : définit la méthode pédagogique utilisée pour l'activité. Elle peut être expositive, interrogative, active.
- *Scénario* : un texte qui décrit l'interaction entre les acteurs pendant l'activité.
- *Description* : une description générale de l'activité.
- *Durée* : le temps estimé pour réaliser l'activité.
- *Synchronisation* : certaines activités sont qualifiées de synchrone, comme par exemple des visioconférences, alors que d'autres sont qualifiées d'asynchrone comme une discussion sur un forum.
- *Acteur* : une activité peut concerner un travail individuel (ex : un élève) ou bien plusieurs acteurs. On considère alors un travail à deux (ex : deux élèves), un travail en petit groupe (ex : 3 à 4 élèves) ou un travail en groupe classe (ex : la classe entière).
- *Feedback* : certaines activités requièrent un retour d'information. Cette information peut être donnée par le système dans le cas d'activités autonomes, par un tuteur dans le cas d'une activité guidée ou bien par des pairs dans le cas d'activités collaboratives.
- *Outil* : certaines activités nécessitent l'utilisation d'outils spécifiques. Par exemple, la validation d'un diagramme UML par simulation requiert un outil de simulation. Cet outil devra être disponible pendant l'activité.

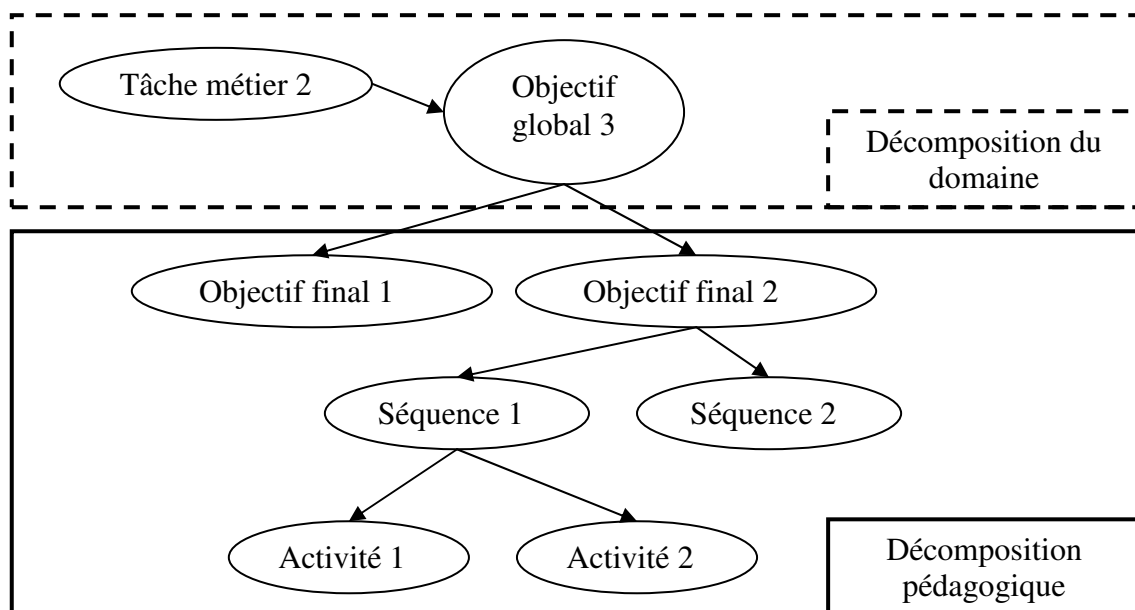


Figure 4-4 : exemple de décomposition pédagogique

4.2.3. Modèle conceptuel d'objets d'apprentissage

Dans notre modèle de contenus, nous considérons qu'à chaque activité est associé un ou plusieurs matériels pédagogiques qu'on appelle des ressources (Figure 4-5). Une ressource est vue comme un objet d'apprentissage. Il peut être un (ou plusieurs) fichier HTML, une animation FLASH, etc. Dans le cas où une activité est associée à plusieurs ressources, chacune vise à répondre au mieux à différents styles d'apprentissage. Une ressource est caractérisée par :

- *URI* : uri de la page ou du fichier à charger.
- *Structure* : il existe deux types de ressources : « composite » et « atomique ». Les ressources « composites » sont composées d'un ensemble de ressources « atomiques ». Une ressource « atomique » est un élément non décomposable. Il peut être une image, une animation, un son, etc.
- *Représentation* : précise le type de représentation principalement utilisé dans la ressource. Cette information est utilisée pour adapter l'activité au style d'apprentissage de l'élève. Elle peut prendre la valeur texte, son, dessin, schéma, image, vidéo, tableau, diagramme, donnée chiffrée.
- *Technique* : les ressources sont classées par leur type. Les différents types utilisés sont : définition, résumé, exemple, contre-exemple, exercice formatif, exercice sommatif.

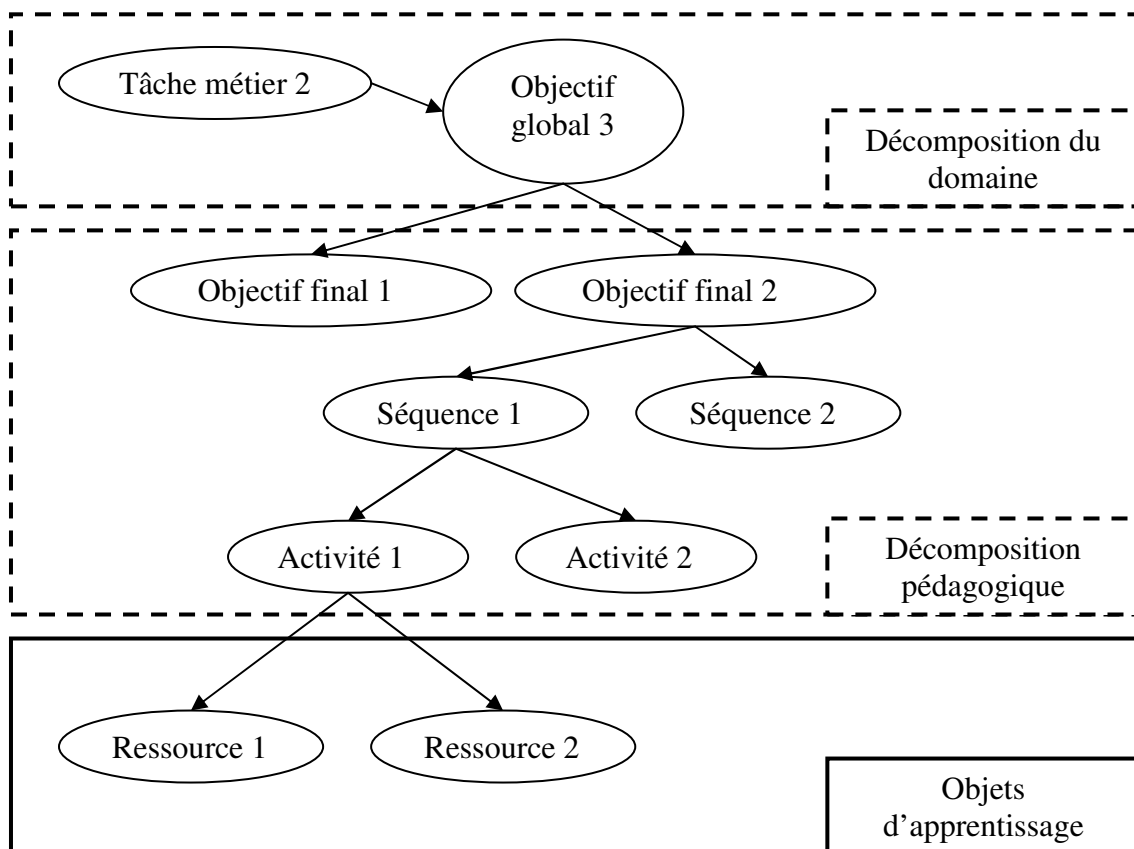


Figure 4-5 : exemple d'objets d'apprentissage

4.2.4. Décomposition Structurale du modèle de contenus

La structure du modèle de contenus fait apparaître deux sortes de relations :

- *Est composé de* : cette relation exprime la notion de composition qui existe entre deux éléments du modèle (un contenant et un contenu). Elle est utilisée pour définir la décomposition structurale du modèle de contenus. Elle montre ce qui est nécessaire de savoir ou savoir faire pour considérer une tâche ou un objectif comme atteint. Elle est utilisée aussi pour décrire le contenu d'une séquence ou d'une activité.
- *Influence* : cette relation exprime la notion de dépendance qui existe entre deux éléments du modèle (précédence dans le temps). Elle est utilisée pour définir l'ordonnancement des tâches métiers ou des objectifs dans une séquence d'apprentissage.

La Figure 4-6 montre une représentation UML de la décomposition structurale du modèle de contenus. Elle est basée sur l'exploitation de la relation de type « est composé de ». On observe qu'une tâche peut être décomposée en sous-tâches et se compose au minimum d'une compétence métier (objectif global). Un objectif global peut se décomposer en sous-objectifs globaux et se compose au minimum d'un objectif pédagogique (objectif final). Un objectif final peut être associé à une ou plusieurs séquences spécifiques adaptées à différents styles d'apprentissage. Chaque séquence est un ensemble d'une ou plusieurs activités qui utilisent elles-mêmes des ressources.

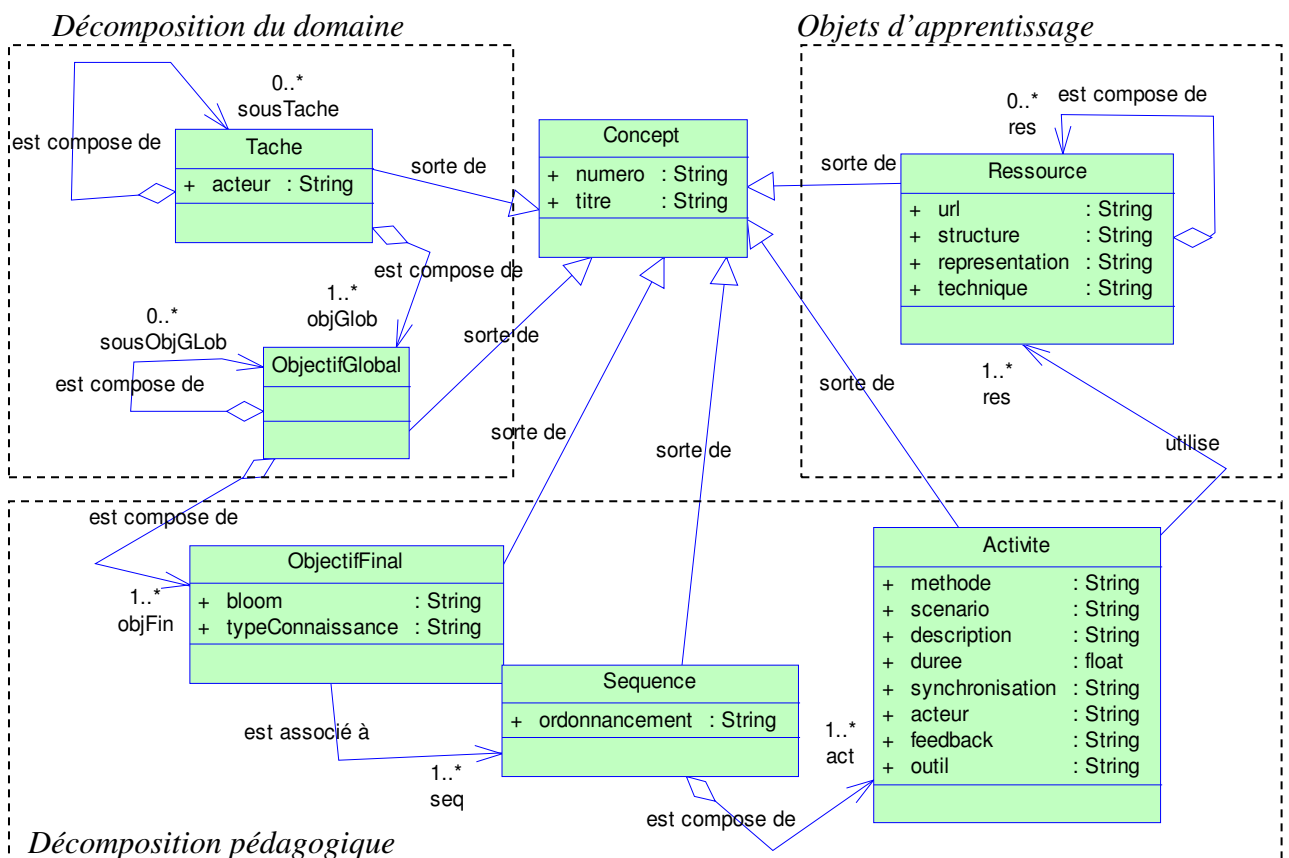


Figure 4-6 : représentation UML de la décomposition structurale du modèle de contenus

Dans ce schéma, *Concept* est une classe abstraite offrant le moyen de structurer les différents éléments. Ce modèle conceptuel peut être assimilé à une méta-ontologie (Figure 4-7) qui peut servir à décrire un cours. Le code OWL de cette ontologie se trouve en Annexe 5.

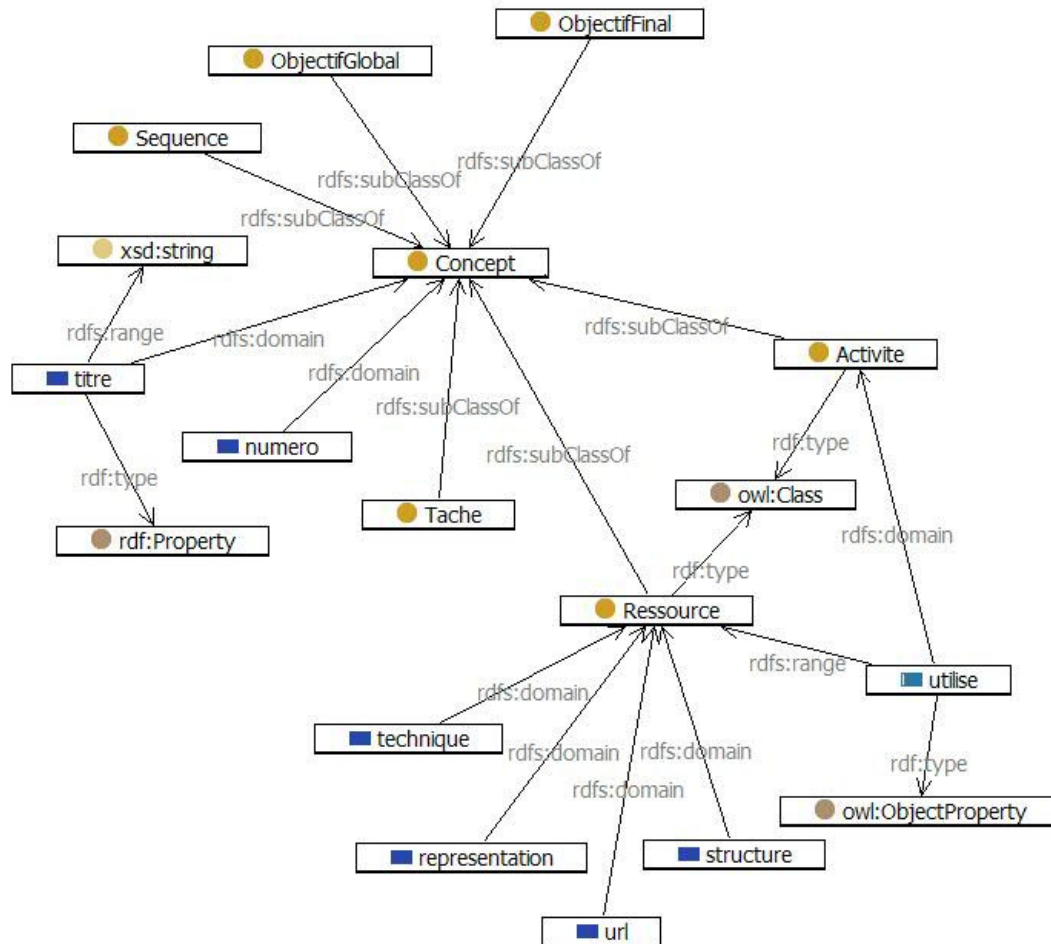


Figure 4-7 : vision parcelaire de la méta-ontologie du modèle de contenu

La Figure 4-8 présente une représentation UML des dépendances du modèle de contenus. Elle est basée sur l'exploitation de la relation de type « influence ». On observe qu'une tâche peut influencer la réalisation d'une ou plusieurs autres tâches. Un objectif global peut influencer la maîtrise d'un ou de plusieurs autres objectifs globaux. Il en est de même pour les objectifs finaux. La relation d'influence entre objectifs est principalement une relation de pré-requis.

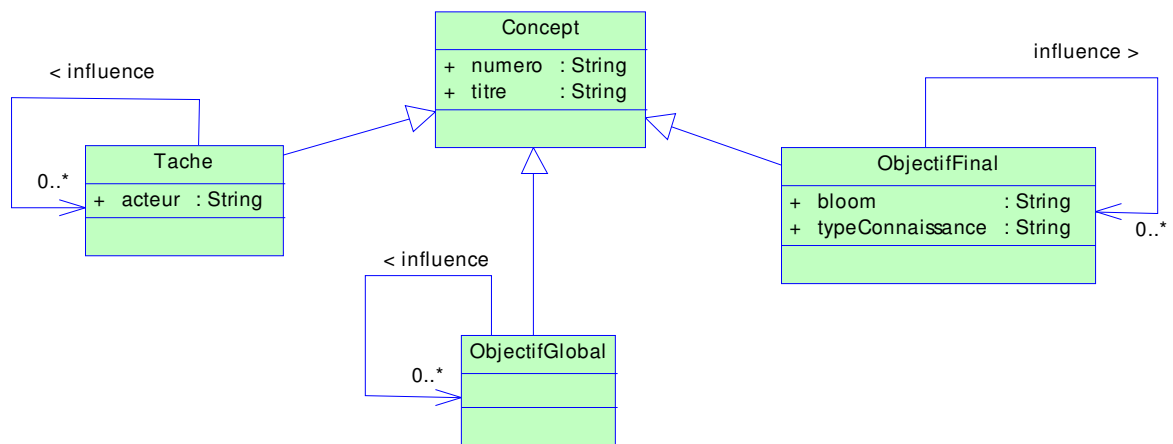


Figure 4-8 : représentation UML des dépendances du modèle de contenus

4.3. Méthode de production des contenus

La conception d'un modèle de contenus nécessite une méthode de production similaire à ce que l'on trouve dans le domaine de la conception de logiciels [Merrill, 2002]. Dans cette thèse, nous définissons un cycle de vie du modèle de contenus, des artefacts à produire tout au long de ce cycle de vie et des outils à utiliser.

Ce cycle de vie est composé de cinq phases (Figure 4-9). Son originalité se fonde sur l'utilisation des ontologies comme artefacts de communication entre les phases. En effet, chaque phase produit ou enrichit un des trois types d'ontologies dont les modèles conceptuels ont été décrits plus haut.

- *Recueil des besoins de formation* : cette phase consiste à identifier les tâches métier et les compétences qui leurs sont associées pour produire l'ontologie du domaine.
- *Analyse pédagogique* : cette phase consiste à déterminer pour chaque compétence métier des objectifs pédagogiques et les activités qui leurs sont associées pour produire l'ontologie pédagogique.
- *Conception pédagogique* : cette phase consiste à organiser/structurer les objectifs pédagogiques en unités d'apprentissage (cours, leçons,...).
- *Développement du matériel pédagogique* : cette phase consiste à créer les ressources pédagogiques que l'élève utilisera pendant la séquence d'apprentissage.
- *Validation et tests du matériel pédagogique* : cette phase consiste à exploiter le matériel pédagogique créé sur un échantillon d'élèves pour relever et corriger les erreurs de conception et de développement mais aussi pour vérifier que les besoins initiaux de formation sont bien couverts.

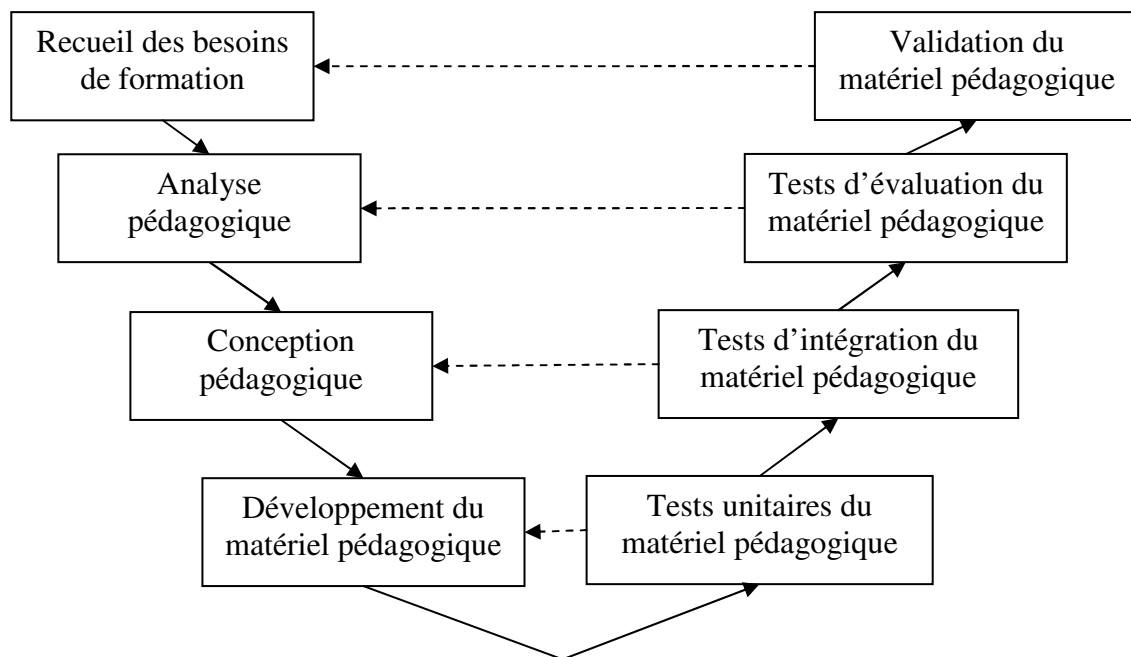


Figure 4-9 : cycle de vie du modèle de contenus

Nous avons appliqué cette méthode dans le cadre du projet européen UP2UML (cf. section 4.5). Un aperçu des artefacts produits par les différentes phases se trouve en Annexe 6. Ces phases sont décrites en détail dans les sections suivantes.

4.3.1. Recueil des besoins de formation du métier

La phase de recueil des besoins de formation consiste à identifier les tâches métier et les compétences métier qui leurs sont reliées.

L'analyse des tâches métier se décompose en trois points :

- identifier les relations entre Acteurs, Tâches et artefacts métier : Cette étape consiste à identifier l'ensemble des tâches métier que doit réaliser un acteur du domaine. Pour chaque tâche, on exprime les artefacts métier qui doivent être produits.
- proposer un modèle de décomposition structurelle des tâches métier : Cette étape consiste à découper l'ensemble des tâches métier en sous-tâches élémentaires. Il est alors possible d'identifier le processus d'apprentissage à suivre pour mener l'élève jusqu'à la maîtrise de la tâche.
- proposer un modèle de dépendances des tâches métier : La dépendance des tâches métier présente les relations de précédence entre une ou plusieurs tâches. Lorsqu'une tâche A est désignée comme dépendante d'une tâche B, cela signifie dans notre modèle que l'élève doit maîtriser la tâche B avant de pouvoir étudier la tâche A.

Après avoir identifié les tâches métier, il faut établir l'ensemble des compétences et connaissances requises pour chaque tâche. Elles sont formulées par des objectifs globaux. Un objectif global exprime un savoir ou un savoir faire, c'est-à-dire par exemple la connaissance d'un concept ou bien la capacité à appliquer une méthode.

Dans cette phase, nous créons une ontologie semi-formelle [Desmoulins et Grandbastien, 2002] en utilisant l'outil TM4L (Topic Maps for Learning) [Dicheva et al, 2006]. Cet outil permet de créer et de visualiser un graphe de concepts au format XTM (Xml Topic Map).

L'ontologie du domaine (Annexe 7) générée se compose de :

- La décomposition structurelle des tâches métier
- La décomposition structurelle des compétences métier
- Les dépendances entre les tâches métier
- Les dépendances entre les compétences métier

4.3.2. Analyse pédagogique

L'analyse pédagogique consiste à exprimer les objectifs pédagogiques associés à chaque objectif global (ou compétence métier), puis à décrire pour chaque objectif pédagogique une ou plusieurs séquences d'activités d'apprentissage.

Un objectif pédagogique correspond dans notre modèle à un *objectif final*. L'objectif final sera exploité par le système pour mesurer le niveau d'acquisition d'une compétence métier (objectif global). Les verbes utilisés dans l'expression de l'objectif pédagogique sont liés aux niveaux de complexité de la tâche d'apprentissage définis selon la taxinomie de Bloom. Nous associons à chaque verbe un scénario particulier qui est défini dans le patron de la séquence.

Une ou plusieurs séquences d'activités sont conçues pour chaque objectif pédagogique. A chaque niveau de complexité de la tâche est associée une ou plusieurs séquences d'activités différentes. Nous avons créé des patrons (ou templates) de séquence que nousinstancions pour chaque objectif pédagogique.

Les patrons de séquence sont composés de trois phases d'apprentissages :

- Phase *expositive* : l'élève apprend par la lecture ou bien l'écoute de définitions, d'exemples ou de contres exemples.
- Phase *active* : l'élève apprend par la pratique. Il réalise des exercices ou des expériences. A la fin de l'exercice, le système met en évidence les erreurs commises par l'élève. Nous avons utilisé deux types d'exercices :
 - *Déplacer-déposer* : l'exercice consiste à proposer à l'élève de choisir dans une liste des objets, et de les déplacer pour les déposer dans des zones libres prévues à cet effet.
 - *Compléter-blanc* : l'exercice consiste à proposer à l'élève de remplir des zones de saisie avec son propre vocabulaire.
- Phase *interrogative* : deux sortes d'interrogations sont proposées à l'élève.
 - *Interrogation d'évaluation formative* : l'élève apprend en ayant une activité réflexive sur son niveau de connaissances. Le système lui pose des questions et lui propose de rejouer les séquences d'apprentissage qui sont liées aux réponses erronées.
 - *Interrogation d'évaluation sommative* : l'élève répond à un questionnaire. Les réponses sont évaluées et l'élève reçoit une note d'évaluation de son niveau de connaissances.

- Dans les deux cas, nous avons utilisé les trois types de tests suivants :
 - *Test Vrai/Faux* : il consiste à proposer à l'élève un questionnaire à choix multiples composé de questions fermées. L'élève n'a pas d'autres solutions que de répondre vrai ou faux à la question.
 - *QCM simple* : il consiste à proposer à l'élève un questionnaire à choix multiples composé de questions ouvertes auxquelles il est possible de cocher une seule réponse parmi les solutions offertes.
 - *QCM multiple* : il consiste à proposer à l'élève un questionnaire à choix multiples composé de questions ouvertes auxquelles il est possible de cocher plusieurs réponses parmi les solutions offertes.

En pratique, cette phase produit l'ontologie semi-formelle pédagogique (Annexe 8) qui se compose de :

- La décomposition structurelle des objectifs pédagogiques
- Des dépendances entre objectifs pédagogiques
- Des associations entre objectifs pédagogiques et les séquences pédagogiques

4.3.3. Conception pédagogique

La conception pédagogique consiste à organiser en unités d'apprentissage les objectifs pédagogiques et à faire la maquette des écrans utilisés pour chaque activité pédagogique.

Les unités d'apprentissage sont les éléments structurels des contenus pédagogiques. Ils regroupent un ensemble d'objectifs pédagogiques en une structure conceptuelle qui possède deux niveaux de granularité :

- Le *cours* : un cours est un ensemble de leçons. Il est vu comme un simple conteneur.
- La *leçon* : une leçon est un ensemble d'objectifs pédagogiques.

Lors de cette phase, il est possible d'utiliser des unités d'apprentissage existantes pour les associer à l'ensemble des unités créées réellement dans ce processus. Il est donc permis de réutiliser dans plusieurs contextes d'apprentissage des cours ou leçons déjà conçus.

La maquette des activités pédagogiques se compose d'un ensemble de représentations graphiques présentant au mieux les éléments à faire apparaître dans les écrans qui seront ensuite développés. Nous avons défini pour chaque type d'apprentissage (ex : expositive, active, interrogative) un modèle d'écran à réaliser. Ce modèle sert de base à la réalisation des ressources physiques dans la phase de développement du matériel pédagogique.

4.3.4. Développement du matériel pédagogique

La phase de développement du matériel pédagogique consiste à réaliser l'ensemble des ressources physiques associées à chaque activité. Ces ressources physiques sont les objets d'apprentissage qui seront exploités par des plateformes d'enseignement pour être présentées à l'élève en cours de séquences pédagogiques. Les concepteurs utilisent les scénarios (ainsi que les autres caractéristiques) des activités contenus dans l'ontologie pédagogique ainsi que les écrans de la maquette, pour créer les ressources à l'aide d'outils de création de contenus.

L'ensemble des ressources créées, ainsi que leur description sont structurés selon l'organisation proposée par ADL dans le modèle de référence SCORM [ADL SCORM, 2004].

4.4. Structuration standard des contenus

Dans [Labat et al., 2006], l'auteur présente une synthèse sur les normes et standards pour la conception, la production et l'exploitation des environnements informatiques pour l'apprentissage humain. Les nombreux travaux présentés mettent en exergue le concept d'objets d'apprentissage comme étant une ressource importante des dispositifs d'apprentissage. C'est pourquoi, il existe dans ce contexte de nombreux travaux qui contribuent aux efforts de normalisation. On trouve ainsi LOM qui cherche à normaliser les langages d'indexation des données pédagogiques ou bien SCORM qui s'attache à normaliser les protocoles de diffusion, d'échange et de suivi d'objets d'apprentissage.

Dans cette thèse, nous proposons pour la première fois d'adopter *SCORM Content Agregation Model* comme langage pour décrire un cours avec ses ontologies et ses ressources. Nous proposons aussi une amélioration des métadonnées LOM pour décrire nos contenus adaptatifs. Dans un premier temps, nous revenons sur la nécessité de structurer les ressources pédagogiques. Ensuite, après avoir montré les faiblesses de LOM et SCORM dans la prise en compte d'une adaptation basée sur les styles d'apprentissage, nous présentons l'amélioration de ces normes.

4.4.1. Nécessité de structurer les ressources

Dans [Simard, 2002], l'auteur précise que la problématique d'utiliser une norme pour structurer des ressources est apparue sous l'effet conjugué de deux nécessités : d'une part améliorer l'efficacité économique des investissements et d'autre part améliorer l'efficacité pédagogique des produits proposés. Dans [Labat et al., 2006], l'auteur ajoute que ces enjeux peuvent se traduire par cinq critères à atteindre :

- Accessibilité : permettre la recherche, l'identification, l'accès et la mise à disposition de contenus et composantes de formation de façon distribuée ;
- Interopérabilité : permettre l'utilisation de contenus et composants développés par une organisation sur une plateforme donnée par d'autres organisations sur d'autres plateformes ;
- Réutilisabilité : permettre la réutilisation des contenus et composantes à différentes fins, dans différentes applications, dans différents produits, dans différents contextes et via différents modes d'accès ;
- Durabilité : permettre aux contenus et composants d'affronter les changements technologiques sans la nécessité d'une réingénierie ou d'un redéveloppement ;
- Adaptabilité : permettre la modulation sur mesure des contenus et des composantes.

4.4.2. Modèle SCORM d'agrégation de contenus

Le modèle SCORM d'agrégation de contenus (CAM⁴³) répond à plusieurs questions :

- Quels types de composants utilisés dans la conception de contenus pédagogiques ?
- Comment packager ces composants pour les rendre interopérables ?
- Comment les indexer pour offrir une recherche aisée ?
- Comment autoriser de multiples séquences et navigations dans les contenus ?

Ainsi, il introduit les concepts de composants, package, indexation, navigation et séquence.

⁴³ CAM : Content Aggregation Model

Les composants

SCORM utilise deux composants pour décrire les ressources physiques : les assets et les objets de contenus partageables SCO⁴⁴

- L'asset représente la forme la plus élémentaire de ressources d'apprentissage. Un asset est une représentation numérique de média tels que texte, images ou son qu'un client Web peut présenter à l'élève (Figure 4-10). Plusieurs assets peuvent être associés pour construire un autre asset. Par exemple, une page Web peut contenir des images GIF et du son WAV.

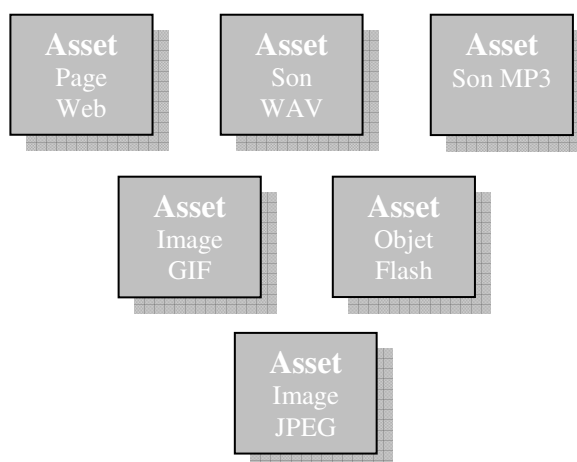


Figure 4-10 : exemples d'assets

- L'objet de contenu partageable (SCO) est une collection d'un ou plusieurs assets (Figure 4-11) représentant le plus bas niveau de granularité de ressource pédagogique utilisable par un LMS⁴⁵ pour suivre les activités d'un apprenant. Un SCO communique avec un LMS en utilisant l'API IEEE ECMA⁴⁶. Pour améliorer sa réutilisation, un SCO doit être indépendant du contexte d'apprentissage dans lequel il est utilisé. Cela permettra son exploitation dans différentes séquences pédagogiques.

⁴⁴ SCO : Sharable Content Object

⁴⁵ LMS : Learning Management System

⁴⁶ Institute for Electrical and Electronics Engineers ECMA Script Application Programming Interface

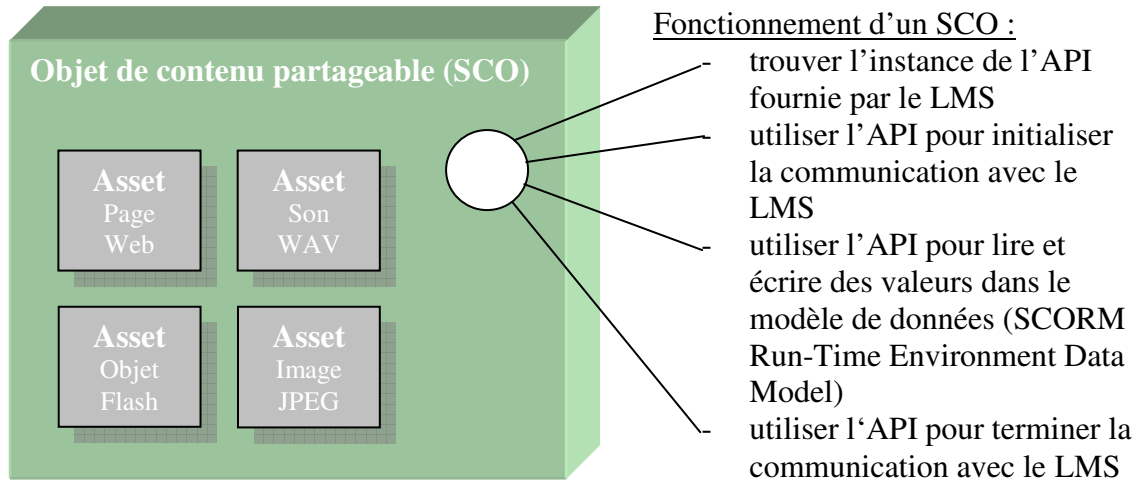


Figure 4-11 : exemple de SCO

Le package

Les contenus sont structurés dans un package sous la forme d'organisations. Une organisation représente une unité d'enseignement dont la structure décrit les activités. Chaque activité est reliée aux ressources qu'elle utilise (Figure 4-12).

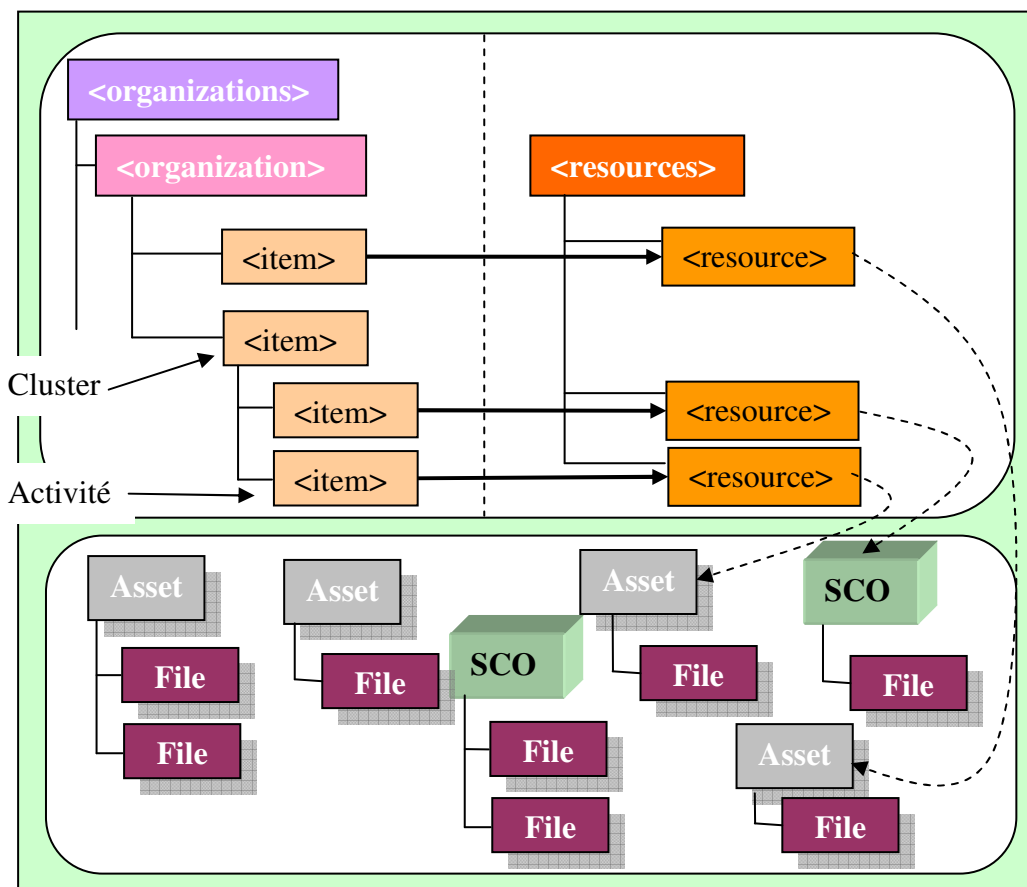


Figure 4-12 : Organisation des contenus

Une activité peut être décomposée en sous-activités (cluster). On peut associer à chaque niveau de décomposition une taxinomie particulière (ex : cours, chapitre, module,...). Une ressource est associée aux activités qui ne sont pas décomposées en sous-activités (feuilles). Ces ressources peuvent être des SCO ou des assets.

L'indexation

SCORM utilise des métadonnées LOM pour décrire l'ensemble des éléments manipulés (asset, SCO, activité, cluster, organisation, ressource, file) (Figure 4-13). L'ensemble des balises proposées par LOM permet d'identifier, de catégoriser et de rechercher chaque élément afin de faciliter une réutilisation et un partage des ressources pédagogiques.

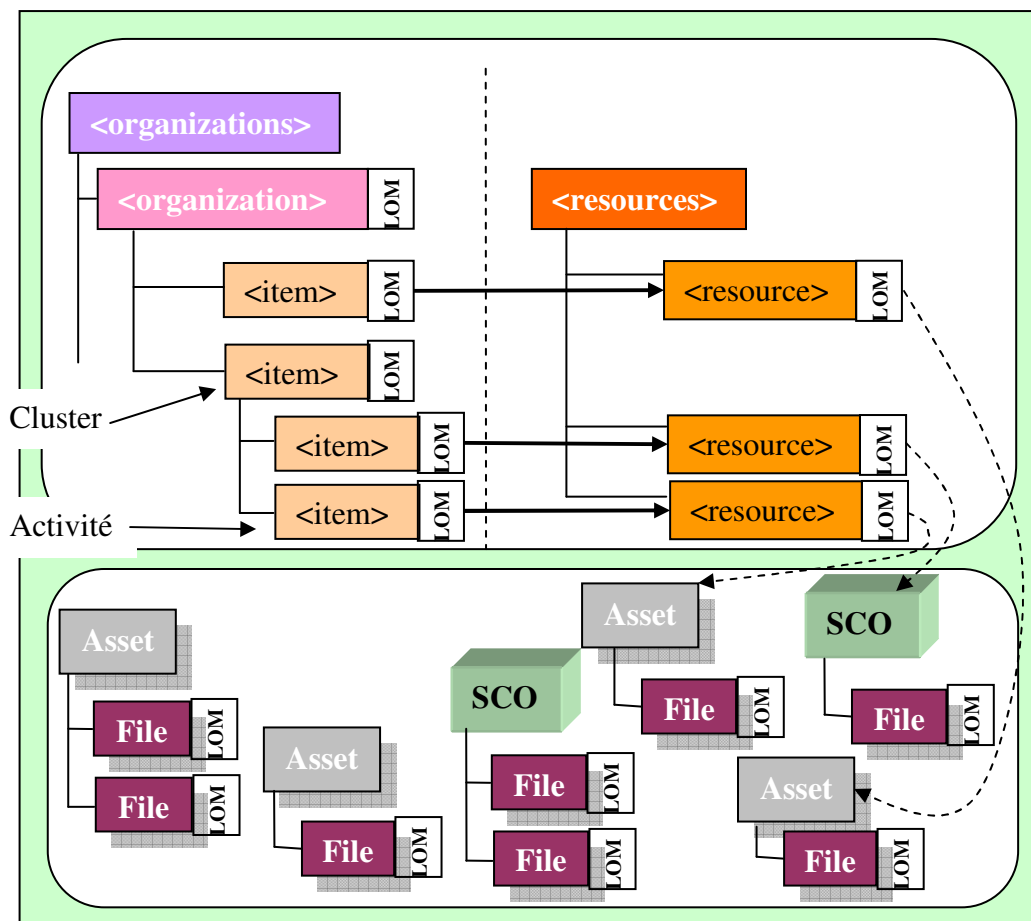


Figure 4-13 : métadonnées LOM

Les métadonnées de LOM sont décomposées en 9 catégories (Figure 4-14) :

1. *general* : permet des informations générales décrivant les composants. Il est possible par exemple de donner un titre et un identifiant unique à un composant.
2. *lifecycle* : comporte les caractéristiques du composant liées à son historique, son état courant et les auteurs ayant participé à sa création et son évolution.
3. *metaMetadata* : permet d'ajouter une description au bloc de métadonnées lui-même.
4. *technical* : utilisée pour identifier les exigences techniques et les caractéristiques intrinsèques du composant. Il est possible, par exemple, de définir le type mime de la ressource (text/html, image/gif, audio/wav,...),
5. *educational* : utilisée pour présenter les caractéristiques pédagogiques associées au composant. Il est possible, par exemple, d'indiquer le type d'élève pour lequel le composant est créé, le niveau de difficulté.
6. *rights* : utilisée pour décrire les droits de propriété intellectuelle sur le composant,
7. *relation* : utilisée pour préciser le type de relations qui lie ce composant à d'autres dans le paquet SCORM (ispartof, isrequiredby,...),
8. *annotation* : utilisée pour fournir un commentaire sur l'utilisation pédagogique du composant et définir la date et l'auteur de ce dernier,
9. *classification* : utilisée pour indiquer si le composant fait partie d'une classification particulière définissant par exemple un vocabulaire restreint pour le renseignement de certains champs.

```
<lom>
  <general />
  <lifecycle />
  <metaMetadata />
  <technical />
  <educational />
  <rights />
  <relation />
  <annotation />
  <classification />
</lom>
```

Figure 4-14 : les 9 catégories de LOM

Les séquences et la navigation

Les informations de séquences et de navigation sont associées aux activités et aux organisations (Figure 4-15). Ceci garantit un niveau de réutilisation élevé pour les composants (asset et SCO). En effet, les conditions de navigation étant contextuelles, les associer aux ressources aurait comme conséquence de ne pouvoir les utiliser dans aucun autre contexte. Les informations ajoutées dans l'organisation des activités consistent à définir certains comportements en fonction de situations données. Il incombe au LMS de jouer la séquence en respectant ces informations.

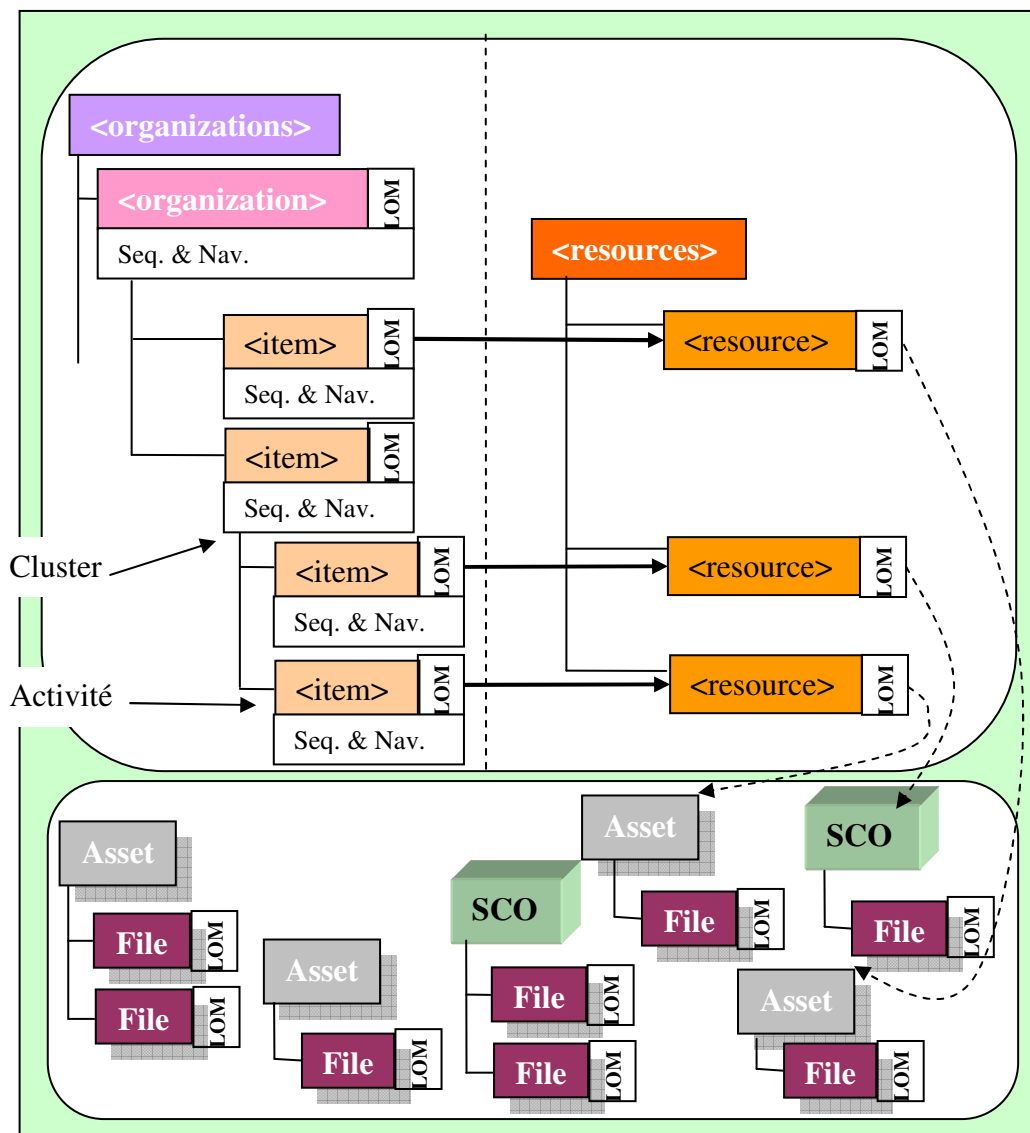


Figure 4-15 : séquence et navigation (Seq. & Nav.) pour les activités

Le modèle de contenu proposé par SCORM ne préconise pas vraiment de structure particulière pour maintenir les informations nécessaires à un LMS pour jouer des séquences d'enseignement *adaptées aux styles d'apprentissage* des apprenants. En utilisant les moyens d'extensions des métadonnées offerts par SCORM, nous proposons une amélioration de SCORM CAM afin de pouvoir implémenter les informations de notre modèle conceptuel de contenus adaptatifs.

4.4.3. Structuration proposée

Afin d'atteindre l'objectif d'indexation des contenus par l'ontologie du domaine et l'ontologie pédagogique, nous avons décidé d'adopter SCORM CAM comme langage. Ainsi, nous avons défini un manifest pour organiser et décrire les unités d'enseignement produites par notre méthode de production d'objets d'apprentissage. Rappelons que notre modèle conceptuel décrit les tâches métiers, les objectifs globaux et les objectifs finaux, ainsi que les diverses séquences pédagogiques composées de différentes activités et ressources physiques associées.

Le principe fondamental de notre structure consiste à dissocier la décomposition des unités d'enseignement (Tâche, Objectif global, Objectif Final) de l'ensemble des séquences, activités et ressources utilisées pour implémenter un objectif final. Cette structure offre comme avantage de fournir un cadre de travail qui peut être utilisé aisément par des composants logiciels pour faire de l'adaptation à plusieurs niveaux.

Dans notre système, l'adaptation (cf. chapitre 5) se fait sur le niveau de connaissance et sur le style d'apprentissage de l'élève (Figure 4-16).

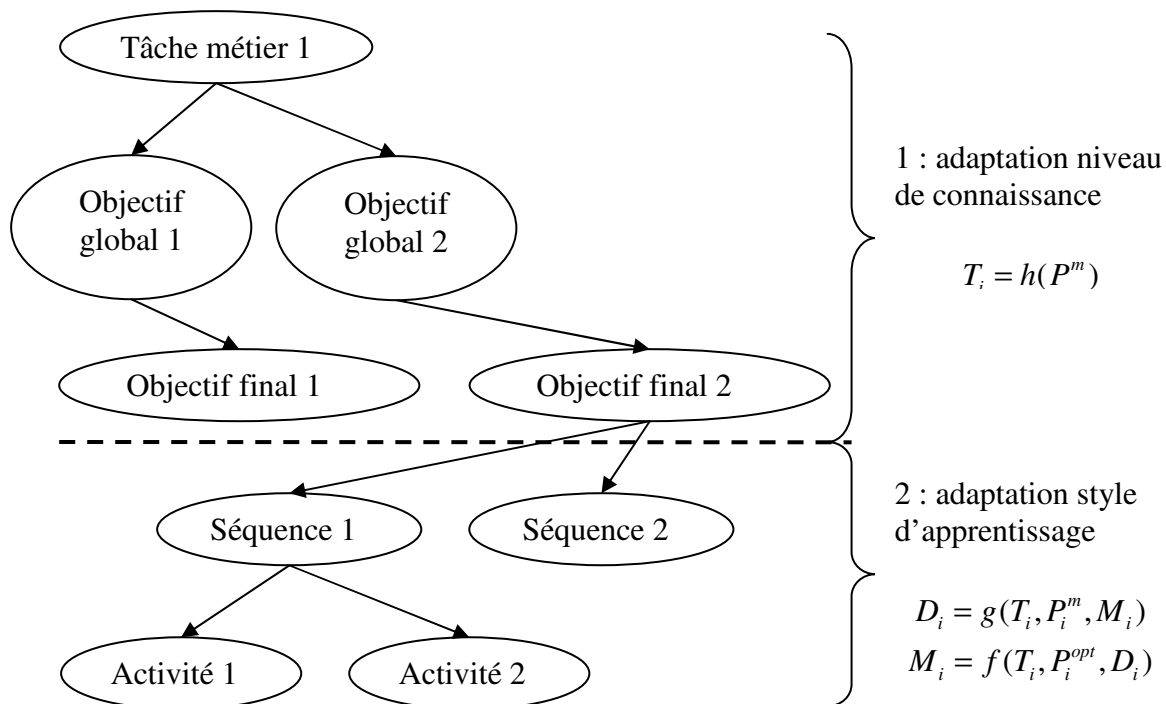


Figure 4-16 : Deux niveaux d'adaptation

Un premier composant logiciel est chargé d'adapter le contenu au niveau de connaissances de l'apprenant. Il utilise un modèle de connaissances de l'élève de type Overlay sur la première partie de la structure : les tâches et les objectifs. Pour se déplacer dans la décomposition des unités d'enseignements, il applique une stratégie $T_i = h(P^m)$ pour choisir le prochain objectif à enseigner. P^m est le comportement observé (mesuré) de l'élève pendant la séquence pédagogique qui vient de se terminer et T_i est l'objectif pédagogique suivant qui sera proposé à l'élève.

Un second composant logiciel est chargé d'adapter la présentation de la séquence au style d'apprentissage de l'apprenant. Il utilise la seconde partie de la structure qui propose pour un objectif final l'ensemble des séquences disponibles dans le paquet. Pour choisir la séquence et l'activité à proposer à l'élève, ce composant fait une première estimation du style d'apprentissage de l'élève $D_i = g(T_i, P_i^m, M_i)$ en fonction de ce qui a été observé lors de la dernière séquence pédagogique. Ensuite sur cette base, il recommande la séquence pédagogique idéale M_i à proposer à l'élève :
$$M_i = f(T_i, P_i^{opt}, D_i).$$

Pour mettre en œuvre cette structure, nous avons utilisé le concept de sous-manifest proposé par SCORM (Figure 4-17). Un manifest principal contient la décomposition des unités d'enseignement jusqu'aux objectifs finaux. A chaque objectif final est alors associé un sous-manifest qui contient l'ensemble des séquences pédagogiques offertes par le paquet. Ainsi, nous restons compatible avec les LMS SCORM tout en enrichissant le contenu d'informations adaptatives.

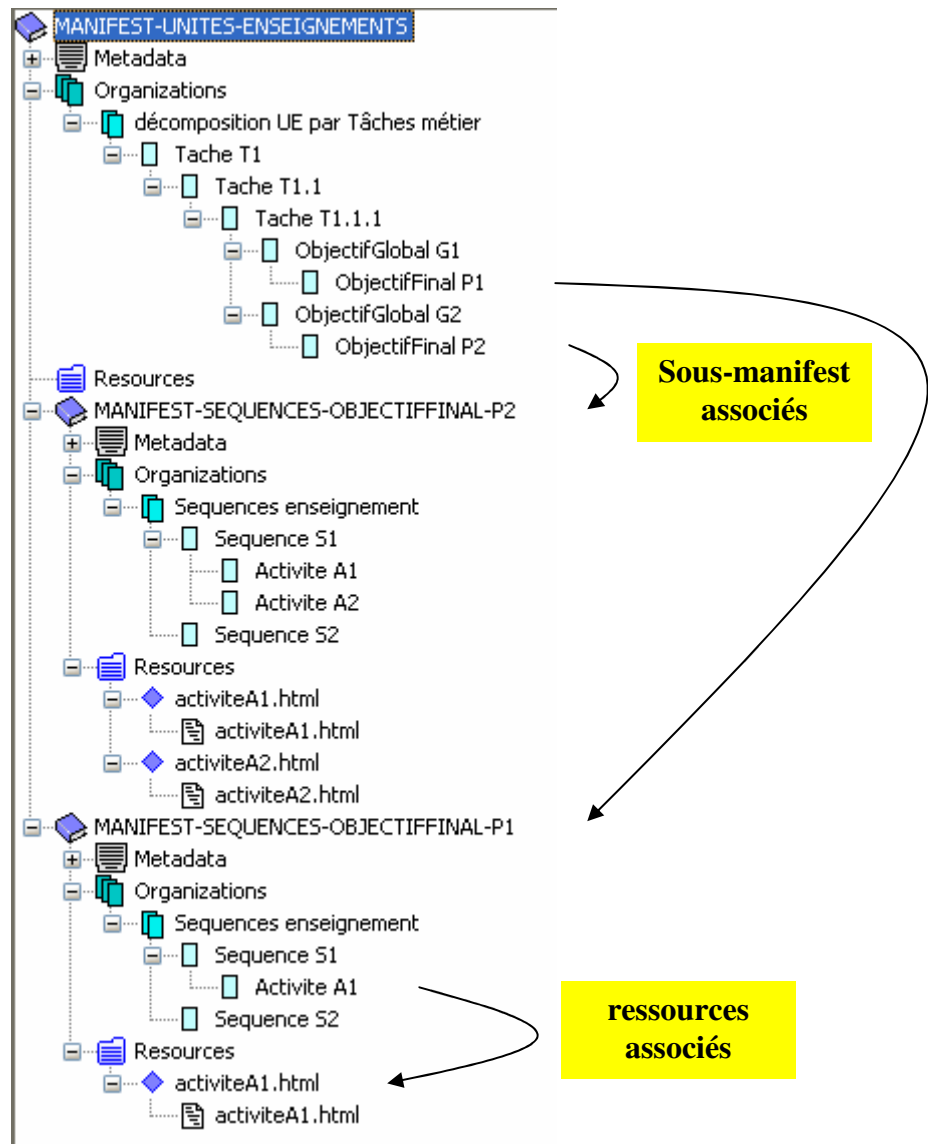


Figure 4-17 : Structure du paquet SCORM

4.4.4. Données d'adaptation proposées

Le modèle de contenus de SCORM propose d'utiliser les métadonnées de LOM pour décrire chaque niveau de composants contenus dans le paquet (Organization, Item, Resource, ...). Ces métadonnées couvrent mal les problématiques liées à l'adaptation des contenus au style d'apprentissage des apprenants. Nous proposons d'étendre les métadonnées LOM pour représenter les informations nécessaires à l'adaptation. En effet, chaque élément de notre manifest est associé à des métadonnées permettant de l'adapter au style d'apprentissage. Dans la suite, nous parcourons l'ensemble des concepts que nous avons défini dans notre modèle de contenus. Nous proposons des métadonnées LOM enrichies pour satisfaire nos besoins liés à l'adaptation.

Les Tâches

Dans la décomposition structurelle de notre modèle de contenus, nous avons considéré que le concept <Tâche> était caractérisé par un numéro, un titre, un acteur et un ensemble d'objectifs Globaux. De plus, il peut être décomposé en sous-Tâches (Figure 4-18).

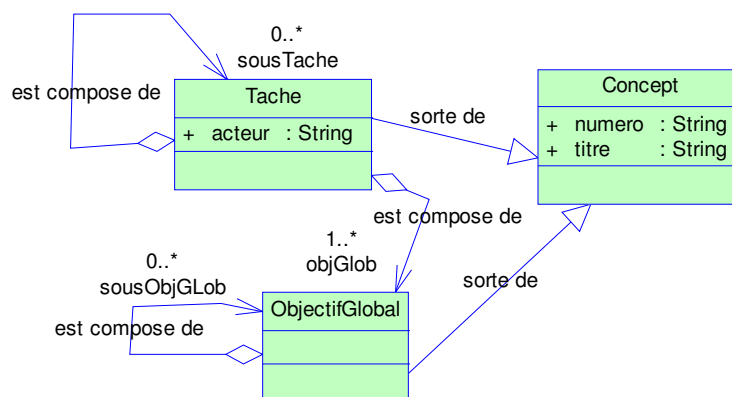


Figure 4-18 : le concept de <Tache>

Le modèle fait apparaître aussi une relation d'influence entre les concepts <Tâche> (Figure 4-19).

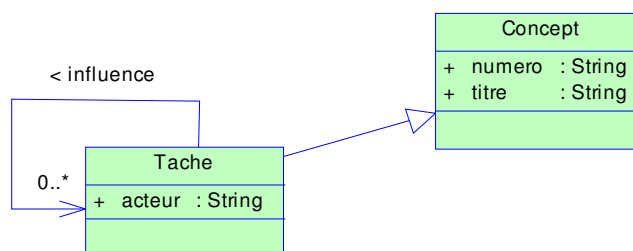


Figure 4-19 : relation d'influence entre concepts <Tâche>

Afin de décrire ses attributs, nous proposons d'utiliser les métadonnées décrites dans le Tableau 4-1 :

Eléments LOM	Eléments Ajoutés	Concepts de l'Ontologie
<imsmd:lom> <imsmd:general>		
<imsmd:identifiant />		numero
<imsmd:title />		titre
<imsmd:aggregationLevel>		type du concept
</imsmd:general>		
<imsmd:educational> </imsmd:educational>	<up2uml :typicalLearnerType/>	Acteur
<imsmd:relation />		est compose de / influence
</imsmd:lom>		

Tableau 4-1 : représentation du concept de <Tache>

L'élément **<imsmd:identifiant />** est utilisé pour stocker le numéro du concept. Il joue le rôle d'un container composé de **<catalog>** et **<entry>**. **<catalog>** prend la valeur DOI (Digital Object Identifier) et **<entry>** contient un numéro unique qui identifie le concept.

L'élément **<imsmd:title />** contiendra le titre du concept (chaîne de caractères).

L'élément **<imsmd:aggregationLevel>** contiendra le type du concept. Il est considéré comme un container. Il se compose d'un premier élément **<source>** définissant la source du vocabulaire utilisé dans le second élément **<value>**. Dans notre cas, la source sera « *up2uml* » et les valeurs autorisées : *Tache*, *ObjectifGlobal*, *ObjectifFinal*, *Sequence*, *Activite* et *Ressource*.

Nous proposons l'élément `<up2uml :typicalLearnerType/>` pour définir le type d'apprenant pour lequel ce concept est utile (chaîne de caractères).

L'élément `<imsmd:relation />` est utilisé pour stocker les relations qui existent entre ce concept et d'autres concepts. Il est considéré comme un container composé d'un premier élément `<kind>` qui sera utilisé pour définir le type de relation dont le vocabulaire est fixé par LOM. Nous utilisons le terme « *haspart* » pour la relation « *est composé de* » et « *isrequiredby* » pour la relation « *influence* ». Un second élément `<resource>` contiendra un élément `<imsmd:identifiant />` pour définir le concept vers lequel la relation s'applique. L'élément `<imsmd:relation />` peut apparaître plusieurs fois. Il est alors possible de décrire pour un concept plusieurs relations.

En guise d'exemple, la Figure 4-21 montre la description de la tâche T0550 de la Figure 4-20.

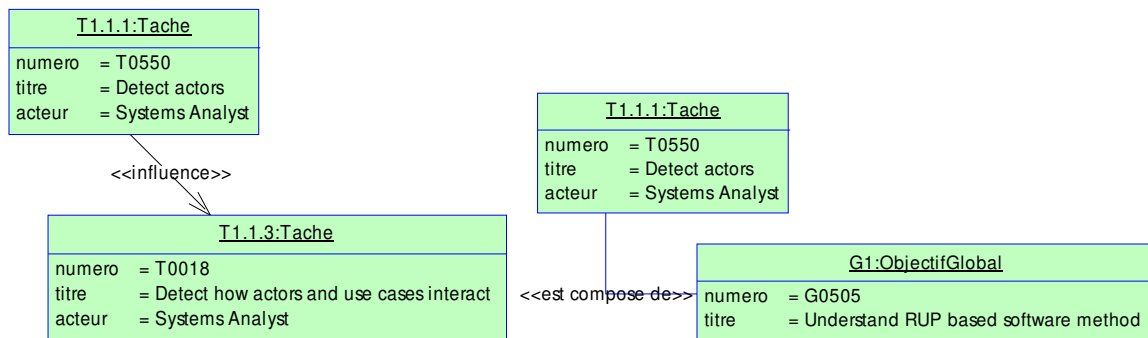


Figure 4-20 : exemple de concept <Tâche>

```

<imsmd:lom>
  <imsmd:general>
    <imsmd:identifiant>
      <imsmd:catalog>DOI</imsmd:catalog>
      <imsmd:entry>T0550</imsmd:entry>
    </imsmd:identifiant>
    <imsmd:title>
      <imsmd:string>Detect actors</imsmd:string>
    </imsmd:title>
    <imsmd:aggregationLevel>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>Tache</imsmd:value>
    </imsmd:aggregationLevel>
  </imsmd:general>
  <imsmd:educational>
    <up2uml:typicalLearnerType>
      Systems Analyst
    </up2uml:typicalLearnerType>
  </imsmd:educational>
  <imsmd:relation>
    <imsmd:kind>
      <imsmd:source>LOMv1.0</imsmd:source>
      <imsmd:value>haspart</imsmd:value>
    </imsmd:kind>
    <imsmd:resource>
      <imsmd:identifiant>
        <imsmd:catalog>DOI</imsmd:catalog>
        <imsmd:entry>G0505</imsmd:entry>
      </imsmd:identifiant>
    </imsmd:resource>
  </imsmd:relation>
  <imsmd:relation>
    <imsmd:kind>
      <imsmd:source>LOMv1.0</imsmd:source>
      <imsmd:value>isrequiredby</imsmd:value>
    </imsmd:kind>
    <imsmd:resource>
      <imsmd:identifiant>
        <imsmd:catalog>DOI</imsmd:catalog>
        <imsmd:entry>T0018</imsmd:entry>
      </imsmd:identifiant>
    </imsmd:resource>
  </imsmd:relation>
</imsmd:lom>

```

Figure 4-21 : exemple de description LOM d'un concept <Tâche>

L'objectif global

Dans la décomposition structurelle de notre modèle de contenus, nous avons considéré que le concept **<ObjectifGlobal>** était caractérisé par un numéro, un titre et un ensemble d'objectifs Finaux. De plus, il peut être décomposé en sous-Objectifs globaux (Figure 4-22).

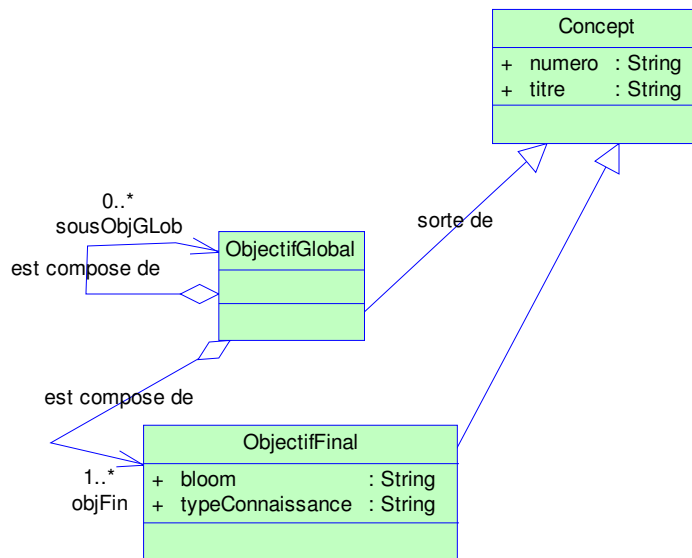


Figure 4-22 : le concept **<ObjectifGlobal>**

Le modèle fait apparaître aussi une relation d'influence entre les concepts **<ObjectifGlobal>** (Figure 4-23).

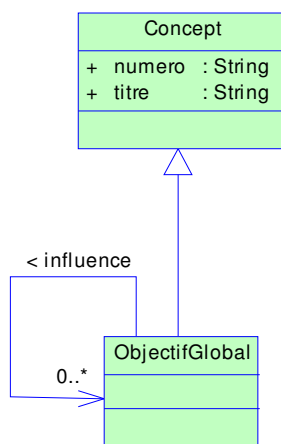


Figure 4-23 : relation d'influence entre concepts **<ObjectifGlobal>**

Nous proposons d'utiliser les éléments suivants pour décrire le concept **<ObjectifGlobal>** (Tableau 4-2)

Eléments LOM	Eléments Ajoutés	Concepts de l'Ontologie
<imsmd:lom> <imsmd:general>		
<imsmd:identifiant />		numero
<imsmd:title />		titre
<imsmd:aggregationLevel>		type du concept
</imsmd:general>		
<imsmd:relation />		est compose de / influence
</imsmd:lom>		

Tableau 4-2 : représentation du concept <ObjectifGlobal>

L'élément **<imsmd:identifiant />** est utilisé pour stocker le numéro du concept. Il joue le rôle d'un container composé de **<catalog>** et **<entry>**. **<catalog>** prend la valeur DOI (Digital Object Identifier) et **<entry>** contient un numéro unique qui identifie le concept.

L'élément **<imsmd:title />** contiendra le titre du concept (chaîne de caractères).

L'élément **<imsmd:aggregationLevel>** contiendra le type du concept. Il est considéré comme un container. Il se compose d'un premier élément **<source>** définissant la source du vocabulaire utilisé dans le second élément **<value>**. Dans notre cas, la source sera « *up2uml* » et les valeurs autorisées : *Tache*, *ObjectifGlobal*, *ObjectifFinal*, *Sequence*, *Activite* et *Ressource*.

L'élément **<imsmd:relation />** est utilisé pour stocker les relations qui existent entre ce concept et d'autres concepts. Il est considéré comme un container composé d'un premier élément **<kind>** qui sera utilisé pour définir le type de relation dont le vocabulaire est fixé par LOM. Nous utilisons le terme « *haspart* » pour la relation « *est composé de* » et « *isrequiredby* » pour la relation « *influence* ». Un second élément **<resource>** contiendra un élément **<imsmd:identifiant />** pour définir le concept vers lequel la relation s'applique. L'élément **<imsmd:relation />** peut apparaître plusieurs fois. Il est alors possible de décrire pour un concept plusieurs relations.

La Figure 4-25 présente la description de l'objectif global G0124 de la Figure 4-24.

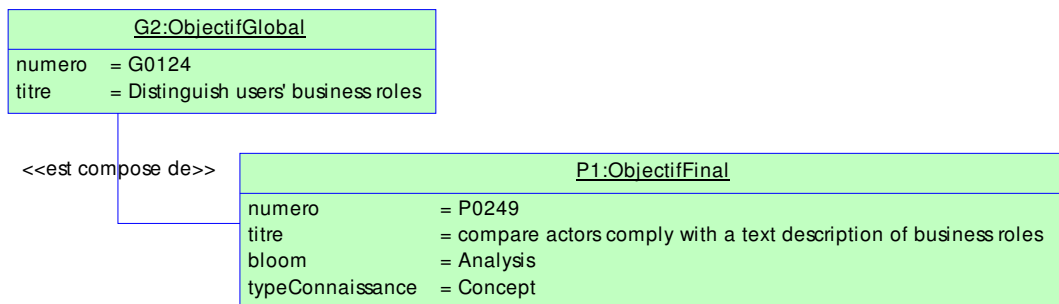


Figure 4-24 : exemple de concept <ObjectifGlobal>

```

<imsmd:lom>
  <imsmd:general>
    <imsmd:identifiant>
      <imsmd:catalog>DOI</imsmd:catalog>
      <imsmd:entry>G0124</imsmd:entry>
    </imsmd:identifiant>
    <imsmd:title>
      <imsmd:string>Distinguish user's business roles</imsmd:string>
    </imsmd:title>
    <imsmd:aggregationLevel>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>ObjectifGlobal</imsmd:value>
    </imsmd:aggregationLevel>
  </imsmd:general>
  <imsmd:relation>
    <imsmd:kind>
      <imsmd:source>LOMv1.0</imsmd:source>
      <imsmd:value>haspart</imsmd:value>
    </imsmd:kind>
    <imsmd:resource>
      <imsmd:identifiant>
        <imsmd:catalog>DOI</imsmd:catalog>
        <imsmd:entry>P0249</imsmd:entry>
      </imsmd:identifiant>
    </imsmd:resource>
  </imsmd:relation>
</imsmd:lom>

```

Figure 4-25 : exemple de description LOM de concept <ObjectifGlobal>

L'objectif final

Toujours dans la décomposition structurelle de notre modèle de contenus nous avons considéré que le concept **<ObjectifFinal>** était caractérisé par un numéro, un titre, un niveau de Bloom, un type de connaissances et un ensemble de séquences pédagogiques (Figure 4-26).

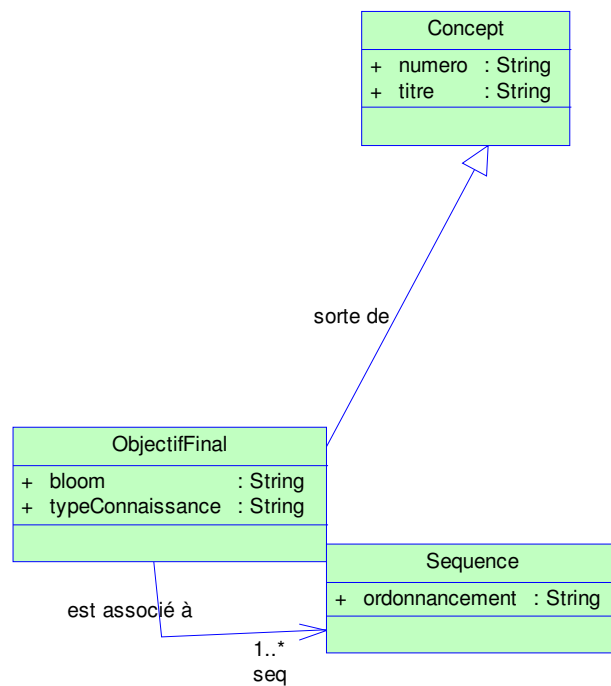


Figure 4-26 : le concept **<ObjectifFinal>**

Le modèle fait apparaître aussi une relation d'influence entre les concepts **<ObjectifFinal>** (Figure 4-27).

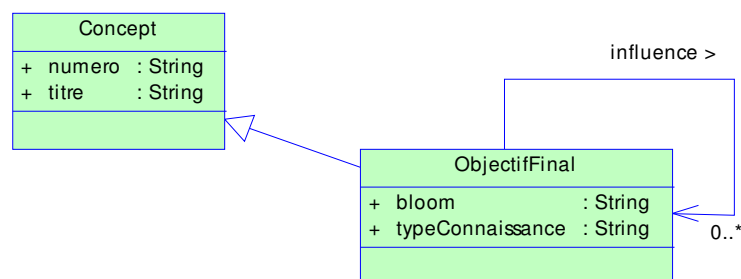


Figure 4-27 : relation d'influence entre concepts **<ObjectifFinal>**

Nous proposons d'utiliser les éléments suivants pour décrire le concept **<ObjectifFinal>** (Tableau 4-3). Il faut considérer que l'ensemble des séquences qui sont associées à un objectif final sont stockées dans un sous-manifest qui sera relié à l'élément **<ObjectifFinal>**. De ce fait les seules relations qu'il sera nécessaires de faire apparaître pour ce type d'élément sont les relations de type « *influence* » entre concepts de type **<ObjectifFinal>**.

Éléments LOM	Éléments Ajoutés	Concepts de l'Ontologie
<imsmd:lom> <imsmd:general>		
<imsmd:identifiant />		numero
<imsmd:title />		titre
<imsmd:aggregationLevel>		type du concept
</imsmd:general>		
<imsmd:educational> </imsmd:educational>	<up2uml :bloom/> <up2uml :typeConnaissance/>	bloom typeConnaissance
<imsmd:relation />		influence
</imsmd:lom>		

Tableau 4-3 : représentation du concept <ObjectifFinal>

L'élément **<imsmd:identifiant />** est utilisé pour stocker le numéro du concept. Il joue le rôle d'un container composé de **<catalog>** et **<entry>**. **<catalog>** prend la valeur DOI (Digital Object Identifier) et **<entry>** contient un numéro unique qui identifie le concept.

L'élément **<imsmd:title />** contiendra le titre du concept (chaîne de caractères).

L'élément **<imsmd:aggregationLevel>** contiendra le type du concept. Il est considéré comme un container. Il se compose d'un premier élément **<source>** définissant la source du vocabulaire utilisé dans le second élément **<value>**. Dans notre cas, la source sera « *up2uml* » et les valeurs autorisées : *Tache*, *ObjectifGlobal*, *ObjectifFinal*, *Sequence*, *Activite* et *Ressource*.

L'élément **<up2uml :bloom/>** est un élément ajouté au modèle LOM. Il contiendra une chaîne de caractère précisant le niveau de bloom que traite le concept concerné. Les valeurs autorisées sont celles définies par la taxinomie de Bloom (*Knowledge*, *Comprehension*, ..., *Evaluation*).

L'élément `<up2uml :typeConnaissance/>` est un élément ajouté au modèle LOM. Il contiendra une chaîne de caractère précisant le niveau de bloom que traite le concept concerné. Les valeurs autorisées sont celles définies par la classification de Gagné (concept, fait, processus, principe, procédé).

L'élément `<imsmd:relation />` est utilisé pour stocker les relations qui existent entre ce concept et d'autres concepts. Il est considéré comme un container composé d'un premier élément `<kind>` qui sera utilisé pour définir le type de relation dont le vocabulaire est fixé par LOM. Nous utilisons le terme « *haspart* » pour la relation « *est composé de* » et « *isrequiredby* » pour la relation « *influence* ». Un second élément `<resource>` contiendra un élément `<imsmd:identifiant />` pour définir le concept vers lequel la relation s'applique. L'élément `<imsmd:relation />` peut apparaître plusieurs fois. Il est alors possible de décrire pour un concept plusieurs relations.

La description de l'objectif final numéro P0249 de la Figure 4-28 est donnée en Figure 4-29 et Figure 4-30.

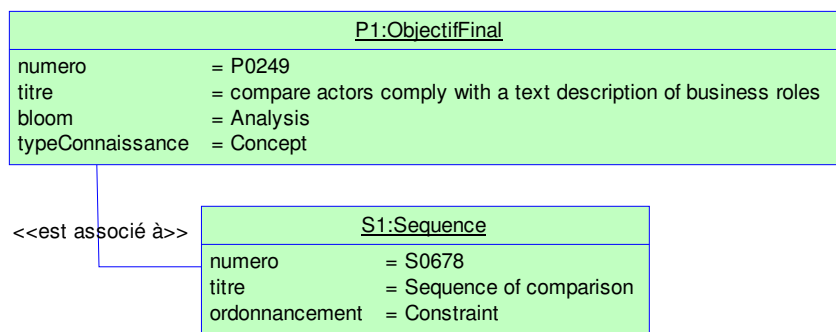


Figure 4-28 : exemple de concept <ObjectifFinal>

```
<imsmd:lom>
  <imsmd:general>
    <imsmd:identifier>
      <imsmd:catalog>DOI</imsmd:catalog>
      <imsmd:entry>P0249</imsmd:entry>
    </imsmd:identifier>
    <imsmd:title>
      <imsmd:string>
        compare actors comply with a text description of business roles
      </imsmd:string>
    </imsmd:title>
    <imsmd:aggregationLevel>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>ObjectifFinal</imsmd:value>
    </imsmd:aggregationLevel>
  </imsmd:general>
  <imsmd:educational>
    <up2uml:bloom>
      Analysis
    </up2uml:bloom>
    <up2uml:typeConnaissance>
      Concept
    </up2uml:typeConnaissance>
  </imsmd:educational>
</imsmd:lom>
```

Figure 4-29 : exemple de description LOM de concept <ObjectifFinal>

```

<manifest identifier="MANIFEST-UNITES-ENSEIGNEMENTS" >
...
  <item identifier="P1" identifierref="MANIFEST-SEQUENCES-
OBJECTIFFINAL-P1">
    <title>ObjectifFinal P1</title>
  </item>
...
  <manifest identifier="MANIFEST-SEQUENCES-OBJECTIFFINAL-P1"
xml:base="submanifest1/">
...
    <item identifier="S1 P1">
      <title>Sequence S1</title>
      ...
    </item>
...
  </manifest>
</manifest>

```

Figure 4-30 : exemple de description SCORM du concept <Sequence> pour un objectif final

La séquence

Notre modèle de contenus considère que le concept <Sequence> était caractérisé par un numéro, un titre, un type d'ordonnement et un ensemble d'activités (Figure 4-31).

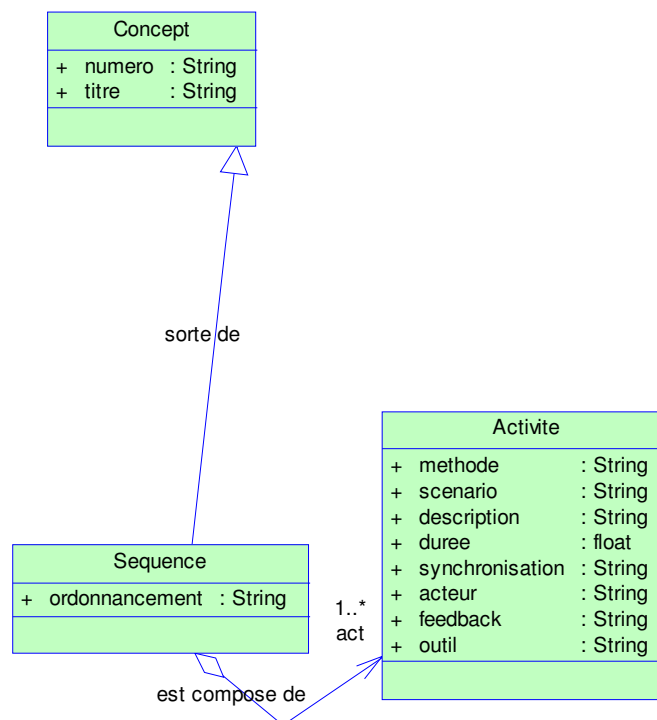


Figure 4-31 : le concept <Sequence>

Nous proposons d'utiliser les éléments suivants pour décrire ce concept (Tableau 4-4).

Eléments LOM	Eléments Ajoutés	Concepts de l'Ontologie
<imsmd:lom> <imsmd:general>		
<imsmd:identifiant />		numero
<imsmd:title />		titre
<imsmd:aggregationLevel>		type du concept
</imsmd:general>		
<imsmd:relation />		Est composé de
</imsmd:lom>		

Tableau 4-4 : représentation du concept <Sequence>

L'élément **<imsmd:identifiant />** est utilisé pour stocker le numéro du concept. Il joue le rôle d'un container composé de **<catalog>** et **<entry>**. **<catalog>** prend la valeur DOI (Digital Object Identifier) et **<entry>** contient un numéro unique qui identifie le concept.

L'élément **<imsmd:title />** contiendra le titre du concept (chaîne de caractères).

L'élément **<imsmd:aggregationLevel>** contiendra le type du concept. Il est considéré comme un container. Il se compose d'un premier élément **<source>** définissant la source du vocabulaire utilisé dans le second élément **<value>**. Dans notre cas, la source sera « *up2uml* » et les valeurs autorisées : *Tache*, *ObjectifGlobal*, *ObjectifFinal*, *Sequence*, *Activite* et *Ressource*.

L'élément **<imsmd:relation />** est utilisé pour stocker les relations qui existent entre ce concept et d'autres concepts. Il est considéré comme un container composé d'un premier élément **<kind>** qui sera utilisé pour définir le type de relation dont le vocabulaire est fixé par LOM. Nous utilisons le terme « *haspart* » pour la relation « *est composé de* » et « *isrequiredby* » pour la relation « *influence* ». Un second élément **<resource>** contiendra un élément **<imsmd:identifiant />** pour définir le concept vers lequel la relation s'applique. L'élément **<imsmd:relation />** peut apparaître plusieurs fois. Il est alors possible de décrire pour un concept plusieurs relations.

La description de la séquence numéro S0678 de la Figure 4-32 est donnée en Figure 4-33 et Figure 4-34.

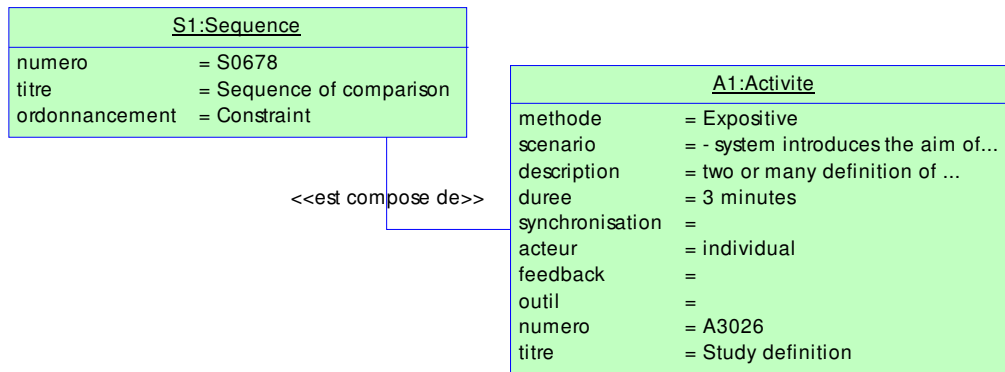


Figure 4-32 : exemple de concept <Sequence>

```

<imsmd:lom>
  <imsmd:general>
    <imsmd:identifieur>
      <imsmd:catalog>DOI</imsmd:catalog>
      <imsmd:entry>S0678</imsmd:entry>
    </imsmd:identifieur>
    <imsmd:title>
      <imsmd:string>sequence of comparison</imsmd:string>
    </imsmd:title>
    <imsmd:aggregationLevel>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>Sequence</imsmd:value>
    </imsmd:aggregationLevel>
  </imsmd:general>
</imsmd:lom>

```

Figure 4-33 : exemple de description LOM de concept <Sequence>

Eléments SCORM	Eléments Ajoutés	Concepts de l'Ontologie
<pre> <imsss:sequencing> <imsss:controlMode choice="false" flow="true" /> </imsss:sequencing> </pre>		ordonnancement

Tableau 4-5 : représentation SCORM du concept <Sequence>

L'élément `<imsss:controlMode />` du modèle de contenus de SCORM est utilisé pour décrire le type d'ordonnancement des activités dans la séquence. Nous utilisons une combinaison des valeurs pour les attributs « **choice** » et « **flow** » pour coder les valeurs liées à notre ontologie, à savoir « *contraint* » et « *libre* ». Si l'attribut « **choice** » est à **true** alors le LMS doit offrir une navigation libre à l'apprenant. Si l'attribut « **flow** » est à **true** alors le LMS doit offrir à l'apprenant une navigation séquentielle.

```
<manifest identifier="MANIFEST-UNITES-ENSEIGNEMENTS" >
...
  <item identifier="P1" identifierref="MANIFEST-SEQUENCES-
OBJECTIFFINAL-P1">
  <title>ObjectifFinal P1</title>
  </item>
...
  <manifest identifier="MANIFEST-SEQUENCES-
OBJECTIFFINAL-P1" xml:base="submanifest1/">
  ...
    <item identifier="S1 P1">
      <title>Sequence S1</title>
      ...
      <imsss:sequencing>
        <imsss:controlMode choice="false" flow="true" />
      </imsss:sequencing>
    </item>
  ...
  </manifest>
</manifest>
```

Figure 4-34 : exemple de description SCORM du concept <Sequence> pour un objectif final

L'activité

Le concept <Activite> est caractérisé, dans notre modèle de contenus, par un numéro, un titre, une méthode, un scénario, une description, une durée, un type de synchronisation, un type d'acteur, un type de feedback, des outils et un ensemble de ressources (Figure 4-35).

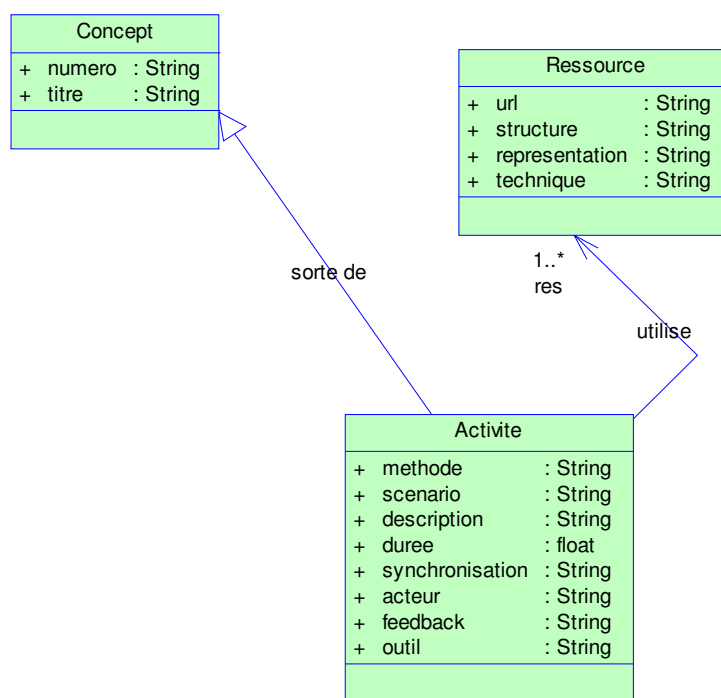


Figure 4-35 : le concept <Activite>

Nous proposons d'utiliser les éléments suivants pour décrire ce concept (Tableau 4-6).

Eléments LOM	Eléments Ajoutés	Concepts de l'Ontologie
<imsmd:lom> <imsmd:general>		
<imsmd:identifiant />		numero
<imsmd:title />		titre
<imsmd:aggregationLevel>		type du concept
</imsmd:general>		
<imsmd:educational>		
<interactivityType/>		methode

<description/>		description
<typicalLearningTime/>		duree
<intendedEndUserRole/>		acteur
	<up2uml :scenario/> <up2uml :synchronisation/> <up2uml :feedback/>	scenario synchronisation feedback
</imsmd:educational>		
<imsmd:technical> <requirement /> <otherPlateformRequirements/> </imsmd:technical>		outil outil
<imsmd:relation />		Est composé de
</imsmd:lom>		

Tableau 4-6 : représentation du concept <Activite>

L'élément <imsmd:identifiant /> est utilisé pour stocker le numéro du concept. Il joue le rôle d'un container composé de <catalog> et <entry>. <catalog> prend la valeur DOI (Digital Object Identifier) et <entry> contient un numéro unique qui identifie le concept.

L'élément <imsmd:title /> contiendra le titre du concept (chaîne de caractères).

L'élément <imsmd:aggregationLevel> contiendra le type du concept. Il est considéré comme un container. Il se compose d'un premier élément <source> définissant la source du vocabulaire utilisé dans le second élément <value>. Dans notre cas, la source sera « up2uml » et les valeurs autorisées : *Tache*, *ObjectifGlobal*, *ObjectifFinal*, *Sequence*, *Activite* et *Ressource*.

L'élément <interactivityType/> contiendra le type de la méthode pédagogique utilisée dans le composant. Les valeurs autorisées sont expositive, interrogative et active.

L'élément <description/> contiendra une description générale de l'activité.

L'élément <typicalLearningTime/> contiendra la durée estimée pour réaliser l'activité.

L'élément <intendedEndUserRole/> est utilisé pour décrire le type d'acteur réalisant l'activité. On pourra exprimer une activité pour un élève, un groupe d'élève ou bien la classe entière.

L'élément <up2uml :scenario/> est un élément ajouté. Il est utilisé pour décrire l'interaction entre les acteurs pendant l'activité.

L'élément <up2uml :synchronisation/> est un élément ajouté. Il est utilisé pour définir si l'activité est synchrone ou asynchrone.

L'élément <up2uml :feedback/> est un élément ajouté. Il définit qui doit faire un feedback sur l'activité.

Les éléments **<requirement />** et **<otherPlatformRequirements/>** sont utilisés pour définir les outils spécifiques requis pour l'activité.

L'élément **<imsmd:relation />** est utilisé pour stocker les relations qui existent entre ce concept et d'autres concepts. Il est considéré comme un container composé d'un premier élément **<kind>** qui sera utilisé pour définir le type de relation dont le vocabulaire est fixé par LOM. Nous utilisons le terme « *haspart* » pour la relation « *est composé de* » et « *isrequiredby* » pour la relation « *influence* ». Un second élément **<resource>** contiendra un élément **<imsmd:identifiant />** pour définir le concept vers lequel la relation s'applique. L'élément **<imsmd:relation />** peut apparaître plusieurs fois. Il est alors possible de décrire pour un concept plusieurs relations.

La description de l'activité numéro A3026 de la Figure 4-36 est donnée en Figure 4-37.

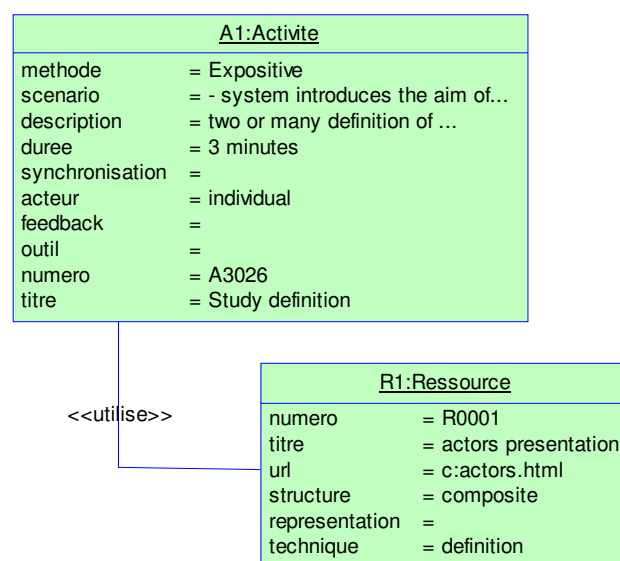


Figure 4-36 : exemple de concept <Activite>

```

<imsmd:lom>
  <imsmd:general>
    <imsmd:identifier>
      <imsmd:catalog>DOI</imsmd:catalog>
      <imsmd:entry>A3026</imsmd:entry>
    </imsmd:identifier>
    <imsmd:title>
      <imsmd:string>study definition</imsmd:string>
    </imsmd:title>
    <imsmd:aggregationLevel>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>Activite</imsmd:value>
    </imsmd:aggregationLevel>
  </imsmd:general>
  <imsmd:technical>
    <imsmd:requirement /><imsmd:otherPlatformRequirements />
  </imsmd:technical>
  <imsmd:educational>
    <imsmd:interactivityType>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>Expositive</imsmd:value>
    </imsmd:interactivityType>
    <imsmd:intendedEndUserRole>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>Individual</imsmd:value>
    </imsmd:intendedEndUserRole>
    <imsmd:typicalLearningTime>
      <imsmd:duration>PT3M</imsmd:duration>
    </imsmd:typicalLearningTime>
    <imsmd:description>
      <imsmd:string>two or many definition of...</imsmd:string>
    </imsmd:description>
    <up2uml :scenario>- system introduce ...</up2uml :scenario>
    <up2uml :synchronisation/>
    <up2uml :feedback/>
  </imsmd:educational>
  <imsmd:relation>
    <imsmd:kind>
      <imsmd:source>LOMv1.0</imsmd:source>
      <imsmd:value>haspart</imsmd:value>
    </imsmd:kind>
    <imsmd:resource>
      <imsmd:identifier>
        <imsmd:catalog>DOI</imsmd:catalog>
        <imsmd:entry>R0001</imsmd:entry>
      </imsmd:identifier>
    </imsmd:resource>
  </imsmd:relation>
</imsmd:lom>

```

Figure 4-37 : exemple de description LOM de concept <Activite>

La ressource

Finalement, le concept **<Resource>** est caractérisé, dans notre modèle de contenus, par un numéro, un titre, un URL, un type de structure, un type de représentation et un type de technique d'apprentissage (Figure 4-38).

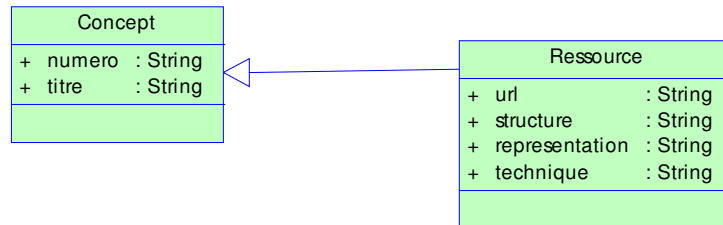


Figure 4-38 : le concept de **<Ressource>**

Nous proposons d'utiliser les éléments suivants pour décrire ce concept (Tableau 4-7).

Eléments LOM	Eléments Ajoutés	Concepts de l'Ontologie
<imsmd:lom> <imsmd:general>		
<imsmd:identifiant />		numero
<imsmd:title />		titre
<imsmd:aggregationLevel>		type du concept
<imsmd:structure/>		structure
</imsmd:general>		
<imsmd:educational> <imsmd:learningResourceType/> </imsmd:educational>		technique
<imsmd:technical> <location /> <format/> </imsmd:technical>		url représentation
</imsmd:lom>		

Tableau 4-7 : représentation du concept **<Ressource>**

L'élément **<imsmd:identifiant />** est utilisé pour stocker le numéro du concept. Il joue le rôle d'un container composé de **<catalog>** et **<entry>**. **<catalog>** prend la valeur DOI (Digital Object Identifier) et **<entry>** contient un numéro unique qui identifie le concept.

L'élément **<imsmd:title />** contiendra le titre du concept (chaîne de caractères).

L'élément `<imsmd:aggregationLevel>` contiendra le type du concept. Il est considéré comme un container. Il se compose d'un premier élément `<source>` définissant la source du vocabulaire utilisé dans le second élément `<value>`. Dans notre cas, la source sera « *up2uml* » et les valeurs autorisées : *Tache*, *ObjectifGlobal*, *ObjectifFinal*, *Sequence*, *Activite* et *Ressource*.

L'élément `<imsmd:structure/>` définit si la structure du composant est atomique ou composite.

L'élément `<imsmd:learningResourceType/>` décrit le type de ressource utilisée (definition, résumé, exemple,...)

L'élément `<location />` définit l'url de la page.

L'élément `<format/>` définit le type de présentation principalement utilisée dans le ressource. On utilise les types mime *text/html*, *video/mpeg*,...

La description de la ressource numéro R0001 de la Figure 4-39 est donnée en Figure 4-40.

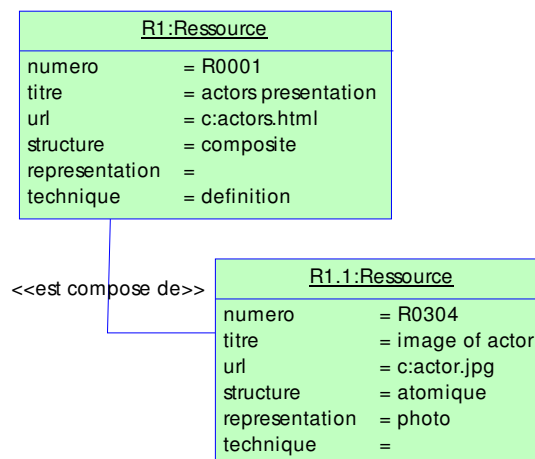


Figure 4-39 : exemple de concept <Ressource>

```

<imsmd:lom>
  <imsmd:general>
    <imsmd:identifiant>
      <imsmd:catalog>DOI</imsmd:catalog>
      <imsmd:entry>R0001</imsmd:entry>
    </imsmd:identifiant>
    <imsmd:title>
      <imsmd:string>actors presentation</imsmd:string>
    </imsmd:title>
    <imsmd:structure>
      <imsmd:source>LOMv1.0</imsmd:source>
      <imsmd:value>collection</imsmd:value>
    </imsmd:structure>
    <imsmd:aggregationLevel>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>Ressource</imsmd:value>
    </imsmd:aggregationLevel>
  </imsmd:general>
  <imsmd:technical>
    <imsmd:format>text/html</imsmd:format>
    <imsmd:location>c:actors</imsmd:location>
  </imsmd:technical>
  <imsmd:educational>
    <imsmd:learningResourceType>
      <imsmd:source>up2uml</imsmd:source>
      <imsmd:value>definition</imsmd:value>
    </imsmd:learningResourceType>
  </imsmd:educational>
</imsmd:lom>

```

Figure 4-40 : exemple de description LOM de concept <Ressource>

```

<resource identifier="R1 A1 P2" adlcp:scormType="sco"
href="actors.html" type="webcontent">

  <file href="c:actors.html" />
  <dependency identifierref="R1.1" />
</resource>

```

Figure 4-41 : exemple de description SCORM de concept <Ressource>

La balise <dependency> est utilisée pour représenter la relation de type « est composé de » (Figure 4-41)

Ainsi, avec cette approche, nous utilisons SCORM CAM comme langage de description d'ontologie d'apprentissage tout en structurant les ressources du contenu pédagogique. Par ailleurs, l'extension proposée de LOM permet de décrire les informations nécessaires à notre modèle d'adaptation.

4.5. Mise en œuvre et résultats

Up2UML [UP2UML, 2006] est un projet Européen financé par la Commission Européenne dans le cadre du programme Leonardo da Vinci. Il se compose de cinq partenaires : Fraunhofer Institute Experimental Software Engineering (IESE) Kaiserslautern, Germany ; Institut National Polytechnique de Toulouse (INPT) Toulouse, France ; National College of Ireland (NCI) Dublin, Ireland ; SOFTWIN Bucharest, Romania ; New Bulgarian University, Sofi, Bulgaria.

Up2UML a pour objectifs de :

- Développer de nouveaux matériels pédagogique de formation pour UML2⁴⁷ et RUP⁴⁸;
- Développer des programmes de formation pour différents groupes cibles;
- Développer des concepts novateurs de formation (par exemple formation à distance comportant des séances en ligne avec tuteur, apprentissage social, cycles d'étude virtuels, communauté de pratiques...)

UP2UML a été le champ expérimental du processus d'ingénierie d'objets d'apprentissage proposé dans cette thèse [Moebs et al., 2007]. Une étude a été menée dans quatre pays (Allemagne, Bulgarie, France, Roumanie) afin de recenser les besoins des entreprises en termes de formation sur RUP et UML. Cette étude a permis d'identifier 33 tâches métiers décomposées en 250 compétences et 700 objectifs pédagogiques. Un aperçu est fourni en Annexe 7 et Annexe 8.

Au moment de la rédaction de ce document, un prototype de contenus sur l'apprentissage du langage UML dans le cadre de RUP dans les entreprises a été développé en suivant notre processus d'ingénierie d'objets d'apprentissage. Il est composé de deux cours de 11 leçons chacun. Les deux cours abordent les deux thèmes suivant : « RUP comme méthode de développement de logiciel » et « définition des exigences fonctionnelles dans le processus RUP ». Ils peuvent être considérés comme représentatifs de l'ensemble des contenus à développer.

Une étude de satisfaction a été menée dans les quatre pays. Les sociétés enrôlées dans l'étude ont répondu aux deux questionnaires fournis en Annexe 9 et Annexe 10. Les résultats détaillés de l'étude sont présentés en Annexe 11.

Cette étude montre que globalement les résultats sont sensiblement équivalents dans les différents pays. 4 points sont abordés dans cette étude : processus d'apprentissage, contenus, exercices et évaluation, navigation dans les contenus.

⁴⁷ UML : Unified Modeling Language version 2

⁴⁸ RUP : Rational Unified Process

Le processus d'apprentissage a été en général apprécié par les participants. Les unités d'apprentissage sont clairement identifiées ainsi que les objectifs d'apprentissage. Malgré tout, les participants déplorent le faible nombre d'informations sur leur progression dans les contenus. Des informations telles que le pourcentage d'écrans visités, le pourcentage de tests faits et réussis seront à reconsidérer.

Les participants ont largement apprécié la structure des contenus et le choix des situations d'apprentissage proposées dans le prototype. Les exemples (diagramme, schéma, ...) choisis expliquent bien les concepts et la terminologie UML. Mais certains participants déplorent l'aspect basique abordé par les contenus du prototype. La version *concepts UML avancés* prévue pour la version finale du système devrait répondre à leurs attentes.

La disposition d'exercices en fin de chaque unité d'enseignement a été accueillie avec beaucoup de plaisir par les participants. Mais ces derniers auraient aimé avoir plus de commentaires sur les erreurs détectées par le système. Cette remarque peut être associée à la remarque générale du manque de feedback sur l'ensemble des activités proposées à l'élève.

Les outils de navigation (index, fonction de recherche) dans les contenus du prototype ont été approuvés par les participants. Mais certains ont éprouvé de la difficulté à revenir au début de l'unité d'enseignement après avoir progressé dans la structure du contenu.

4.6. Conclusion

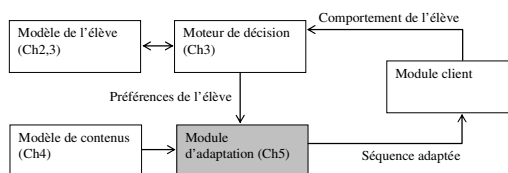
Ce chapitre a traité du modèle de contenus de notre système adaptatif. Nous avons proposé un modèle fondé sur trois modèles complémentaires : métier, pédagogique et ressources. Deux types de dépendances relient les concepts de ce modèle : composition et influence. Notre modèle a été présenté comme une méta-ontologie guidant la création d'un cours.

Une méthode d'ingénierie pédagogique a été proposée pour réaliser le modèle de contenus sous forme d'une ontologie. Cette méthode se compose de cinq étapes produisant ou enrichissant chacune l'ontologie à la base des trois modèles conceptuels. La phase de production crée les objets pédagogiques nécessaires.

L'ensemble des connaissances du domaine représenté par l'ontologie et les ressources pédagogiques est structuré et décrit. A cette fin, SCORM CAM est proposé pour la première fois comme langage de description d'ontologie assurant en même temps la structuration et l'indexation de contenus. Par ailleurs, nous avons proposé une amélioration des métadonnées LOM pour décrire les informations nécessaires pour l'adaptation de nos contenus. Cette approche normative a été choisie pour permettre l'exploitation de notre cours par des systèmes d'enseignement (LMS) répandus.

La méthode et le modèle proposés ont été mis en œuvre dans le projet européen UP2UML. Un cours de formation professionnelle en UML et RUP a été créé avec quatre autres partenaires européens. Le prototype produit a été évalué dans quatre pays par des utilisateurs cibles. L'analyse des résultats a permis de fournir un retour principalement aux phases de conception et de production.

Le chapitre suivant montre enfin le lien entre l'adaptation, ce modèle et les autres composants du système. On y décrit un modèle original d'adaptation qui se fonde sur le moteur de décision et les informations du modèle de contenus.



CHAPITRE

5. Adaptation

Résumé :

Ce chapitre traite de la stratégie d'adaptation. Nous y formulons le problème de l'adaptation comme un problème de recherche de séquence pertinente dans une base. Nous établissons l'architecture du module d'adaptation pour compléter notre système adaptatif. Nous proposons un modèle d'adaptation de la séquence en fonction de la navigation, l'activité, le média et le type de ressource. Pour mesurer la distance entre deux séquences, nous proposons une nouvelle mesure de similarité sémantique qui exploite l'indexation des contenus par des métadonnées. Enfin, nous présentons la simulation que nous avons effectuée pour valider l'approche. Des mesures quantitatives de performance sont utilisées.

5.1. Introduction

L'objet de la recherche documentaire (IR)⁴⁹ est de trouver dans une base documentaire, l'ensemble des documents qui sont les plus pertinents au regard des exigences formulées par un utilisateur [Boughanem et al., 2006]. La méthode la plus utilisée consiste à trier les documents en fonction de leur pertinence que l'on mesure en calculant une similarité [Tombros et van Rijsbergen, 2004] entre la requête de l'utilisateur et la description du contenu du document [Habegger et Quafafou, 2004].

Dans cette thèse, nous formulons le problème de l'adaptation comme étant la recherche de la séquence pédagogique la plus pertinente parmi celles dont on dispose. Pour évaluer cette pertinence, nous proposons d'utiliser une mesure de similarité sémantique originale. Celle-ci permet de mesurer une distance entre chacune des séquences disponibles et le vecteur de métadonnées retourné par le moteur de décision. Ainsi, nous recommanderions les activités pédagogiques qui s'adaptent le mieux au style d'apprentissage de l'élève, tel qu'on l'aurait estimé.

Ce chapitre présente cette approche. Nous commençons par rappeler les diverses mesures de pertinence proposées en recherche documentaire. Nous présentons ensuite, un modèle d'adaptation pédagogique original qui se fonde sur deux niveaux et quatre dimensions d'adaptation. Nous développerons alors la similarité sémantique proposée. La dernière section montre le fonctionnement de notre méthode à travers une simulation. Les mesures quantitatives de précision et de rappel sont utilisées pour évaluer les performances du système.

5.2. Mesures de pertinence en IR

Un système de recherche documentaire est souvent défini comme un logiciel qui aide l'utilisateur à trouver un ensemble d'informations correspondant à son besoin. Ces informations sont enfouies dans un certain nombre de documents que le système rassemble et classe par ordre d'importance dans une liste qu'il retourne à l'utilisateur. Généralement, la recherche ne se produit pas directement sur le contenu du document, mais plutôt sur un ensemble de métadonnées décrivant le document. Dans [Genest et Chein, 2004], les auteurs précisent que ces métadonnées peuvent être de deux sortes. Des données objectives telles que les noms des auteurs, le titre du document, les références sur l'éditeur, la date de publication, etc. Mais elles peuvent aussi être subjectives en tentant de décrire au mieux le contenu du document. La mesure de pertinence d'un document est alors déterminée par la correspondance entre une requête et la représentation du contenu du document.

Une représentation simple du contenu, appelée *index* est souvent utilisée. Un *index* $W = \{w_{t_1}, \dots, w_{t_n}\}$ est un ensemble de termes pondérés choisi parmi un ensemble fini. Les w_{t_i} expriment le niveau de pertinence du terme t_i faisant partie d'une requête de recherche de documents. Une requête est alors représentée par une expression booléenne de termes $q = t_1 \wedge t_2 \wedge t_3$ [Wondergem et al., 2000]. Le choix de la représentation du contenu a une influence directe sur les performances du système documentaire. Cette performance est communément mesurée par deux grandeurs : la précision (*Préc*) et le rappel (*Rap*), dont les expressions sont :

⁴⁹ IR : Information Retrieval

$$\left\{ \begin{array}{l} \text{Pr } ec = \frac{nf_{pert}}{nf} \\ \text{Rap} = \frac{nf_{pert}}{n_{pert}} \end{array} \right.$$

avec nf_{pert} le nombre de documents pertinents retrouvés, nf le nombre total de documents retrouvés et n_{pert} le nombre total de documents pertinents.

La pondération des termes est calculée par une analyse statistique. Plus précisément, le poids w_t d'un terme t dans un document est généralement estimé en combinant sa fréquence tf dans le document, la fréquence df des documents contenant t , ainsi que la longueur dl du document [Farah et Vanderpooten, 2006]. Dans ce sens, le système *Mercur*e [Boughanem et al., 1998] utilise une formule dérivée du système *Okapi* [Robertson et Walker, 1999] :

$$w_t = \frac{tf}{0.2 + 0.7 \times \frac{dl}{\Delta_l} + tf} \times \log\left(\frac{n_{tot}}{df}\right)$$

avec Δ_l la moyenne des longueurs de documents dans la collection et n_{tot} le nombre de documents dans la collection.

Ainsi, une mesure simple de similarité estimant la pertinence du document d pour la requête q pourrait être :

$$S_{qd} = \sum_{t \in q} \lambda_t \times w_{td}$$

avec λ_t un poids qui indique l'importance du terme t dans la requête q et w_{td} le poids du terme t dans l'index du document d .

Nous trouvons dans la littérature au moins deux façons de calculer cette mesure de similarité. La première consiste à arranger les termes du vecteur index afin de créer des formules de calculs ; la seconde consiste à comparer directement les termes dans le vecteur index.

5.2.1. Mesures de similarité par calcul

Pour évaluer par calcul, dans quelle mesure un document est pertinent pour une requête, une valeur de statut de recherche rsv ⁵⁰ est utilisée [Boughanem et al., 2006]. Le calcul de rsv est fait par des fonctions qui utilisent divers opérateurs reliant les poids des termes présents dans la requête. La liste des documents trouvés sera classée en fonction de la valeur rsv . Différents opérateurs sont utilisés pour combiner les poids des termes présents dans la requête. On trouve les opérateurs min et max [Fox et Show, 1994], l'opérateur de moyenne [Yager, 1988], l'opérateur linéaire pondéré [Craswell et al., 2005], l'opérateur de mesure de similarité [Frakes et Baeza-Yates, 1992], l'opérateur P-norms [Salton et al., 1983 ; Robertson, 1977] ou les opérateurs de disjonction / conjonction à logique floue [Dubois et Prade, 1984 ; Fodor et al., 1997].

L'opérateur de minimum pondéré Gödel $OW \min_G$ [Dubois et Prade, 1996] développe l'idée de donner moins d'importance aux faibles poids de l'index. Cet opérateur utilise un vecteur auxiliaire Γ qui exprime un niveau d'importance pour minimiser l'impact des poids faibles dans le calcul de la pertinence. Soit l'index $W = \{w_1, \dots, w_n\}$, avec w_i représentant le degré de pertinence du document pour le i^{th} terme de l'index. Nous considérons que les termes du vecteur sont classés (c'est à dire $w_i \geq w_{i+1}$). Et soit le vecteur auxiliaire $\Gamma = \{\varphi_1, \dots, \varphi_n\}$ définissant le niveau d'importance de chaque terme de l'index, avec $\varphi_i \geq \varphi_{i+1}$ et $\varphi_1 = 1$. L'idée est de donner plus d'importance (avec φ_i élevé) aux termes ayant un niveau de pertinence élevé.

L'opérateur de minimum pondéré Gödel $OW \min_G$ est alors défini par :

$$\varphi_i \rightarrow w_i = \begin{cases} 1 & \text{si } \varphi_i \leq w_i \\ w_i & \text{sin on} \end{cases}$$

et

$$OW \min_G(W, \Gamma) = \min_i(\varphi_i \rightarrow w_i)$$

En pratique, si les plus petits φ_i sont mis à 0, cet opérateur restreint le traitement aux w_i ayant les plus fortes valeurs (puisque les w_i ayant une faible valeur seront remplacés par 1 et ceux proches de 1 resteront inchangés).

La mesure de similarité entre un document d et une requête q est alors :

$$rsv(q, d) = OW \min_G(W, \Gamma)$$

⁵⁰ RSV : Retrieval Status Value

5.2.2. Mesures de similarité par comparaison

Le modèle possibiliste suggéré dans [Prade et Testemale, 1987] et utilisé dans [Boughanem et al., 2002] ne fournit pas une mesure de similarité par un calcul basé par les poids de l'index, mais par une paire de degrés de possibilité et de nécessité directement obtenue par l'exploitation des valeurs de l'index.

$$rsv(q, d) = (\Pi(q, d), N(q, d))$$

L'idée est de distinguer deux aspects de la notion de pertinence. Si le poids d'un terme dans un document est suffisamment élevé, alors ce terme est considéré comme certainement plus ou moins représentatif du contenu du document. Si le poids du terme n'est pas suffisamment élevé, alors le terme est simplement considéré comme une possible représentation du contenu du document. Dans ce cas, la mesure $rsv(q, d)$ exprime donc à quel degré il est possible et à quel degré il est certain que le document d soit pertinent pour la requête q .

L'exploitation de ce modèle, nécessite l'utilisation d'un index de représentation d'un document pour estimer le degré de possibilité et de certitude avec une requête. Chaque document est donc considéré comme un ensemble flou de termes [Kraft et al., 1999 ; Buell, 1982] obtenus par normalisation de l'index dans l'intervalle $[0,1]$. Une solution pour évaluer le degré de possibilité $\Pi(t, d)$ et de nécessité $N(t, d)$ d'un terme t dans un document d , est d'utiliser un système paramétré de transformation linéaire par morceau de $w_t = \frac{tf}{df}$ tel que :

$$\Pi(t, d) = \begin{cases} 0 & \text{si } w_t = 0 \\ 1 & \text{si } w_t \geq \alpha \\ \frac{w_t}{\alpha} & \text{sin on} \end{cases}$$

Et

$$N(t, d) = \begin{cases} 1 & \text{si } w_t = 1 \\ \frac{w_t - \alpha}{1 - \alpha} & \text{si } \alpha < 1 \text{ et } w_t \geq \alpha \\ 0 & \text{sin on} \end{cases}$$

Il est à noter que lorsque $\alpha = 0$ alors $\Pi = 1$ et $N = w_t$. Lorsque $\alpha = 1$ alors $\Pi = w_t$ et $N = 0$. Ainsi, l'évaluation du niveau de pertinence d'un document d pour une requête q composée des termes t_1, \dots, t_n consiste à calculer la paire de vecteurs $(\Pi(t_1, d), \dots, \Pi(t_n, d))$ et $(N(t_1, d), \dots, N(t_n, d))$. Ensuite les documents sont ordonnés en fonction des valeurs contenues dans les deux vecteurs en utilisant l'opérateur de comparaison de Pareto [Boughanem et al., 2006] :

$$(t_1, \dots, t_n) >_{\text{pareto}} (t'_1, \dots, t'_n) \text{ ssi } \forall i, t_i \geq t'_i, \exists j, t_j > t'_j$$

5.3. Adapter une séquence pédagogique au style d'apprentissage

Pour nous, adapter une séquence pédagogique au style d'apprentissage d'un élève consiste à rechercher dans un ensemble de séquences celle qui paraît être la plus proche des conditions idéales d'apprentissage exprimées pour un style d'apprentissage donné. Il faut donc pour cela être capable d'inférer le style d'apprentissage d'un élève en fonction des données observées, puis pour chaque style être capable d'inférer les caractéristiques optimales d'une séquence pédagogique.

Dans cette section, nous présentons l'architecture de notre système adaptatif qui fait apparaître deux niveaux d'adaptation. Ensuite, nous détaillons plus particulièrement le modèle d'adaptation du contenu en présentant les quatre dimensions qui le composent. Enfin, nous proposons une extension de la notion de mesure de similarité proposée en IR pour concevoir une fonction qui calcule la pertinence d'une séquence pédagogique pour un ensemble de caractéristiques de recherche.

5.3.1. Architecture du système adaptatif

L'architecture du système d'adaptation est composée de six parties (Figure 5-1) : un modèle de l'élève, un modèle de contenus et une base de contenus, un moteur de décision, un module d'adaptation et un module client.

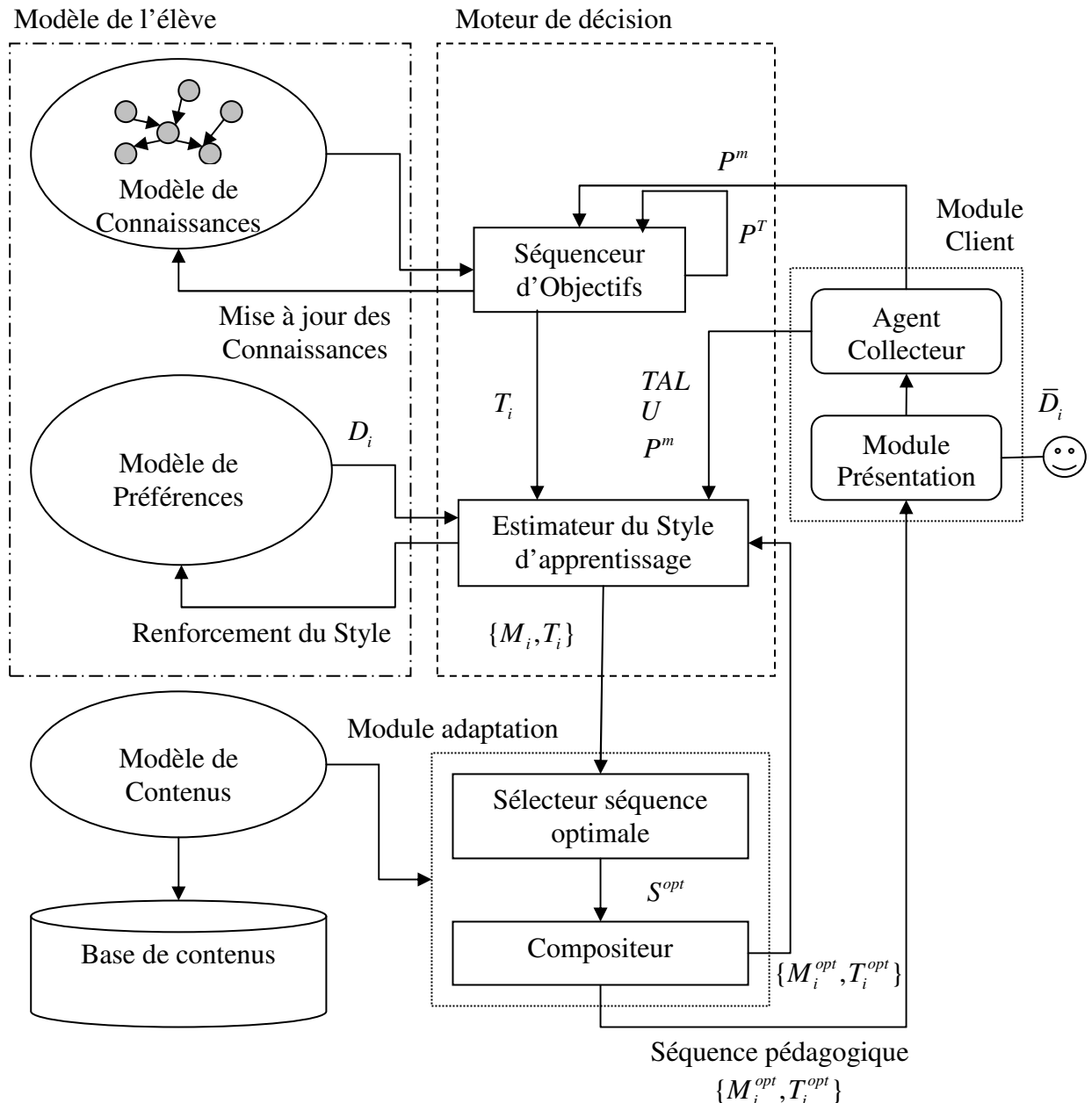


Figure 5-1 : architecture d'un système adaptatif (\bar{D}_i est le style réel inconnu de l'élève)

Le module client du système se charge de présenter les objets d'apprentissage à l'utilisateur et de collecter les informations comportementales de ce dernier. Nous rappelons qu'il s'agit là d'indicateurs observables tels que TAL, U, P^m (cf. chapitre 2). Ces indicateurs sont envoyés au moteur de décision qui estime la valeur des métadonnées $\{M_i, T_i\}$ décrivant le contenu de la séquence pédagogique optimale pour le style d'apprentissage inféré. Ces métadonnées forment le vecteur requête $q = M_i \vee T_i$ qui est envoyé au module d'adaptation. Ce dernier compose dynamiquement la séquence pédagogique qui est transmise au client. Le modèle de contenus est conçu à partir des ontologies du domaine et pédagogique que nous avons décrites au chapitre 4. L'ensemble

des combinaisons nécessaires pour couvrir les différents styles d'apprentissage, n'étant pas nécessairement disponible dans la base de contenus, le module d'adaptation tente de choisir la séquence pédagogique qui contient le maximum d'objets d'apprentissage avec des caractéristiques proches de celles souhaitées. Il utilise le modèle de contenus comme un index sur la base de contenus. La séquence pédagogique réellement composée sera au final, conçue avec les objets d'apprentissage optimaux découverts dans la base $\{M_i^{opt}, T_i^{opt}\}$.

5.3.2. Modèle d'adaptation

Notre système adaptatif utilise un modèle d'adaptation qui se compose de 2 niveaux (Figure 5-2). Le premier niveau d'adaptation vise à choisir la complexité de la tâche à proposer à l'apprenant. Le choix consiste donc à trouver l'objectif pédagogique que l'apprenant est capable d'atteindre en fonction de son niveau actuel de connaissances. Nous rappelons que ceci est réalisé par le séquenceur d'objectifs du moteur de décision en fonction de la performance observée $T_i = h(P^m)$ (cf. section 3.4).

Le second niveau d'adaptation est l'œuvre conjointe de l'estimateur de style, le sélecteur de séquences et le compositeur. Il vise quant à lui, à adapter le contenu de la séquence pédagogique au style d'apprentissage de l'élève. Nous rappelons que l'estimateur de style est responsable de l'estimation de la valeur des métadonnées de la séquence théorique idéale $M_i = f(T_i, P_i^{opt}, D_i)$ (cf. section 3.5). Cependant, il ne garantit pas le contenu réel de la séquence qui sera jouée à l'élève. C'est le moteur d'adaptation qui a cette responsabilité.

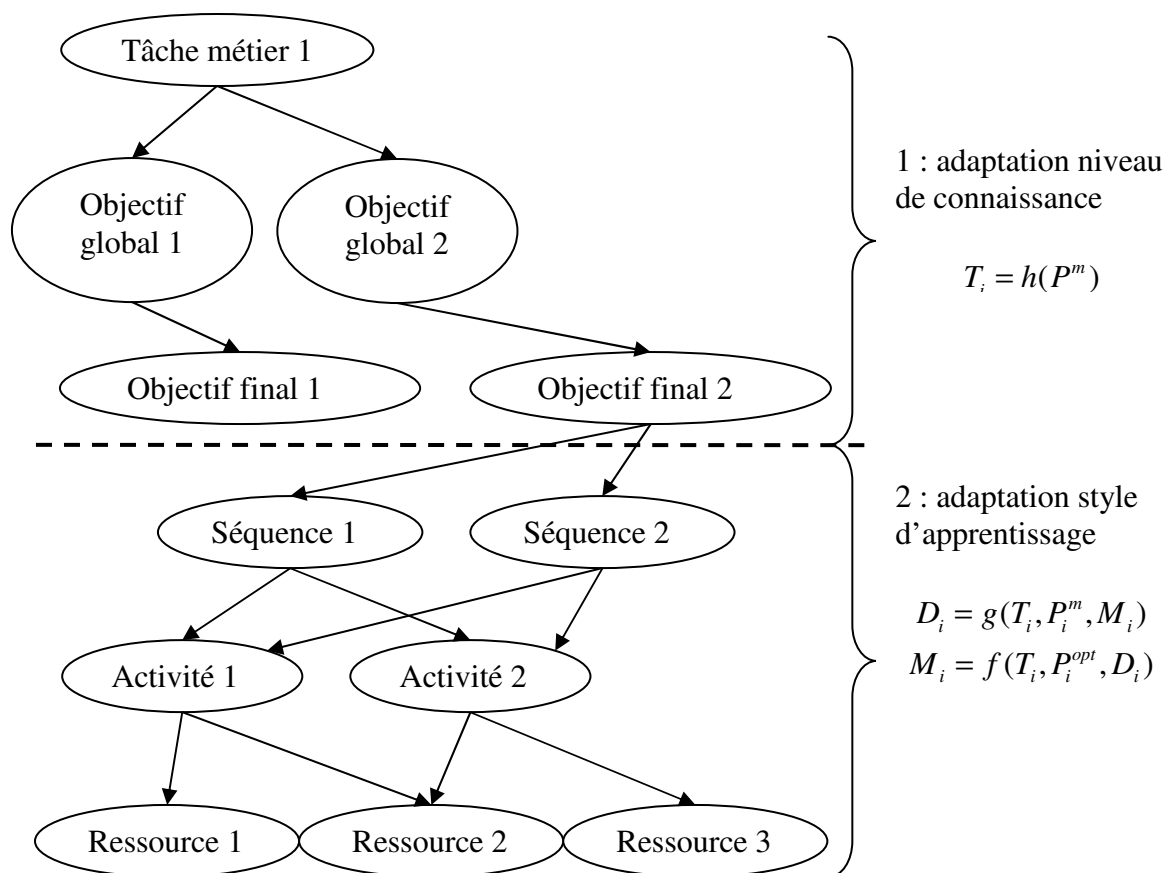


Figure 5-2 : Deux niveaux d'adaptation

Le module d'adaptation se compose de deux étages : un sélecteur de séquences et un compositeur. A travers son sélecteur de séquence, le module d'adaptation utilise comme critères de sélection, les données caractérisant l'activité pédagogique T_i et la séquence pédagogique M_i (cf. chapitre 2), telles que définies dans l'ontologie du domaine et l'ontologie pédagogique (cf. chapitre 4).

La recherche d'une séquence pédagogique consiste à trouver, dans la base de contenus, la séquence dont la représentation $\{M_i, T_i\}$ est semblable à la description inférée par le modèle de l'élève (cf. chapitre 3). En pratique, cette recherche s'effectue parmi un nombre limité d'activités. En effet, pour une séquence donnée, toutes les combinaisons nécessaires pour couvrir l'ensemble des styles d'apprentissage, ne sont pas obligatoirement disponibles. Il faut alors, sélectionner la séquence dont la sémantique est la plus proche de la sémantique exprimée par la requête.

Afin d'effectuer cette sélection, nous avons défini quatre dimensions sémantiques d'adaptation de séquence pédagogique (Tableau 5-1). Nous les utilisons conjointement pour calculer la similarité sémantique d'une séquence théorique optimale à une séquence réelle décrite dans la base de contenus. Cette tâche est la responsabilité du sélecteur de séquence optimale. Il évalue une fonction de pertinence basée sur ces quatre dimensions. Le Tableau 5-2 présente la correspondance entre les éléments structurels du contenu (cf. chapitre 4) et les points d'adaptation désirés.

Dimensions sémantiques d'adaptation	description
1 – Navigation	cette dimension offre le moyen de choisir la façon avec laquelle un élève se déplacera dans la structure des contenus pédagogiques. Son déplacement pourra être libre, dans ce cas l'élève choisira sans contrainte particulière le contenu pédagogique à visualiser. Dans le cas où l'élève désire être guidé par le système, une progression contextuelle lui sera proposée.
2 – Activité	cette dimension permet de choisir l'activité pédagogique à proposer à l'élève pour qu'il soit dans les meilleures conditions pour atteindre l'objectif visé. Le choix se porte sur des activités qui peuvent être de type expositive, active ou interrogative.
3 – Média	cette dimension consiste à choisir le type de media à utiliser comme support de l'information. Ici, on considère comme media potentiel, les dessins, le son, le texte, les schémas, la vidéo,...
4 – Type de ressource	cette dimension consiste à choisir les types de ressources qui seront utilisés dans la séquence. Ici, on considère comme type de ressource potentiel, les définitions, les exemples, les exercices, les évaluations,...

Tableau 5-1 : dimensions d'adaptation sémantique d'une séquence pédagogique

Éléments structurels des contenus	Adaptation sémantique	Termes représentant le contenu
Séquence	Environnement	M_2
Activité	Activité	T_2
Ressource	Média, Type de ressource	$M_1, M_3, M_4, M_5, M_6, M_7$

Tableau 5-2 : correspondances adaptation / structure de contenus

Ainsi dans notre système, le sélecteur reçoit du moteur de décision la requête, $q = T_2 \vee M_1 \vee M_2 \vee M_3 \vee M_4 \vee M_5 \vee M_6 \vee M_7$ composée toujours du même nombre de termes. Ces termes représentent le contenu de la séquence souhaitée (Tableau 5-2) pour les quatre dimensions d'adaptation désirées : la navigation, l'activité, le média et le type de ressource. Les séquences pédagogiques sont décrites en respectant la structure de contenus proposée dans le chapitre 4. Pour un objectif pédagogique donné, il est donc possible d'avoir plusieurs séquences pédagogiques mettant en relation différentes activités et ressources. Le compositeur, quant à lui, part de cette séquence optimale et tente de composer dynamiquement le contenu réel à proposer à l'élève. Il choisira les objets d'apprentissage les plus compatibles avec la séquence optimale. On notera ces objets optimaux (M_i^{opt}, T_i^{opt}).

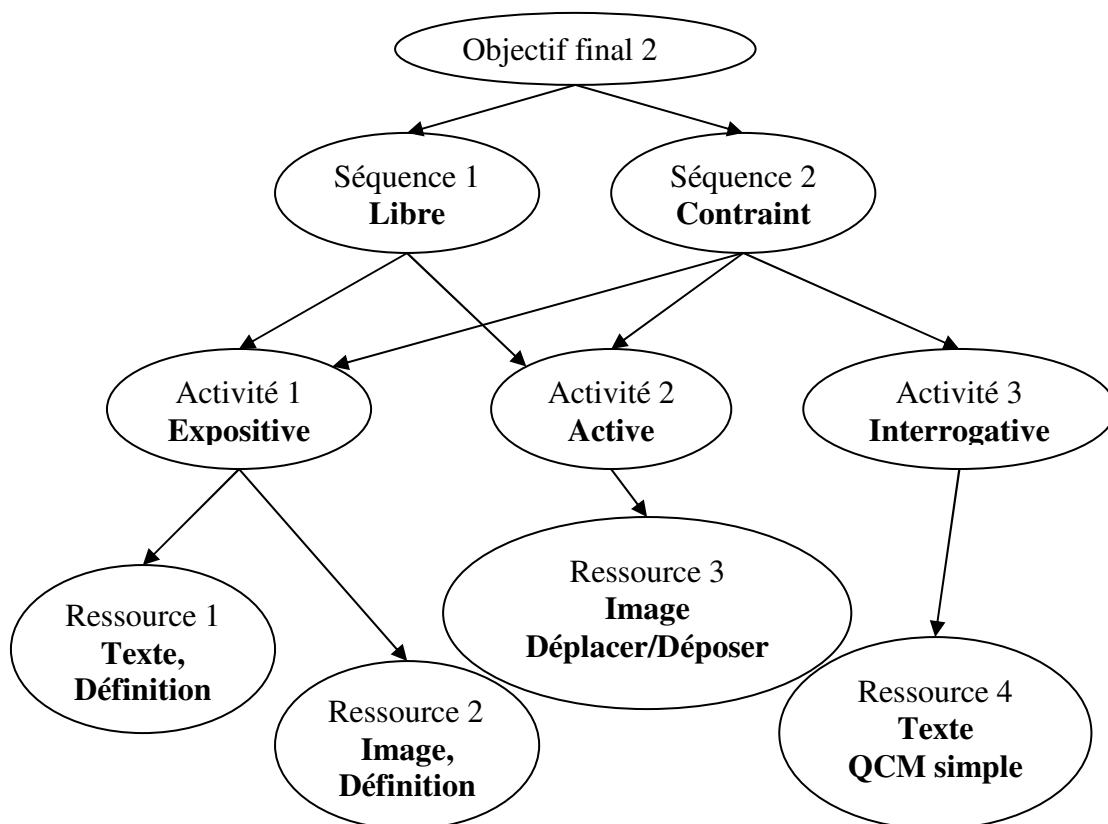


Figure 5-3 : Structure de contenus adaptatifs

Dans l'exemple de la Figure 5-3, si le sélecteur reçoit la requête $q = Expositive \vee Libre \vee Image \vee \neg Re\ summary \vee Définition \vee DéplacerDéposer \vee \neg Evaluation$, alors il proposera la séquence 1 et aura le choix pour la conception de l'activité 1. En effet, dans la base de contenus, il y a deux versions de ressources. Le compositeur choisira la ressource 2 qui est proche de la demande (Image) pour l'activité expositive. Par contre, l'activité 2 sera obligatoirement composée de la ressource 3 car elle est la seule disponible.

Il est à noter que dans ce contexte, nous n'avons jamais de grande « masse » de données à manipuler. Il existera uniquement une dizaine de séquences pédagogiques qui couvriront un certain nombre de situations différentes d'apprentissage. La difficulté se situe ailleurs. Le sélecteur doit être capable dans un ensemble restreint de séquences pédagogiques, de choisir celle qui comporte le plus de similarités avec la requête $q = M_i \vee T_i$. Pour ce faire, il utilise la mesure de similarité sémantique décrite dans la section suivante.

5.3.3. Mesure de similarité sémantique

Dans [Cullot et al., 2003], les auteurs précisent que la modélisation de la sémantique de données structurées passe par la définition de métadonnées sur lesquelles il est possible de définir des mesures de similarité ou de distance sémantique. En s'appuyant sur ces travaux, nous présentons dans ce chapitre une méthode de calcul d'une mesure de similarité de la description sémantique de la séquence pédagogique ; celle-ci étant décrite par les métadonnées des ontologies pédagogiques et du domaine.

Nous proposons de représenter le contenu d'une séquence pédagogique par l'index W , qui est vu comme un ensemble de termes pondérés choisi parmi l'ensemble des métadonnées T_i et M_i :

$$W = \{w_{T_2}, w_{M_1}, w_{M_2}, w_{M_3}, w_{M_4}, w_{M_5}, w_{M_6}, w_{M_7}\}$$

Les w_i expriment un niveau de pertinence du terme i faisant partie d'une requête de recherche de séquence pédagogique. Nous rappelons qu'une requête est une expression booléenne de termes $q = T_2 \vee M_1 \vee M_2 \vee M_3 \vee M_4 \vee M_5 \vee M_6 \vee M_7$.

La pondération w des termes est calculée par une analyse statistique en s'appuyant sur la formule proposée dans le système *Mercur* [Boughanem et al., 1998] :

$$w_t = \frac{tf}{0.2 + 0.7 \times \frac{sl}{\Delta_i} + tf} \times \log\left(\frac{n_{tot}}{sf}\right)$$

avec tf la fréquence d'apparition du terme dans la séquence, sl la longueur de la séquence (nombre d'activités x nombre de ressources), Δ_i la moyenne des longueurs des séquences pour l'objectif pédagogique cible, n_{tot} le nombre de séquences dans la collection pour l'objectif pédagogique cible, sf le nombre de séquences dans lesquelles le terme apparaît.

A partir de là, il est possible d'utiliser n'importe quelle méthode de calcul de similarité utilisée en IR. Nous avons adopté pour une mesure simple de similarité entre une requête q et une séquence s donnant une estimation de la pertinence de la séquence pour la requête :

$$S_{qs} = \sum_{t \in q} \lambda_t \times w_{ts}$$

avec λ_t un poids qui indique l'importance du terme t dans la requête q et w_{ts} le poids du terme t dans l'index pour la séquence s . Dans nos expérimentations, nous avons fixé le vecteur λ_t à $\{\lambda_{T_2} = 1, \lambda_{M_1} = 1, \lambda_{M_2} = 0.8, \lambda_{M_3} = 1, \lambda_{M_4} = 1, \lambda_{M_5} = 1, \lambda_{M_6} = 1, \lambda_{M_7} = 1\}$. λ_{M_2} est mis à 0.8 pour exprimer l'effet modéré de la dimension navigation tel que démontré par l'étude empirique du chapitre 2.

Le module d'adaptation utilise cette mesure de similarité sémantique pour sélectionner la séquence optimale à présenter à l'élève pour un style d'apprentissage donné. Les Figure 5-4 et Figure 5-5 présentent les algorithmes respectivement du sélecteur de séquence et du compositeur.

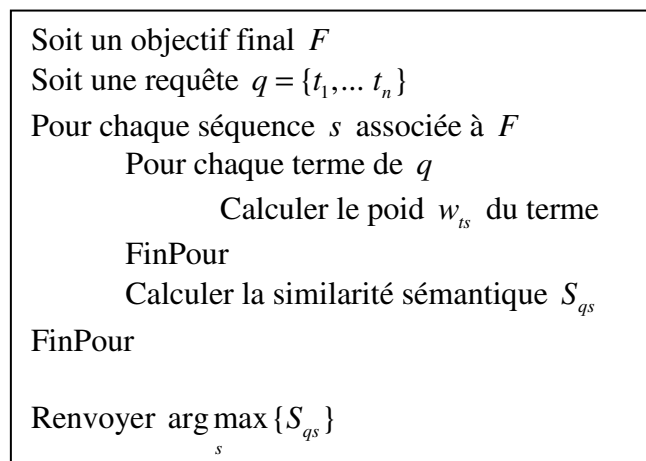


Figure 5-4 : algorithme du sélecteur de séquence

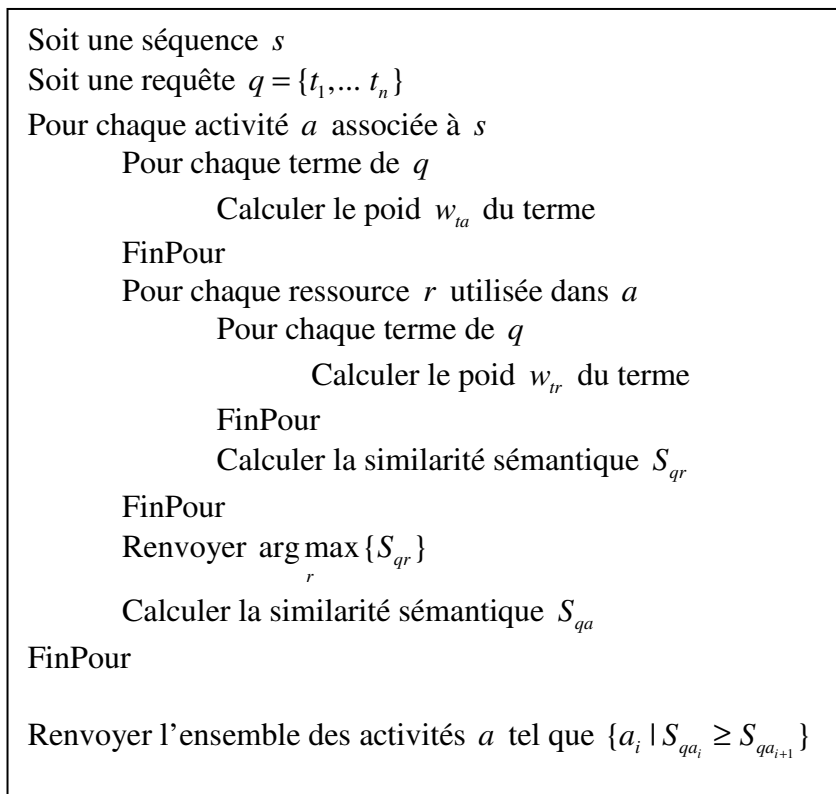


Figure 5-5 : algorithme du compositeur

Dans cette solution, nous considérons que toutes les activités de la séquence choisie par le sélecteur seront proposées à l'élève. Pour chaque activité, le compositeur choisira la ressource la plus pertinente pour la requête. L'ensemble des activités sera classé par similarité décroissante. Ainsi, sera proposée en premier à l'élève, l'activité la plus pertinente.

5.4. Simulations et Résultats

Pour évaluer le système d'adaptation, nous avons utilisé un générateur de requêtes qui simule la sortie du moteur de décision. Cette requête est passée au sélecteur de séquence qui retourne une liste de séquences pédagogiques classée par ordre décroissant suivant la valeur de la mesure de similarité sémantique S_{qs} .

Pour évaluer la qualité de la réponse, nous avons utilisé les mesures de précision et de rappel dont nous avons adapté les expressions :

$$\begin{cases} Prec = \frac{nf_{pert}}{nf} \\ Rap = \frac{nf_{pert}}{n_{pert}} \end{cases}$$

avec nf_{pert} le nombre de séquences pertinentes retrouvées, nf le nombre total de séquences retrouvées et n_{pert} le nombre total de séquences pertinentes.

Afin d’effectuer cette simulation, nous avons établi une base virtuelle d’objets d’apprentissage, qui suit la structure de la Figure 5-3. Nous y avons mis toutes les combinaisons de caractéristiques possibles couvrant les 16 classes de styles d’apprentissage formées par les quatre axes (Tableau 5-3). En effet, pour chacune des classes, nous avons simulé une séquence pédagogique appropriée. C’est-à-dire que les métadonnées sont en adéquation avec les dimensions du style.

D_1	D_2	D_3	D_4	Environnement	Activité	Média, Type Ressource
visuel	global	réfléchi	déductif	Libre	Expositive	Image, Définition, exemple
visuel	global	réfléchi	inductif	Libre	Expositive	Image, Expérience, Simulation
visuel	global	actif	déductif	Libre	Active	Image, Définition, exemple
.....
verbal	global	actif	inductif	Libre	Active	Texte, Expérience, Simulation
verbal	sequentiel	réfléchi	déductif	Contraint	Expositive	Texte, Définition, exemple
verbal	sequentiel	réfléchi	inductif	Contraint	Expositive	Texte, Expérience, Simulation
verbal	sequentiel	actif	déductif	Contraint	Active	Texte, Définition, exemple
verbal	sequentiel	actif	inductif	Contraint	Active	Texte, Expérience, Simulation

Tableau 5-3 : 16 modèles de séquences pédagogiques

Par exemple, pour un style {visuel, global, réfléchi, inductif}, nous avons créé les caractéristiques {libre, expositive, image, expérience, simulation}.

Pour chacune des combinaisons, nous avons alors varié le nombre et le type des exercices, le nombre et le type d'évaluation, et le nombre de définitions. Ainsi, pour chaque classe de style nous avons abouti à plusieurs séquences virtuelles ayant des caractéristiques variées. Cependant, dans une classe, les quatre dimensions d'adaptation restent fixes. Seulement, les autres caractéristiques varient. Par exemple, pour un style actif nous avons créé plusieurs activités pratiques en variant le type d'exercices : déplacer-déposer, QCM, remplir les vides etc. Lors de la recherche de séquences, on devrait sélectionner toutes ces activités pour un style actif.

Au final, la base d'objets d'apprentissage a été composée de n_{tot} séquences pédagogiques définies pour un certain objectif pédagogique. Ensuite, nous avons défini un ensemble de 16 requêtes spécialement conçues pour identifier un sous-ensemble n_{pert} des n_{tot} séquences pédagogiques. En fait, chaque requête identifie une classe de style d'apprentissage. Nous notons $Prec_1$ et $Prec_5$ le calcul de la précision du système respectivement sur la première et les 5 premières séquences retrouvées. Le Tableau 5-4 présente les résultats obtenus sur les 16 requêtes.

	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15	q16
$Prec_1$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$Prec_5$	0.8	0.7	1	0.6	1	0.9	0.7	0.8	0.9	1	0.6	0.9	1	0.8	1	0.9
Rap	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Tableau 5-4 : Précision et Rappel du système

On observe que le système obtient une mesure de rappel et de précision sur la première séquence retournée de 100%. On note ensuite une moyenne de 85% de précision sur les 5 premières séquences retournées sur les 16 requêtes. Nous pouvons donc conclure que le niveau de qualité du système est très satisfaisant puisqu'en pratique nous chercherons à présenter à l'élève, la première séquence retrouvée par le système.

Ces bons résultats observés sont bien sûr dû au fait que nous avons dans notre base au moins une séquence pour chaque classe de style d'apprentissage. L'idéal est donc de veiller à ce que se soit toujours le cas. Mais s'il arrivait qu'un style ne trouve pas de séquence pédagogique adaptée, notre mesure serait capable de choisir une séquence qui soit la plus proche de celle souhaitée.

5.5. Conclusion

Ce chapitre a traité du chaînon final de notre système : la stratégie d'adaptation. Nous y avons formulé l'adaptation comme étant un problème de recherche d'une séquence pertinente dans une base de contenus. Ainsi, nous avons complété l'architecture du système en ajoutant un module d'adaptation en deux étages : un sélecteur de séquence optimale et un compositeur de contenu. Ce module implante un modèle d'adaptation de la séquence fondé sur quatre dimensions ; la navigation, l'activité, le média, et le type de ressources. Pour l'étage de sélection de la séquence optimale, nous avons établi une mesure de similarité sémantique nouvelle. Elle est fondée sur le principe d'indexation fréquentielle des séquences par les métadonnées définies au chapitre 4. Une pondération a été proposée pour prendre en considération l'importance des axes du style. L'approche a été validée par une simulation et les performances ont été évaluées à l'aide des mesures de précision et de rappel.

Cependant, cette approche reste théorique car elle n'a pas été testée en situation réelle. Elle nécessiterait l'ajustement des critères de sélection ainsi que leur pondération. En effet, on pourrait envisager la personnalisation des poids des métadonnées pour exprimer la prédominance des axes du style chez chaque élève.

Conclusion générale

La problématique posée à l'origine de cette thèse a été l'adaptation de la pédagogie à la manière dont l'élève *préfère* apprendre. Nous avons tenté d'aller au-delà des méthodes classiques d'adaptation qui ajuste le contenu au niveau de connaissances de l'élève. Notre approche a été de s'inspirer des travaux sur la pédagogie différenciée pour proposer aux élèves des activités leur permettant d'exploiter au mieux leur capacité. Nous nous sommes donc posés les questions suivantes :

- comment caractériser les préférences d'apprentissage de l'élève ?
- comment représenter un cours de manière à y intégrer une richesse d'informations et d'activités permettant l'adaptation ?
- comment choisir les activités pédagogiques pour s'adapter aux préférences de l'élève ?

Ces questions ont constitué la trame qui a guidé l'ensemble de nos travaux. Pour répondre à la première question, nous avons étudié le style d'apprentissage. Nous avons mené une étude empirique afin d'établir un modèle de dépendance entre le style, la pédagogie, et le comportement de l'élève. Nous avons établi que l'ajustement de la séquence pédagogique au style de l'élève influence son comportement et ses performances. Nous avons aussi défini trois nouvelles mesures pour quantifier le comportement de l'élève : le degré d'attention, l'énergie d'apprentissage, et le niveau d'adaptation. Le degré d'attention évalue l'effort que l'élève a fourni lors d'une activité pédagogique. L'énergie d'apprentissage mesure la quantité de travail effectué. Alors que le niveau d'adaptation mesure à quel point l'activité et les ressources utilisées ont été compatibles avec le style de l'élève.

Ce modèle de dépendance est devenu la base de la suite de la thèse. En premier, nous l'avons exploité pour créer un modèle probabiliste des préférences de l'élève répondant ainsi à la première question de recherche. Nous avons montré comment on peut apprendre par supervision ce modèle à partir des données empiriques et le renforcer continuellement lors de l'exploitation. Nous avons créé un algorithme qui exploite ce modèle de deux manières complémentaires : affiner l'estimation du style d'apprentissage de l'élève, et inférer les préférences pédagogiques. Nous en avons proposé deux implémentations non paramétriques différentes. La première adopte la technique des réseaux bayésiens. Alors que la deuxième se fonde sur le formalisme de machines à vecteurs de support.

Pour traiter la deuxième question, nous avons repris les caractéristiques pédagogiques que nous avons structurées et complétées pour former un modèle de contenus. Ce modèle consiste en une ontologie qui combine le domaine à apprendre, les informations pédagogiques, et les propriétés des ressources physiques. Le domaine est représenté par les tâches et les compétences métier structurées en terme de dépendances et de composition. La pédagogie est décrite par les objectifs pédagogiques au sens de la théorie de Bloom, les types de connaissances selon la classification de Gagné, la nature de l'activité : expositive, interrogative ou active. La ressource est décrite, entre autres, en termes de navigation, type de contenu, média, type de pratique et type d'évaluation. Nous avons développée une méthode d'ingénierie en cinq étapes qui permet de produire l'ontologie et les objets d'apprentissage nécessaires. Nous avons structuré l'ensemble du cours, incluant ontologie et ressources, en utilisant SCORM content aggregation model.

Afin d'intégrer les informations nécessaires à l'adaptation, nous avons étendu les métadonnées LOM.

Pour être capable de proposer la séquence pédagogique appropriée au style de l'élève, nous avons élaboré une stratégie d'adaptation en quatre dimensions : activité, navigation, média et ressource. Pour implémenter cette approche, nous avons formulé l'adaptation comme étant un problème de recherche de la séquence pertinente étant données les préférences de l'élève. Nous avons associé les dimensions d'adaptation à l'ensemble des métadonnées décrivant la sémantique du contenu. Nous avons proposé une mesure de similarité sémantique qui nous permet d'évaluer la pertinence d'une séquence pour un ensemble de préférences. L'adaptation se fait alors en deux étapes. Premièrement, on sélectionne la sélection pertinente pour le style. Deuxièmement, on compose le contenu à proposer en ajustant au mieux les ressources disponibles aux préférences de l'élève.

Les trois modèles proposés ont été validés de différentes manières. Le modèle de préférences a été validé à l'aide d'une partie des données de notre étude empirique. Le modèle de contenu a été mis en œuvre concrètement dans le projet européen, multi-partenaires, UP2UML. Un cours sur le sujet de RUP et UML a été créé et évalué par un échantillon de la population cible. Le modèle d'adaptation a été, quant à lui, validé par simulation.

Limitations

Les travaux présentés dans cette thèse présentent quelques limitations, certaines dues à des facteurs extérieurs, d'autres à la portée assez large du sujet traité. Nous citons en particulier :

- vu le nombre trop important de variables à étudier, notre étude empirique a porté uniquement sur cinq caractéristiques pédagogiques
- la formalisation du modèle de connaissances reste à compléter
- du, aussi, au manque de données pour l'entraînement, notre modèle de préférences a été uniquement validé sur les données de l'étude empirique
- la méthode de création de cours nécessite une boîte à outils flexible et exploitable par des concepteurs non informaticiens
- les aspects métiers ne sont pas pris en considération dans notre stratégie d'adaptation

Perspectives

Le travail fait dans cette thèse ouvre la voie à plusieurs perspectives de recherche.

Un modèle étendu de préférences

Le modèle de préférences que nous avons développé peut être amélioré de plusieurs manières. Premièrement en mettant en œuvre une étude empirique plus large par le biais d'internet. Une activité d'apprentissage et résolution de problème grand public doit être conçue selon toutes les caractéristiques pédagogiques que nous avons énoncées. Cette activité ne sera pas scolaire mais plutôt liée à un domaine d'intérêt partagé par un grand public comme l'environnement. Le questionnaire du style d'apprentissage doit alors être adapté au public et au domaine. Des dispositions particulières doivent être prises pour maximiser la fiabilité des mesures par internet. La méthode d'apprentissage des paramètres

du modèle sera améliorée en prenant en considération les degrés de certitude liée à la mesure du style d'apprentissage. Les dimensions du style d'apprentissage seront codées sur une plage de valeurs plus étendue permettant de détecter des nuances de préférences.

Modèle de connaissances

Le développement du modèle de connaissances selon la méthode que nous avons esquissée dans cette thèse nécessite un effort significatif. Deux contributions sont nécessaires à cela. Une doit porter sur la manière d'évaluer les probabilités conditionnelles à partir de données factuelles comme la longueur des séquences d'un objectif pédagogique, la durée théorique nécessaire à son accomplissement, un niveau de difficulté qui complète le niveau de bloom, le nombre d'objectifs prérequis... La deuxième contribution portera sur la méthode pour prédire la performance théorique d'un élève en fonction de son niveau de connaissance. Un algorithme de résolution par programmation dynamique individualisée à l'élève devrait être investigué.

Modèle de contenus « blended »

Dans le cadre du projet UP2UML, nous avons spécifié et développé des activités pédagogiques variées et novatrices comme la discussion dans des forums, le brainstorming, les visioconférences, le chat, les projets, les travaux pratiques, le tutorat, l'annotation sémantique etc. Des métadonnées ont été créées pour décrire ces activités et la manière de les mettre en œuvre. Notre objectif est d'étendre notre modèle de contenus en y intégrant une sémantique plus riche, en ajoutant par exemple des opérateurs ET/OU/XOU pour relier les concepts pédagogiques. Nous utiliserions un formalisme plus flexible que SCORM pour décrire l'ontologie. La méthode d'ingénierie proposée sera affinée en créant des outils d'aide et de conception intuitifs. Il sera alors possible de généraliser notre modèle d'adaptation pour créer un framework généralisé d'apprentissage adaptatif. Cette perspective a été à la base de la proposition de projet européen ABLE actuellement en deuxième phase d'évaluation dans le cadre de l'appel à projet FP7 - IST 2007 (ICT).

Modèle d'adaptation pour gérer les compétences métier

La couche métier dans notre modèle de contenus n'a pas été intégrée dans la stratégie d'adaptation. En effet, l'un des objectifs du projet UP2UML a été de fournir une formation adaptée au profil métier des apprenants (exemple analyste, programmeur, concepteur...). Cette adaptation s'est faite d'une manière statique en créant des organisations SCORM distinctes pour chaque profil. Nous souhaitons étendre notre approche à l'adaptation pour permettre de cibler les compétences métier appropriées pour le profil métier de l'individu. Pour cela, il est nécessaire d'inclure des informations liées au profil dans le modèle de contenu. Le système d'adaptation servirait alors comme système d'apprentissage mais aussi comme système d'aide à la gestion et au développement des compétences dans le milieu professionnel.

Bibliographie

- ADAPT project. "Adaptivity and adaptability in ODL based on ICT" (European Community Socrates Minerva 101144-CP-1-2002-NL-MINERVA-MPP). <http://www.wis.win.tue.nl/~alex/HTML/Minerva/>.
- ADL SCORM V1.3 : Sharable Content Object Reference Model, www.adlnet.org
- Alani H., Kim, S., Millard D.E., Weal, M.J., Hall, W., Lewis, P.H., Shadbolt, N.R.: Automatic Ontology based knowledge extraction from Web documents. IEEE computer society, IEEE intelligent system (2003)
- Albrecht D., Zukerman I., Nicholson A. , 'Bayesian Models for Keyhole Plan Recognition in an Adventure Game', *User Modeling and User-Adapted Interaction*, 8, (1 – 2), 5-, 1998
- Alesso, P.H., Smith, C.F.: Developing Semantic Web Services. A K Peters ltd, Wellesey MA, Canada (2004) 165-272
- Alexandraki C., Paramythis A., Maou N., Stephanidis C., 'Web Accessibility through Adaptation', *Lecture Notes in Computer Science*, 3118, 302 – 309, 2004
- Ardissono L., Barbero C., Goy A, and Petrone G.. An agent architecture for personalized web stores. *In Proc. 3rd Int. Conf. On Autonomous Agents*, pp. 182–189, Seattle, WA, 1999.
- Ardissono L., Brna P., and Mitrovic A. Towards User Modeling Meta-ontology, (Eds.): *UM 2005, LNAI 3538*, pp. 448 – 452, 2005. © Springer-Verlag Berlin Heidelberg, 2005
- ARIADNE project. Available online at: <http://ariadne.unil.ch/Metadata/> [02/2000
- Arteaga, C., Fabregat, R., Eyzaguirre, G. & Merida, D. Adaptive Support for Collaborative and Individual Learning (ASCIL): Integrating AHA! and CLAROLINE, AH'04, Eindhoven, The Netherlands, 2004
- Azouaou Faïçal, Desmoulins Cyrille Using and Modeling Context with Ontology in e-learning: the Case of Personal Annotations, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), Dublin, Ireland, June 21-23 2006a
- Azouaou, F., Desmoulins, C.: MemoNote; a context aware annotation tool for teachers. IEEE 1-4244-0406-1, 2006b
- Azouaou, F., Desmoulins, C.: A flexible and extensible architecture for context aware annotation in e-learning. IEEE international conference on advanced learning technologies, 2006c

- Bajraktarevic, N., Hall W. and Fullick P. Incorporating learning styles in hypermedia environment: Empirical evaluation. *AH 2003 workshop session at HT'03 conference, Nottingham, UK*, available online at <http://www.wis.win.tue.nl/ah2003/schedule-ht.html>, 2003
- Baker M., the roles of models in artificial intelligence and education research: a prospective view, *international journal of artificial intelligence in education*, 11, 122-143, 2000
- Balacheff N., Luengo V., Vadcard L., *Environnements informatiques pour l'apprentissage humain*, Chapitre 2, Cognition et traitement de l'information, Hermes Lavoisier, ISBN 2-7462-1171-8, mars 2006
- Barbe, W. B., and M. N. Milone, What we know about modality strengths, *Educational Leadership*, 378-380, 1981.
- Barker T., Jones S., Britton C., Messer D., 'The Use of a Co-operative Student Model of Learner Characteristics to Configure a Multimedia Application', *User Modeling and User-Adapted Interaction*, 12(2 – 3), 207 – 241, 2002
- Batatia, H., Piombo, C., Ayache, A. (2006). Probabilistic estimation of learning styles from observable data, TICE2006, Toulouse, octobre 2006
- Scott Bateman, Christopher Brooks and Gord McCalla, Collaborative Tagging Approaches for Ontological Metadata in Adaptive E-Learning Systems, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), June 21-23 2006, Dublin, Ireland
- Beenen F., Van Der Werff J.J., Statisticophobie. Een onderzoek naar de weerstanden tegen statistiek. *Nederlands Tijdschrift Psychologie*, 22, 209-217, 1976
- Bermudez L., Piasecki M., Metadata Community Profiles for the Semantic Web, *Geoinformatica*, Springer 10:159-176, 2006
- T. Berners-Lee, J. Hendler, O. Lassila, The Semantic Web, *Scientific American*, May 2001, pp. 28-37. <http://www.scientificamerican.com/2001/0501issue/0501berners-lee.html>, 2001
- Berners-Lee, T. Semantic Web Status and Direction *ISWC2003 keynote, ISWC'03*, <http://www.w3.org/2003/Talks/1023-iswc-tbl/slide26-0.html> , 2003
- Billsus, D. and Pazzani, M., A hybrid user model for news story classification. In: *User Modeling (Proceedings of the Seventh International Conference)*. Banff, Canada, pp. 99-108, 1999.
- Bloom B. S. *Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain*. New York ; Toronto: Longmans, Green, 1956.
- Bormuth, J. R., Cloze Test Readability: Criterial reference scores, *Journal of Educational Measurement* 5(3), 189-196, 1968.

- Boser B.E., Guyon I.M., Vapnik V., A Training Algorithm for Optimal Margin Classifiers, Proc. Fifth Ann. Workshop Computational Learning Theory, ACM Press, pp. 144-152, 1992
- Boughanem M., Dkaki T., Mothe J., Soule-Dupuy C., Mercure at TREC-7. In : Proc. Of TREC-7, 135-141, 1998
- Boughanem M., Loiseau Y., Prade H., Graded pattern matching in a multilingual context. Inc.: Proc. 7th Meeting Euro Working Group on Fuzzy Sets, Eurofuse, Varena, 121-126, 2002
- Boughanem M., Loiseau Y., Prade H., Rank-Ordering Documents according to their relevance in information retrieval using refinements of ordered-weighted aggregations, AMR 2005, LNCS3877, Springer-Verlag, pp. 44-54, 2006
- Bourdieu P., le sens pratique, les éditions de minuit, Paris, 1980
- Bouzidi D., Ajhoun, R., Benkiran, A.: The use of the annotations in a collaborative activity of reading an online hypermedia course document. IEEE 1-4244-0406-1, 2006
- Bransford, J. D., Brown, A. L. and Cocking, R. R. (eds.):, How People Learn. Brain, Mind, Experience, and School. Committee on Developments in the Science of Learning, Commission on Behavioral and Social Sciences and Education, National Research Council. National Academy Press, Washington, 1999.
- Broisin J., Vidal P., Sibilla M., A management framework for tracking user activities in a web-based learning environment based on a model driven approach. Dans: educational multimedia, hypermedia & telecommunications (EDMEDIA 2006), Orlando, AACE, juin2006.
- Brooke, A., Wei, D.: Architecture for annotation and Ontology based querying of semantic Web Resources. IEEE/WIC/ACM international Conference on Web Intelligence (2005)
- Brusilovsky, P. & Pesin, L. ISIS-Tutor: An adaptive hypertext learning environment. *JCKBSE'94, Japanese-CIS Symposium on knowledge-based software engineering*. Pereslavl-Zalesski, Russia, 83-87, 1994.
- Brusilovsky, P., & Anderson, J. ACT-R electronic bookshelf: An adaptive system for learning cognitive psychology on the Web. In H. Maurer, & R. G. Olson (Eds.), *Proceedings of WebNet'98, World Conference of the WWW, Internet, and Intranet*, November 7-12, 1998. Orlando, FL, AACE. - pp. 92-97, 1998a.
- Brusilovsky, P., Eklund, J., & Schwarz, E. Web-based education for all: A tool for developing adaptive courseware. *Computer Networks and ISDN Systems*, 30 (1-7), 291-300, 1998b.
- Brusilovsky P. Adaptive and Intelligent Technologies for Web-based Education. In C. Rollinger and C. Peylo (eds.), *Special Issue on Intelligent Systems and Teleteaching*, Künstliche Intelligenz, 4, pp.19-25, 1999

- Brusilovsky P., 'Adaptive Hypermedia', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 87-110, 2001
- Brusilovsky P., Peylo C. Adaptive and intelligent Web-based educational systems. In P. Brusilovsky and C. Peylo (eds.), *International Journal of Artificial Intelligence in Education* **13** (2-4), Special Issue on Adaptive and Intelligent Web-based Educational Systems, pp.159-172, 2003a
- Brusilovsky P. Adaptive navigation support in educational hypermedia: The role of student knowledge level and the case for meta-adaptation. *British Journal of Educational Technology*, **34** (4), pp.487-497, 2003b
- Brusilovsky P., Vassileva J. Course sequencing techniques for large-scale web-based education. *International Journal of Continuing Engineering Education and Lifelong Learning* **13** (1-2), 75-94, 2003c.
- Brusilovsky P.. Authoring Tools for Advanced Technology Learning Environments, *chapter Developing Adaptive Educational Hypermedia Systems: from Design Models to Authoring Tools*, pages 377{409. Kluwer Academic Publishers, 2003.
- Brusilovsky P. Adaptive navigation support: From adaptive hypermedia to the adaptive Web and beyond. *Psychology Journal*, volume 2 (Number 1), pp.7-23, 2004a
- Brusilovsky P., Tasso C., 'Preface to Special Issue on User Modeling for Web Information Retrieval', *User Modeling and User-Adapted Interaction*, 14(2 – 3), 147 – 157, 2004b
- Brusilovsky P, KnowledgeTree: A Distributed Architecture for Adaptive E-Learning, WWW 2004, New York, New York, USA. ACM 1-58113-912-8/04/0005, 2004c.
- Brusilovsky, P., Chavan, G., Farzan, R.: Social adaptive navigation support for open corpus electronic textbooks. Adaptive Hypermedia, Eindhoven 2004d
- Buell D., An analysis of some fuzzy subset applications to information retrieval systems. *Fuzzy Sets and Systems* 7, 35-42, 1982
- Burton R. and Brown J. A tutoring and student modelling paradigm for gaming environments, *ACM SIGCSE bulletin* 8(1) : 236-246, 1978
- Burton J. Diagnosing bugs in a simple procedural skill, *Intelligent tutoring systems*, Academic press, 1982
- Cantador I, Castells P, Multilayered Semantic Social Network Modeling by Ontology-Based User Profiles Clustering: Application to Collaborative Filtering, EKAW 2006, LNAI 4248, 334-349, 2006
- Carberry S., 'Techniques for Plan Recognition', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 31 – 48, 2001

- Carro, R.M., Pulido, E. & Rodríguez, P. (1999). Task-based Adaptive learner Guidance On the WWW: the TANGOW System. *Second Workshop on Adaptive Systems and User Modeling on the Web, WWW'99*, Toronto, Canada, May 11-14, Computer Science Report 99-07, Eindhoven University Technology, (pp. 49-57), 1999.
- Castello Branko Neto Wilson and Alvaro Ostuni Gauthier Fernando Sharing and Reusing Information on Web-Based Learning, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), June 21-23 2006, Dublin, Ireland
- Castillo G., Gamma J., Breda A. (2003). Adaptive bayes for a student modelling prediction task based on learning styles, *user modelling 2003, 9th International Conference*, Springer Verlag, 2003.
- Castillo Gladys, Gama João, Breda Ana M. An Adaptive Predictive Model for Student Modeling, book *Adaptable and Adaptive Hypermedia Systems*, Eds: S. Chen and G.D.Magoulas, Idea Group Inc, 2004
- Catteau O., Vidal P., Marquié D., Broisin J., Production et gestion collaborative d'objets pédagogiques dans le cadre d'un dispositif international de FOAD, distance et savoir, Lavoisier, V. 5, 2007
- Chebotko, A., Deng, Y., Lu, S., Fotouhi, F.: OntoElan: an Ontology based Linguistic Multimedia Annotator. *IEEE International Symposium on Multimedia Software Engineering 2004*
- Chen, S. Y., and R. D. Macredie, Cognitive styles and hypermedia Navigation : development of a learning model, *Journal of the American society for information science and technology*, 53(1), 3-15, 2002.
- Cisco system inc., reusable learning object strategy, definition, creation process and guideline for building, 2000
- Clifford, R. Adaptive Hypermedia for Music Instruction, 7th International Technological Directions in Music Learning Conference, TDML ejournal, retrieved August 11, 2004 from <http://music.utsa.edu/tdml/conf-VII/VII-Clifford/VII-Clifford.html>, 2000
- Conati, C., A. S. Gertner, K. Vanlehn, M. J. Druzdel, On-line student modelling for coached problem solving using Bayesian networks, *International Conference on User modelling*, Chia Laguna (Italy), 231-242, 1997.
- Conati C., Gertner A., Vanlehn K., 'Using Bayesian Networks to Manage Uncertainty in Student Modeling', *User Modeling and User-Adapted Interaction*, 12(4), 371 – 417, 2002
- Corno, L., and R. E. Snow, Adapting teaching to individual differences among learners. In Wittrock, M (Ed.), *Handbook of research on teaching*. NY: Macmillan, 1986.
- Cornuéjols A., Miclet L., Kodratoff Y., Apprentissage artificiel, Eyrolles, 2002. Chapitres 2 et 9.

- Craswell N., Robertson S., Zaragoza H., Taylor M., Relevance weighting for query independent evidence. In : SIGIR 05: Proc. Of the 28th int. conf. on Research and development in information retrieval, New York, ACM Press, 416-423, 2005
- Cristea, A.I., and De Bra, P. Towards Adaptable and Adaptive ODL Environments. Proceedings of AACE E-Learn'02. Montreal, Canada, October. 232-239, 2002a
- Cristea, A., and L. Aroyo, Adaptive Authoring of Adaptive Educational Hypermedia, *International Conference on Adaptive Hypermedia*, Malaga, 122-132, 2002b
- Cristea A., Calvi L. The three Layers of Adaptation Granularity, *UM'03*, Pittsburgh, 2003a.
- Cristea, A.I., De Mooij, A., Adaptive Course Authoring: My Online Teacher, *In Proceedings of ICT'03, Papeete, French Polynesia, IEEE, IEE*, 2003b.
- Cristea, A.I., and Kinshuk Considerations on LAOS, LAG and their Integration in MOT. Proceedings of ED-MEDIA'03, Honolulu, Hawaii, AACE, 2003c.
- Cristea A. Adaptive Course Creation for All, *International Conference on Information Technology: Coding and Computing (ITCC'04) Volume 1*, pp. 718, 2004a.
- Cristea A., Verschoor M. The LAG Grammar for Authoring the Adaptive Web, *International Conference on Information Technology: Coding and Computing (ITCC'04) Volume 1*, pp. 382, 2004b.
- Cristea & F. Garzotto, ADAPT Major Design Dimensions for Educational Adaptive Hypermedia, *ED-MEDIA'04, AACE*, 2004c
- Cristea A, Evaluating Adaptive Hypermedia Authoring while Teaching Adaptive Systems, *SAC'04, ACM, Nicosia, Cyprus*, 2004d
- Cristea, A. IWhat can the Semantic Web do for Adaptive Educational Hypermedia? *Educational Technology & Society*, 7 (4), 40-58, 2004e.
- Cronbach L., Snow R. Aptitudes and Instructional Methods: A Handbook for Research on Interactions. *New York: Irvington Publishers*, 1977.
- Cullot N., Jouanot F., Yétongnon K., Une méthode de réconciliation sémantique pour l'extraction de connaissances, *RSTI – 17/2003. EGC2003*, pages 481 à 493, 2003
- Daconta, M. C, Obrst, L. J. & Smith K.T. *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*, Wiley, 2003.
- Y. Darcy, E. Monfrini, Y Guermeur, Borne "rayon-marge" sur l'erreur "leave-one-out" des SVM multi-classes. *JDS 2006*
- De Bra, P. M. E. Teaching Hypertext and Hypermedia through the Web. *Journal of Universal computer science*, 2(12), 797-804, 1996.

- De Bra, P. and Calvi, L. AHA! An open Adaptive Hypermedia Architecture. *The New Review of Hypermedia and Multimedia*, 4, Taylor Graham Publishers, 115-139, 1998.
- De Bra, P., & Ruiter, J.-P. AHA! Adaptive hypermedia for all. In W. Fowler, & J. Hasebrook (Eds.), *Proceedings of WebNet'2001*, World Conferen, 2001
- De Bra P., Aerts A., Berden B., De Lange B., Rousseau B., Santic T., Smits D., and Stash N.. AHA! the adaptive hypermedia architecture. *In HT'03: Proc. of the 14th ACM Conf. on Hypertext and Hypermedia*, pages 81-84. ACM Press, 2003a.
- De Bra, P., Santic, T., Brusilovsky, P. AHA! meets Interbook, and more..., *AACE ELearn 2003 Conference*, Phoenix, Arizona, November, 57-64, 2003b.
- De Landsheere V. et G., Définir les objectifs de l'éducation, P.U.F., Paris, 1989
- Dempster, A., Laird, N., Rubin, D., Maximum likelihood from incomplete data via EM algorithm. *Journal of the royal statistical society, series B*, 39(1):1-38, 1977.
- Desmoulins C. et Grandbastien M., Des ontologies pour la conception de manuels de formation à partir de documents techniques, *Sciences et Techniques Educatives*, vol. 9, n°3, 37-86, Hermès, Paris 2002
- Desmoulins C. et Grandbastien M., Environnements informatiques pour l'apprentissage humain, Chapitre 7, Cognition et traitement de l'information, Hermes Lavoisier, ISBN 2-7462-1171-8, mars 2006
- Dicheva, D. & Dichev, C. TM4L: Creating and Browsing Educational Topic Maps, *British Journal of Educational Technology - BJET*, 37(3), 391-404, 2006.
- Dicheva Darina and Dichev Christo Confronting Some Ontology-Building Problems in Educational Topic Map Authoring, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), Dublin, Ireland, June 21-23 2006b
- Dolog, P., Gavriiloaie, R., Nejdl, W. & Brase, J. Integrating Adaptive Hypermedia Techniques and Open RDF-Based Environments, WWW'03, Alternate Education Track, retrieved 12 August 2004 from <http://www2003.org/cdrom/papers/alternate/P810/p810-dolog.html>, 2003
- Downes S., Learning Objects : Resources for Distance Education Worldwide, *International Review of research in open and distance learning*, vol. 2, n°1, 2000
- Dron, J.: Social Software and the Emergence of Control. IEEE international Conference on Advanced Learning Technologies (2006)
- Dubois D. et Prade H., Criteria aggregation and ranking of alternatives in the framework of fuzzy set theory. In Zimmermann H. Zadeh L., Gaines B. (eds) *Fuzzy Sets and Decision analysis*. North-Holland, 209-240, 1984
- Dubois D. et Prade H., Semantic of quotient operators in fuzzy relational databases. *Fuzzy Sets and Systems* 78, 89-93, 1996

- Dufresne, A., and S. Turcotte, Cognitive style and its implications for navigation strategies, *World Conference on Artificial Intelligence and Education*, Amstredam, 1997.
- Dunn, R., How children learn: how learning styles impact student achievement, *Girl's fast-pitch softball and education magazine*, Lake Charles, Louisiana, 1(1), 5-8, 1999.
- Farah M. et Vanderpooten D., A multiple criteria approach for information retrieval, Crestani F., Ferragina P., Sanderson M. (Eds) : SPIRE, Springer-Verlag, LNCS 4209, pp. 242-254, 2006
- Felder R.M. and Silverman L.K., 'Learning and Teaching Styles in Engineering Education', *Engr. Education*, 78(7), 674-681, 1988.
- Felder R.M., 'Reaching the Second Tier: Learning and Teaching Styles in College Science Education', *J. College Science Teaching*, 23(5), 286-290, 1993.
- Felder, R., and B. A. Soloman, Learning styles and strategies. <http://www2.ncsu.edu/unity/lockers/users/ff/felder/public/ILSdir/styles.htm>, December, 2001.
- Felder R.M., Felder G.N., Dietz E.J. The Effects of Personality Type on Engineering Student Performance and Attitudes. *J. Engr. Education*, 91(1), pp.3-17, 2002.
- Felder R.M. and Brent R., 'Understanding Student Differences', *Journal of Engineering Education*, 94(1), 57-72, 2005a.
- Felder R.M. and Spurlin J., 'Applications, Reliability and Validity of the Index of Learning Styles', *Int.J. Engng Ed.*, 21(1), 103-112, 2005b.
- Fink J., Kobsa A., 'A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web', *User Modeling and User-Adapted Interaction*, 10(2 – 3), 209 – 249, 2000.
- Fischer G., 'User Modeling in Human–Computer Interaction', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 65 – 86, 2001.
- Fodor J., Yager R., Rybalov A., Structure of uni-norms. *International Journal of Uncertainty, Fuzzyness and Knowledge Based Systems*, 5, 411-427, 1997
- Foell, S., Boessling, P., Linner, D., Radush, I., Steglich, S.: Self organizing pervasive online communities. *IEEE International symposium on autonomous decentralized system*, 2007
- Fox E.A. et Show J.A., Combination of multiple searches In : Proc. Of TREC-3, Gaithersburg, Maryland, USA, NIST, 1994
- Frakes W. et Baeza-Yates R., *Information Retrieval : Data Structures and algorithms*. Prentice Hall, 1992.
- Francis L.J., Measuring attitudes towards computers among undergraduate college students: the affective domain. *Computers & Education*, 20(3), 251-255, 1993.

- Frasincar, F., Houben, G.-J. Peter Barna, P. & Cristian Pau, C. (). RDF/XML-based Automatic Generation of Adaptable Hypermedia Presentations, *International Conference on Information Technology: Computers and Communications*, April 28 - 30, IEEE, Las Vegas, Nevada, 410-414, 2003
- Frasincar F., Barna P., Houben G., Fiala Z. (2004). Adaptation and Reuse in Designing Web Information Systems, *International Conference on Information Technology: Coding and Computing (ITCC'04) Volume 1*, pp. 387, 2004.
- Gagné R.M. *Conditions of learning* (4 th ed.), New York : Holt, Rinehart and Winston, 1985
- Gamboa, H., and A. Fred, Designing intelligent tutoring systems : a bayesian approach, *ICEIS Artificial Intelligence and Decision Support Systems*, 452-458, 2001.
- Gandon, F.: Le Web sémantique n'est pas antisocial. Semaine de la connaissance. Nantes (2006)
- Gangemi Aldo, Catenacci Carola, Ciaramita Massimiliano, and Lehmann Jos, Modelling Ontology Evaluation and Validation, Y. Sure and J. Domingue (Eds.): *ESWC 2006*, LNCS 4011, pp. 140 – 154, 2006. © Springer-Verlag Berlin Heidelberg 2006
- Gardner H., *Frames of Mind: the theory of Multiple Intelligences*. Basic Books, New York, 1983
- Gardner H., *Multiples Intelligences: the theory in Practice*. Basic Books, New York, 1993
- Gardner H. *Intelligence Reframed: Multiple Intelligences for the 21st Century*, *Basic Books*, 1999.
- Gaudioso, E., Boticario, G. User Data Management and Usage Model Acquisition in an Adaptive Educational Collaborative Environment. *Proceedings of the second International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*. Spain, pp.143-152, , may 2002.
- Genest D. et Chein M., A content search information retrieval process based on conceptual graphs, *Knowledge and information systems*, Springer-Verlag, DOI 10.1007/s10115-004-0179-0, 8 : 292-309, 2004
- Gomez, J.M., Alor-Hernandez, G., Posada-Gomez, R., Abud-Figueroa, M., Garcia-Crespo, A.: SITIO: A Social Semantic Recommendation Platform. *IEEE conference on electronics, Communications and Computers 2007*
- Grandbastien M. et Labat J.M., *Environnements informatiques pour l'apprentissage humain*, Cognition et traitement de l'information, Hermes Lavoisier, ISBN 2-7462-1171-8, mars 2006
- Grandbastien M. et Labat J.M., *Environnements informatiques pour l'apprentissage humain*, Chapitre 9, Cognition et traitement de l'information, Hermes Lavoisier, ISBN 2-7462-1171-8, mars 2006
- Grasha A., *Teaching with Style : A practical guide*, *Pittsburgh: Alliance Publishers*, 1996.

- Grigoriadou, M., Papanikolaou, K., Kornilakis, H., & Magoulas, G. INSPIRE: An Intelligent System for Personalized Instruction in a Remote Environment. In P. D. Bra, P. Brusilovsky, & A. Kobsa (Eds.), *Proceedings of Third workshop on Adaptive Hypertext and Hypermedia*, Sonthofen, Germany, Technical University Eindhoven. - pp. 13-24, 2001.
- Grigorenko, E. L., and R. J. Stenberg, Thinking styles, In Sakloske and Zeidner (Eds.), *International handbook of personality and intelligence*, NY: Plenum, 205-230, 1995.
- Guermeur Y., Elisseeff A., Paugam-Moisy H., A new multi-class SVM based on a uniform convergence result, Neural Networks, IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on, Volume 4, 183 - 188 vol.4, 2000
- Guermeur Y., Large margin multi-category discriminant models and scale-sensitive Ψ - dimensions, Rapport de recherche n°5314, Systèmes biologiques, Projet MODBIO, INRIA, 2004
- Guermeur Y., Elisseeff A., Zelus D., A comparative study of multi-class support vector machines in the unifying framework of large margin classifiers. Applied stochastic models in business and industry, 21(2) : 199-214, 2005
- Habegger B., Quafafou M., Context generalization for information extraction from the web, proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'04), 2004
- Habel Genevieve, LaPalme Reyes Marie, Magnan Francois & Reyes Gonzalo General Poncelet meets the Semantic Web: A concrete example of the usage of ontologies to support creation and dissemination of eLearning contents, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), June 21-23 2006, Dublin, Ireland
- Hartley, H., Maximum likelihood estimation from incomplete data. Biometrics, 14:174-194, 1958.
- Henze, N., & Nejd, W. Student modeling for KBS Hyperbook system using Bayesian networks (Report # . *University of Hannover*, available online at <http://www.kbs.uni-hannover.de/paper/99/adaptivity.html>, 1999.
- Hibou M. et Labat J.M., Embedded Bayesian network student models, conference ITHET04, Istanbul
- Holt P., Dubs S., Jones M. and Greer J. The state of student modelling, *Student modelling : the Key to individualized knowledge-based instruction*, 3-38, 1994
- Honey P., Mumford A., The Manual of Learning Styles, 3 rd Ed., *Maidenhead*, Peter Honey, 1992.
- Höök, K., Rudström, Å. & Waern, A. Edited Adaptive Hypermedia: Combining Human and Machine Intelligence to Achieve Filtered Information, 8th ACM International Hypertext Conference (Hypertext'97), Flexible Hypertext Workshop, <http://www.sics.se/~kia/papers/edinfo.html>, 1997

- Horvitz, E., J. Breese, D. Heckerman, D. Hovel, and R. Rommelse, The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users, *Conference on Uncertainty in Artificial Intelligence*, Madison, Morgan Kaufmann: San Francisco, 256-265, 1998.
- Hu, H., Du, X.: ConAnnotator: Ontology Aided Collaborative Annotation System. IEEE International Conference on Computer Supported Cooperative Work I Design 2006
- Hui Min Lee C., Cheng Y.W., Rai S., Depickere A. What affect student cognitive style in the development of hypermédia learning system ?, *computer & education*, Elsevier, 2004.
- Jacquot, C., Bourda, Y. & Popineau, F. (). Reusability in GEAHs, International Workshop on Adaptive Hypermedia and Collaborative Web-based Systems (AHCW'04), International Conference on Web Engineering (ICWE 2004), July 26-27, Munich, Germany, retrieved 12 August 2004 from <http://www.ii.uam.es/~rcarro/AHCW04/Jacquot.pdf>., 2004
- Jensen, Finn V. An Introduction to Bayesian Networks, Springer-Verlag, New York, 1996
- Jonassen, D. H., and B. L. Grabowski, Handbook of individual differences, learning, and instruction, Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.
- Kabel S., Riemersma J., Wielinga B. Second didactic ontology, rapport final, communauté Européennes ESPRIT contrat IMAT, 2001
- Kass R. Student Modelling in intelligent tutoring systems-implications for user modelling, *User Models in Dialog Systems*, Berlin, Springer-Verlag, pp.386-410, 1989.
- Katz, S., A. Lesgold, G. Eggan, M. Gordan, Modelling the student in SHERLOCK II, *Journal of Artificial Intelligence and Education*, 3(4), 495-518, 1993.
- Kay J., Stereotypes, Student models and Scrutability, *Proceedings of fifth International Conference on Intelligent Tutoring Systems*, Springer-Verlag, 19-30, 2000a.
- Kay J., Accretion representation for scrutable student modelling, *Proceedings of fifth International Conference on Intelligent Tutoring Systems*, Springer-Verlag, 292 – 303, 2000b.
- Kay J., 'Learner Control', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 111 – 127, 2001.
- Kelly D., Tangney B., Empirical evaluation on adaptive multiple intelligence based tutoring system, Third International Conference on Adaptive Hypermedia and Adaptive Web Based System, AH 04, Eindhoven, Netherlands, 308-311, 2004a
- Kelly D., Tangney B., Predicting learning characteristics in a multiple intelligence based tutoring system, International conference on intelligent tutoring system, ITS 04, Maceio, Brazil, 679-688, 2004b

- Klamma, R., Chatti, M.A., Duval, E., Fiedler, S., Hummel, H. & al.: Social Software for Professional Learning: Examples and Research Issues. IEEE International conference on Advance Learning Technologies 2006
- Klyne G., Reynolds F., Woodrow C., Hidetaka O., Hjelm J., Butler M.H., Tran L. (2003). Composite capability/preference profiles (cc/pp):Structure and vocabularies. *W3C Working Draft* 25 March 2003.
- Kobsa, A., Generic User Modeling Systems, *User Modeling and User-Adapted Interaction*, 11, 49- 63, 2001.
- Kolb D.A., LSI Learning-Style Inventory, *McBer & Company*, 1985.
- Koper R., Modeling units of study from a pedagogical perspective. The pedagogical meta-model behind EML, Open university of the Netherlands, <http://eml.ou.nl/introduction/docs/ped-metamodel.pdf>, 2001
- Kostelnik R. and Bielikova M.. Web-based environment using adapted sequences of programming exercises. In M. Bene_s, editor, Proc. of Int. Conf. On Information Systems Implementation and Modelling, pages 33-40, 2003.
- Kraft D., Bordogna G., Pasi G., Fuzzi set techniques in information retrieval. In : Fuzzi Sets in Approximate Reasoning and information systems. Kluwer Academic Publisher, 469-510, 1999
- Kristofic Andrej, Bielikov Maria, Improving Adaptation in WebBased Educational Hypermedia by means of Knowledge Discovery, HT'05 ACM 1595931686/05/0009, Salzburg, Austria September 6–9, 2005, 2005
- Küpper D., Kobsa A., 'Tailoring the Presentation of Plans to Users' Knowledge and Capabilites'. In: A. Günter, R. Kruse, B. Neumann, eds.: *KI 2003: Advances in Artificial Intelligence, 26th Annual German Conference on Artificial Intelligence. Springer Verlag (LNCS 2821)*, Heidelberg, Germany, 606-617, 2003.
- Labat J.M., Pernin J.P., Gueraud V., Environnements informatiques pour l'apprentissage humain, Chapitre 3, Cognition et traitement de l'information, Hermes Lavoisier, ISBN 2-7462-1171-8, mars 2006
- Lauer F. et Bloch G., Méthodes SVM pour l'identification, in Journées Identification et Modélisation Expérimentale, JIME'2006, Poitiers, France Journées Identification et Modélisation Expérimentale (JIME'2006), Poitiers : France 2006
- Philippe Leray et Olivier François. Bayesian network structural learning and incomplete data. In in the proceedings of the International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning (AKRR 2005), Espoo, Finland, 2005.
- Lin T., Kinshuk, Ashok Patel. Cognitive Trait Model for Persistent Student Modelling, *EdMedia 2003 Conference Proceedings*, Norfolk, USA, 2003.

- MacGregor, S. K., Hypermedia navigation profiles: cognitive characteristics and information processing strategies, *Journal of Educational Psychology*, 15(3), 222-241, 1999.
- Manber, Udi; Patel, Ash; Robison, John: Experience with Personalization on Yahoo!, *Zeitschrift Communications - The ACM, Volume 43*, Nummer 8, August 2000
- Maneeawattthana, T., Wills, G. & Hall, W., Adaptive Link Services for the Semantic Web, EKAW'04, Northamptonshire, UK, Springer, retrieved from <http://www.ecs.soton.ac.uk/~tm03r/documents/EKAW04-Thanyalak.pdf>, 12 August 2004.
- Marinilli, M., Micarelli, A. and Sciarrone, F., A case-based approach to adaptive information filtering for the WWW *Proceedings of Second Workshop on Adaptive Systems and User Modeling on the WorldWideWeb*, Toronto and Banff, Canada. Computer Science Report 99-07, Eindhoven University of Technology, pp. 81-87, 1999.
- Mayer R.E., Multimedia learning, Cambridge University Press, New York, 2001
- Mayer R.E. The promise of multimedia learning : using the same instructional design methods across different media, *Learning and instruction 13 (2003) 125-139*, Elsevier Science, 2003.
- Mayo, M., and A. Mitrovic, Using a probabilistic student model to control problem difficulty, *International Conference on Intelligent Tutoring Systems*, Berlin, 524-533, 2000.
- Mayo M., Mitrovic A. Optimising ITS Behaviour with Bayesian Networks and decision Theory, *International Journal of Artificial Intelligence in Education*, 12, pp.124-153, 2001.
- Mayo M. Bayesian student modelling and decision theoretic selection of tutorial actions in intelligent tutoring systems, *thesis*, 2001.
- McLachlan, G. et Krishnan T., the EM algorithm and extensions. Wiley series in probability and statistics. Jhon Wiley & Sons, 1997.
- McLoughlin C., Hutchinson H., Koplín M. (2002). Different Media for Language Learning: Does Technology Add Quality?, *International Conference on Computers in Education (ICCE'02)*, pp. 681, December 2002.
- McLoughlin C. The implications of the research literature on learning styles for the design of instructional material, *Australian Journal of Educational Technology*, 15(3), pp.222-241, 1999.
- Mehta B., Niederee C., Stewart A., Deggemis M., Lops P., Semeraro G. Ontologically-Enriched Unified User Modeling for Cross-System Personalization, L. Ardissono, P. Brna, and A. Mitrovic (Eds): UM 2005, LNAI 3538, pp. 119-123, 2005.

- Mercer J.. Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 209 :415_446, 1909.
- Merrill M.D., A peddle in the pond : Model for instructional Design, Performance improvement, vol. 41, n°7, 39-44, août 2002.
- Messina Enza, Toscani Daniele, and Archetti Francesco, UP-DRES : User Profiling for a Dynamic REcommendation System, P. Perner (Ed.): *ICDM 2006, LNAI 4065*, pp. 146 – 160, 2006. © Springer-Verlag Berlin Heidelberg 2006
- Millán E, Pérez-de-la-Cruz JL. And Suarez E. Adaptive Bayesian networks for multilevel student modelling, *Proceedings of fifth International Conference on Intelligent Tutoring Systems*, Springer-Verlag, 534 – 543, 2000.
- Millán E., Pérez-de-la-Cruz JL., ‘A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation’, *User Modeling and User-Adapted Interaction*, 12(2 – 3), 281 – 330, 2002.
- Miller G.A., The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychological review*, 63, 81-97, 1956
- Miller M. Using learning styles to evaluate computer-based instruction, *Computers in Human Behavior*, Elsevier, 2004.
- Minka, T., Expectation-Maximization as lower bound maximization. Tutorial published on the web at <http://www.white.media.mit.edu/tp-minka/papers/em.html>, 1998
- Mizoguchi, R. Ontology Engineering Environments. *Handbook on Ontologies*, S. Staab, R. Studer (Eds.). *International Handbooks on Information Systems*, Springer, ISBN 3-540-40834-7, 275-298, 2004.
- Moallem M. Applying Constructivist and Objectivist Learning Theories in the Design of a Web-Based Course: Implications for Practice, *Educational Technology & Society* 4 (3), 2001.
- Moebis S., Piombo C., Batatia H. and Weibelzahl S., A Tool Set Combining Learning Styles Prediction, a Blended Learning Methodology and Facilitator Guidebooks – Towards a best mix in blended learning, *ICL2007, Austria, September 2007*
- Moellem M., The implications of research literature on learning styles for the design and development of a web-based course, *IEEE Intern. Conference on Computer in Education*, Auckland (NZ), 2002.
- Morales Rafael, van Labeke Nicolas and Paul Brna Towards a Learner Modelling Engine for the Semantic Web, *Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06)*, June 21-23 2006, Dublin, Ireland
- Murray T. Authoring Intelligent Tutoring Systems: An analysis of the state of the art, *International Journal of Artificial Intelligence in Education*, 10, 98-129, 1999.

- Neal, R., et Hinton, G., A view of the EM algorithm that justifies incremental, sparse and other variants. In Jordan, M., editor, *learning in Graphical Models*. Kluwer Academic Press, 1998.
- Neapolitan, Richard E. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*, John Wiley & Sons, New York, 1990
- Newell A., The knowledge level, *AI Mag.* 2, 1-20, 1981
- Nodenot T., Gaiot M., Bessagnet M.N., Sallaberry C., Creating rich collaborative learning scenarios: a model driven approach for contextualizing software components and documents, WCCE, 2005
- Noriaki, K.: Latent Semantic Indexing Based on Factor Analysis. *Soft Computing Systems - Design, Management and Applications*, Chile (2002) <http://iiscs.wssu.edu/o4e/viewhome.do?tm=O4E.xtm>
- Ounnas Asma, Liccardi Ilaria, Davis Hugh, Millard David and White Su Towards a Semantic Modeling of Learners for Social Networks, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), June 21-23 2006, Dublin, Ireland
- Papanikolaou K., Grigoriadou M., Kornilakis H., Magoulas G., 'Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE', *User Modeling and User-Adapted Interaction*, 13(3), 213 – 267, 2003.
- Paquette G., *L'ingénierie du télé-apprentissage, pour construire l'apprentissage en réseaux*, Presses de l'Université du Québec, 2002
- Paquette G. et Tchounikine P., Contribution à l'ingénierie des systèmes conseillers : une approche méthodologique fondée sur l'analyse du modèle de la tâche, *Revue Sciences et Techniques Educatives*, Hermès, vol. 9,n°3-4, 157-184, 2002
- Paredes, P., and P. Rodriguez, Considering sensing-intuitive dimension to exposition-exemplification in adaptive sequencing. *International Conference on Adaptive Hypermedia*, Malaga (Spain), 556-559, 2002.
- Parkinson A., Redmond J. Do cognitive styles affect learning performance in different computer media ?, *ACM SIGCSE Bulletin*, *Proceedings of the 7th annual conference on Innovation and technology in computer science education*, Volume 34 Issue 3, June 2002.
- Pasin Michele and Dzbor Martin A Task Based Approach to Support Situating Learning for the Semantic Web, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), June 21-23 2006, Dublin, Ireland
- Pastre P., L'analyse du travail en didactique professionnelle, *Revue Française de Pédagogie*, 138, 9-17, 2002

- Pazienza Maria Teresa and Stellato Armando, an Environment for Semi-automatic Annotation of Ontological Knowledge with Linguistic Content, Y. Sure and J. Domingue (Eds.): ESWC 2006, LNCS 4011, pp. 442 – 456, 2006. © Springer-Verlag Berlin Heidelberg 2006
- Peachey D. et McCalla G.I., Using planning techniques in intelligent tutoring system, *International journal of man-machine studies*, vol 24, 77-98, 1986
- Pearl J., *Probabilistic Reasoning in Expert Systems : Networks of Plausible Inference*. San Francisco: Morgan Kaufmann publishers, inc., 1988
- Pernici B., Locatelli P., Marinoni C., The eCCO Sytem : An eCompetence Management Tool Based on Semantic Networks, OTM Workshops 2006, LNCS 4278, 1088-1099, 2006.
- Pernin J.P. et Lejeune A., *Modèle pour la réutilisation de scénarios d'apprentissage*, colloque TICE Méditerranée, Nice, novembre 2004
- Pierrakos D., Paliouras G., Papatheodorou C., Spyropoulos C., 'Web Usage Mining as a Tool for Personalization: A Survey', *User Modeling and User-Adapted Interaction*, Volume 13 (4), 311 – 372, 2003.
- Piombo, C., Batatia, H., Ayache, A. (2003). A framework for adapting instruction to cognitive learning styles, *ICALT2003*, Athènes, july 2003
- Piombo C., Batatia H., Ayache A. 'Réseau bayésien pour la modélisation de la dépendance entre complexité de la tâche, style d'apprentissage et approche pédagogique', *3rd International Conference : Science of Electronic, Technologies of Information and Telecommunications*, Tunisia, 2005.
- Piombo, C., Batatia, H., Dayre, P., Ayache, A., An ontology based Web Annotation System to create new learning practices, *EC-TEL07*, Crete – Greece, September 2007
- Prade H. et Testemale C., Application of possibility and necessity measures to documentary information retrieval, LNCS 286, 265-275, 1987
- Quafafou M., Naouali S., Knowledge datawarehouse:Web usage OLAP Application, *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligenece (WI'05)*, 2005
- Ramoni M. et Sebastiani P., *Robust Learning with Missing Data*, Machine Learning, Kluwer Academic Publishers, 45, 147-170, 2001
- Razek, M. A., Frasson, C., Kaltenbach, M., Using a Machine Learning Approach Support an Intelligent Cooperative Multi-Agent System, *Symposium TICE'02*, Lyon (France), 119-123, 2002.
- Razmerita, L., Angehrn, A., Maedche, A., *Ontology-Based Modeling for Knowledge Management System*, UM 2003, LNAI 2702, 213-217, 2003

- Recker, M.M., Walker, A., Wiley, D.A.: An interface for collaborative filtering of educational resources. International Conference on Artificial intelligence, Las Vegas, Nevada, USA (2000)
- Redmond J., Walsh C., Parkinson A. Equilibrating instructional media for cognitive styles, *ACM SIGCSE Bulletin*, *Proceedings of the 8th annual conference on Innovation and technology in computer science education*, Volume 35 Issue 3, juin 2003.
- Riding, R., and S. Rayner. Cognitive styles and learning strategies: understanding differences in learning and behaviour, *London: David Fulton Pub.*, 1998.
- Riding, R., and M. Grimely, Cognitive style and learning from multimedia materials in 11-year children, *British Journal of Educational Technology*, 30(1), 43-59, 1999.
- Riding R., Rayner, S. *Cognitive Styles and Learning Strategies*, David Fulton Publishers London, 2001.
- Robertson S.E., The probability ranking principle in IR. *Journal of Documentation*, 33(1977), 294-304
- Robertson S.E., Walker S., Okapi-keenbow at TREC-8. In: Proc. 8th Text Retrieval Conf., TREC-8, 60-67, 1999
- Rodriguez, M.A., Steinbock, D.J., Watkins, J.H., Gershenson, C., Bollen, J., Grey, V., Degraf, B.: Smartocracy: Social networks for collective decision making. Hawaii international conference on system sciences (2007)
- Ross, P., J. Jones, and M. Millington, User modeling in Intelligent Teaching and Tutoring. In R. Lewis, R. Tagg. (Eds.), *Trends in computer assisted education*, Blackwell Scientific Publications, 32-44, 1987.
- Rumetshofer H., Wöß W. (2003). Digital contents in education: An approach for adaptable learning systems with respect to psychological aspects, *Proceedings of the 2003 ACM symposium on Applied computing*, March 2003.
- Salton G., Fox E.A., Wu H., Extended Boolean information retrieval. *Commun ACM* 26(11), 1022-1036, 1983
- Sarasin L. C. *Learning Style Perspectives: Impact in the Classroom*, Madison : Atwood Publishing, 1998.
- Schiaffino S., Amandi A., User profiling with Case-Based Reasoning and Bayesian Networks, international joint conference IBERAMIA-SBIA 2000 – Atibaia, Brazil, 12-21, 2000
- Scholkopf B. et Smola A.J., *Learning with kernels: support vector machines, regularization, optimization, and beyond*, the MIT Press, 2002

- Schuyten G., Dekeyser H.M., Preference for textual information and acting on support devices in multiple representations in a computer based learning environment for statistics, *Computers in human behaviour* 23, 2285-2301, Elsevier ScienceDirect, 2007
- Schwartz, D. G. From Open IS Semantics to the Semantic Web: The Road Ahead. *IEEE Intelligent Systems*, 52-58, 2003.
- Self, J., Cognitive diagnosis for tutoring systems, *Proceedings of the 10th European Conference on Artificial Intelligence, ECAI-92*, B. Neumann (_ed.), John Wiley and Sons Ltd, 1992, p. 699-703.
- Semeraro Giovanni, Lops Pasquale, and Degemmis Marco, Personalization for the Web: Learning User Preferences from Text, M. Hemmje et al. (Eds.): *E.J. Neuhold Festschrift, LNCS 3379*, pp. 162–172, 2005. _c Springer-Verlag Berlin Heidelberg 2005
- Shawe-Taylor J. et Cristianini N., *Support Vector Machines and other kernel-based learning methods*, Cambridge University Press, 2000
- Shearer B., *The MIDAS handbook of multiple intelligences in the classroom*, Greyden Press, Colombius, Ohio, 1996
- Shen, B., Yao, M., Wu, Z., Zhang, Y., Yi, W.: Ontology based association rules retrieval using protégé tools. *IEEE international conference on Data mining workshop 2006*
- Simard C., Normalisation de la formation en ligne. Enjeux, tendances et perspectives, AUF, BAN, NordSud.org, février 2002, profetic.org/file/norm-0210-d-RAPPORT.pdf
- W3C, SMIL, Synchronized Multimedia Language. <http://www.w3.org/AudioVideo/>
- School Mathematics Study Group, No 3. Z-Population Test Batteries. NLSMA reports, Leland Stanford Junior University, 1968.
- Sleeman, D.H., Assessing Competence in Basic Algebra, *Intelligent Tutoring Systems*, D.Sleeman et J.S.Brown (eds), Academic Press, New York, 1982, p. 185-200
- Smyth B.and Cotter P.. A Personalized TV Listings Service for the Digital TV Age. *Journal of Knowledge-Based Systems*, 13(2-3):53{59}, 2000.
- Snow R. Aptitude-Treatment Interaction as a framework on individual differences in learning. In P. Ackermann, R.J. Sternberg, & R. Glaser (eds.), *Learning and Individual Differences*. New York: W.H. Freeman, 1989.
- Song, R., Chen, E., Zhao, M., SVM Based Automatic User Profile Construction for Personalized Search, *ICIC 2005, Part I, LNCS 3644*, 475-484 , 2005
- Souto, M. Towards an Adaptive Web Training Environment Based on Cognitive Style of Learning : An Empirical Approach. *Proceedings of the second International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*. Spain, pp.338-347, may 2002

- Sosnovsky Sergey Integration of Two Domain Conceptualizations in a Single Adaptive Hypermedia System, Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@AH'06), June 21-23 2006, Dublin, Ireland
- Sparacino F., 'Sto(ry)chastics: A Bayesian Network Architecture for User Modeling and Computational Storytelling for Interactive Spaces', *Lecture Notes in Computer Science*, 2864, 54 – 72, 2003.
- Specht, M., & Klemke, R. ALE - Adaptive Learning Environment. In W. Fowler, & J. Hasebrook (Eds.), *Proceedings of WebNet'2001, World Conference of the WWW and Internet*, Orlando, FL, AACE. - pp. 1155-1160, 2001.
- Spiegelhalter, David J., A. Philip Dawid, Steffen L. Lauritzen and Robert G. Cowell "Bayesian analysis in expert systems" in *Statistical Science*, 8(3), 219-283, 1993
- Stach, N., Cristea, A.I. & P. De Bra, P. Authoring of Learning Styles in Adaptive Hypermedia: Problems and Solutions, WWW'04, 13th International World Wide Web Conference, May, New York, US, 104-113, 2004.
- Stephanidis C., 'Adaptive Techniques for Universal Access', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 159 – 179, 2001.
- Stewart, C., Cristea, A., Moore, A., Brailsford, T. & Ashman, H. Authoring and Delivering Adaptive Courseware, AH'04, *2nd International Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia*, <http://www.wis.win.tue.nl/~acristea/AH04/workshopAH.htm>, 2004.
- Sweller J., Cognitive load theory, learning difficulty and instructional design, learning and instruction, 4(4), 295-312, 1994
- Tanner, M., *Tools for statistical inference*. Springer Verlag, New York. Third edition, 1996.
- Taylor, W. L., Cloze Procedure: A new tool for measuring readability, *Journalists' Quarterly* 30, 415-433, 1953.
- Tchounikine P., Pour une ingénierie des environnements informatiques pour l'apprentissage humain, revue 13, vol. 2(1), 59-95, <http://ArchivesEiah.org>, 2002
- Tchounikine P., *Environnements informatiques pour l'apprentissage humain*, Chapitre 6, Cognition et traitement de l'information, Hermes Lavoisier, ISBN 2-7462-1171-8, mars 2006
- Tepper, M.: The rise of social software. *netWorker* (2003), 7(3):18-23
- Timothy J., Hafner W., Mitropoulos F. eCAD : A Knowledge-Based Course Engineering System, *proceedings of the 37th Hawaii International Conference on System Sciences*, 2004.
- Tombros A. et van Rijsbergen C.J., Query-Sensitive similarity measures for information retrieval, *Knowledge and Information Systems*, Springer-Verlag, DOI 10.1007/s10115-003-0115-8, 6 : 617-642, 2004

- Trella, M., Conejo, R., Bueno, D., Guzmán, E. An autonomous component architecture to develop WWW-ITS. In Brusilovsky, P., Henze, N., Milln, E. (eds.) Proceedings of the Workshops on Adaptive Systems for Web-Based Education. Malaga (this volume). <http://citeseer.ist.psu.edu/trella02autonomous.html>, 2002
- Triantafyllou E., Pomportsis A., Demetriadis S. The design and the formative evaluation of an adaptive educational system based on cognitive styles, *Computers & Education* 41 : 87-103, Elsevier, 2003. <http://www.up2uml.org/>
- Vassileva, J., A classification and synthesis of student modelling techniques in intelligent computer-assisted instruction, *International Conference on Computer Assisted Learning*, Hagen(Germany), 202-203, 1990.
- Vaughn, J., *Assessing Reading: Using cloze procedure to assess reading skills*. Adult Literacy and Basic Skills Unit, London, 1995.
- Virvou M. and Tsirgara V. Involving effectively teachers and students in the life cycle of an intelligent tutoring system, *Educational technology & society* 3(3) : 511-521, 2000
- W3C, RDF (Resource Description Framework). <http://www.w3.org/RDF/> Les cas et conditions d'utilisation du langage d'ontologie Web OWL, Recommandation du W3C du 10 février 2004, <http://www.w3.org/TR/2004/REC-webont-req-20040210/>
- Wærn A., 'User Involvement in Automatic Filtering: An Experimental Study', *User Modeling and User-Adapted Interaction*, 14(2 – 3), 201 – 237, 2004.
- Weber, G., & Specht, M.: "User modeling and adaptive navigation support in WWW-based tutoring systems"; In Jameson, A., Paris, C., & Tasso, C. (Eds.), *User Modeling*, Springer-Verlag, Wien 289-300, 1997.
- Webb G., Pazzani M., Billsus D., 'Machine Learning for User Modeling', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 19 – 29, 2001.
- Weber, G. & Brusilovsky, P. ELM-ART: An adaptive versatile system for Web-based instruction. *International Journal of Artificial Intelligence in Education* 12 (4), *Special Issue on Adaptive and Intelligent Web-based Educational Systems*, (pp. 351-384), 2001a.
- Weber, G., Kuhl H.-C., Weibelzahl, S. Developing adaptive internet based courses with the authoring system NetCoach. In REICH, S., TZAGARAKIS, M., and DE BRA, P., editors, *Hypermedia : Openness, Structural Awareness, and Adaptivity* : 226-238, Berlin, 2001b
- Weibelzahl, S., Weber, G. Adapting to Prior Knowledge of Learners. Proceedings of the second *International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*. Spain, pp.448-451, may 2002
- Wenger, E., *Artificial Intelligence and Tutoring Systems*, Morgan Kaufmann Publisher, Los Altos, California, 1987

- Winter M. and McCalla G. The emergence of student models from an analysis of ethical decision making in a scenario-based learning environment, *proceedings of th seventh international conference on user modeling (UM99)*, Springer-Verlag, 265-274, 1999
- Wondergem B., Van Bommel P., Van der Weide T., Nesting and defoliation of index expressions for information retrieval, *Knowledge and information systems*, Springer-Verlag, 2 : 33-52, 2000
- Yager R., On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18, 183-190, 1988
- Yuxin Mao, Zhaohui Wu, Huajun Chen, and Xiaoqing Zheng, Concept Map Model for Web Ontology Exploration, X. Zhou et al. (Eds.): *APWeb 2006*, LNCS 3841, pp. 544 – 555, 2006. © Springer-Verlag Berlin Heidelberg 2006
- Zakaria, M.R. & Brailsford, T.J. User Modelling and Adaptive Educational Hypermedia Frameworks for Education. *New Review of Hypermedia and Multimedia*, 8, 83-97, 2002.
- Zhang, L.F., Thinking styles: their relationship with modes of thinking and academic performance, *Educational Psychology*, 22(3), 331-348, 2002.
- Zhang Y., Liu Z. A Model of Web Oriented Intelligent Tutoring System for Distance Education, *Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'03)*, pp 78, 2003.
- Zhang L. Does teaching for a balanced use of thinking styles enhance students' achievement?, *personality and individual differences*, Elsevier, 2004.
- Zhou Y., and M. W. Evens, A practical student model in intelligent tutoring system, *IEEE International Conference on Tools with Artificial Intelligence*, Chicago, 13-18, 1999.
- Zukerman I., Albrecht D., 'Predictive Statistical Models for User Modeling', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 5-, 2001a.
- Zukerman I., Litman D., 'Natural Language Processing and User Modeling: Synergies and Limitations', *User Modeling and User-Adapted Interaction*, 11(1 – 2), 129 – 158, 2001b.

Annexes

Annexe 1

INDEX OF LEARNING STYLES*

Barbara A. Soloman
First-Year College
North Carolina State University
Raleigh, North Carolina 27695

Richard M. Felder
Department of Chemical Engineering
North Carolina State University
Raleigh, NC 27695-7905

DIRECTIONS

Circle "a" or "b" to indicate your answer to every question. Please choose only one answer for each question.

If both "a" and "b" seem to apply to you, choose the one that applies more frequently.

1. I understand something better after I
 - (a) try it out.
 - (b) think it through.
2. I would rather be considered
 - (a) realistic.
 - (b) innovative.
3. When I think about what I did yesterday, I am most likely to get
 - (a) a picture.
 - (b) words.
4. I tend to
 - (a) understand details of a subject but may be fuzzy about its overall structure.
 - (b) understand the overall structure but may be fuzzy about details.
5. When I am learning something new, it helps me to
 - (a) talk about it.
 - (b) think about it.
6. If I were a teacher, I would rather teach a course
 - (a) that deals with facts and real life situations.
 - (b) that deals with ideas and theories.
7. I prefer to get new information in
 - (a) pictures, diagrams, graphs, or maps.
 - (b) written directions or verbal information.
8. Once I understand
 - (a) all the parts, I understand the whole thing.
 - (b) the whole thing, I see how the parts fit.
9. In a study group working on difficult material, I am more likely to
 - (a) jump in and contribute ideas.
 - (b) sit back and listen.

10. I find it easier
 - (a) to learn facts.
 - (b) to learn concepts.
11. In a book with lots of pictures and charts, I am likely to
 - (a) look over the pictures and charts carefully.
 - (b) focus on the written text.
12. When I solve math problems
 - (a) I usually work my way to the solutions one step at a time.
 - (b) I often just see the solutions but then have to struggle to figure out the steps to get to them.
13. In classes I have taken
 - (a) I have usually gotten to know many of the students.
 - (b) I have rarely gotten to know many of the students.
14. In reading nonfiction, I prefer
 - (a) something that teaches me new facts or tells me how to do something.
 - (b) something that gives me new ideas to think about.
15. I like teachers
 - (a) who put a lot of diagrams on the board.
 - (b) who spend a lot of time explaining.
16. When I'm analyzing a story or a novel
 - (a) I think of the incidents and try to put them together to figure out the themes.
 - (b) I just know what the themes are when I finish reading and then I have to go back and find the incidents that demonstrate them.
17. When I start a homework problem, I am more likely to
 - (a) start working on the solution immediately.
 - (b) try to fully understand the problem first.
18. I prefer the idea of
 - (a) certainty.
 - (b) theory.
19. I remember best
 - (a) what I see.
 - (b) what I hear.
20. It is more important to me that an instructor
 - (a) lay out the material in clear sequential steps.
 - (b) give me an overall picture and relate the material to other subjects.
21. I prefer to study
 - (a) in a study group.
 - (b) alone.
22. I am more likely to be considered
 - (a) careful about the details of my work.
 - (b) creative about how to do my work.
23. When I get directions to a new place, I prefer
 - (a) a map.
 - (b) written instructions.
24. I learn
 - (a) at a fairly regular pace. If I study hard, I'll "get it."
 - (b) in fits and starts. I'll be totally confused and then suddenly it all "clicks."
25. I would rather first
 - (a) try things out.
 - (b) think about how I'm going to do it.

-
26. When I am reading for enjoyment, I like writers to
 - (a) clearly say what they mean.
 - (b) say things in creative, interesting ways.
 27. When I see a diagram or sketch in class, I am most likely to remember
 - (a) the picture.
 - (b) what the instructor said about it.
 28. When considering a body of information, I am more likely to
 - (a) focus on details and miss the big picture.
 - (b) try to understand the big picture before getting into the details.
 29. I more easily remember
 - (a) something I have done.
 - (b) something I have thought a lot about.
 30. When I have to perform a task, I prefer to
 - (a) master one way of doing it.
 - (b) come up with new ways of doing it.
 31. When someone is showing me data, I prefer
 - (a) charts or graphs.
 - (b) text summarizing the results.
 32. When writing a paper, I am more likely to
 - (a) work on (think about or write) the beginning of the paper and progress forward.
 - (b) work on (think about or write) different parts of the paper and then order them.
 33. When I have to work on a group project, I first want to
 - (a) have "group brainstorming" where everyone contributes ideas.
 - (b) brainstorm individually and then come together as a group to compare ideas.
 34. I consider it higher praise to call someone
 - (a) sensible.
 - (b) imaginative.
 35. When I meet people at a party, I am more likely to remember
 - (a) what they looked like.
 - (b) what they said about themselves.
 36. When I am learning a new subject, I prefer to
 - (a) stay focused on that subject, learning as much about it as I can.
 - (b) try to make connections between that subject and related subjects.
 37. I am more likely to be considered
 - (a) outgoing.
 - (b) reserved.
 38. I prefer courses that emphasize
 - (a) concrete material (facts, data).
 - (b) abstract material (concepts, theories).
 39. For entertainment, I would rather
 - (a) watch television.
 - (b) read a book.
 40. Some teachers start their lectures with an outline of what they will cover. Such outlines are
 - (a) somewhat helpful to me.
 - (b) very helpful to me.
 41. The idea of doing homework in groups, with one grade for the entire group,
 - (a) appeals to me.
 - (b) does not appeal to me.
 42. When I am doing long calculations,
 - (a) I tend to repeat all my steps and check my work carefully.
 - (b) I find checking my work tiresome and have to force myself to do it.

43. I tend to picture places I have been
- (a) easily and fairly accurately.
 - (b) with difficulty and without much detail.
44. When solving problems in a group, I would be more likely to
- (a) think of the steps in the solution process.
 - (b) think of possible consequences or applications of the solution in a wide range of areas.

Annexe 2

Version Française du questionnaire ILS

Questionnaire adapté par Christophe Piombo (Doctorant INPT) du ILS de Richard M. Felder

Consignes

Entourez la réponse "a" ou "b" pour chacune des questions. *Veillez choisir une seule réponse pour chaque question.*

Si "a" et "b" semblent toutes les deux s'appliquer à vous, choisissez la réponse qui s'applique à vous le plus souvent.

Répondez le plus naturellement possible. Il n'y a pas de bonnes ou mauvaises réponses.

1. Pour comprendre la règle d'un jeu de société
 - (a) Je dois faire une partie.
 - (b) La lecture de la notice explicative me suffit.
2. Si je devais dessiner un arbre, j'aurai tendance à
 - (a) faire une reproduction qui soit la plus proche de la réalité.
 - (b) faire une reproduction avec une touche personnelle.
3. Si je pense à ce que j'ai fait hier, ce qui me vient à l'esprit ce sont
 - (a) des images.
 - (b) des mots.
4. Si j'étudie l'histoire de France
 - (a) je comprends les détails des évènements historiques, mais je reste dans le flou sur les périodes historiques.
 - (b) je comprends les périodes historiques, mais je reste dans le flou sur les détails des évènements historiques.
5. Quand j'étudie une nouvelle leçon, je comprends mieux
 - (a) si j'en parle avec ma famille.
 - (b) si j'y réfléchi seul dans ma chambre.
6. Si j'étais un(e) enseignant(e), j'enseignerais plutôt un cours
 - (a) sur les sciences de la vie et de la terre.
 - (b) sur la philosophie.
7. Lorsqu'un ami m'explique l'itinéraire pour se rendre chez lui
 - (a) je préfère qu'il me fasse un plan.
 - (b) je préfère qu'il me donne des instructions écrites ou verbales.

8. Lorsque j'étudie un texte en Français
 - (a) dès que je comprends tous les paragraphes, je comprends l'idée générale.
 - (b) dès que je comprends l'idée générale, je comprends mieux les relations entre les paragraphes.
9. Lors d'un travail de groupe au collège
 - (a) je participe activement et suggère des idées à mes camarades.
 - (b) j'ai tendance à écouter les idées proposées par mes camarades.
10. Si je dois apprendre la pluviométrie d'un pays,
 - (a) je préfère étudier un tableau de chiffres présentant les précipitations annuelles du pays.
 - (b) je préfère comparer le climat du pays avec les climats connus.
11. Dans une bande dessinée
 - (a) je me concentre sur l'image de la scène.
 - (b) je me concentre sur le texte des bulles.
12. Lorsque je résous des problèmes de mathématiques
 - (a) j'ai l'habitude d'élaborer la solution étape par étape.
 - (b) je vois souvent d'emblée la solution, mais j'éprouve des difficultés à trouver les étapes qui me mèneraient à cette solution.
13. A l'école,
 - (a) j'ai souvent eu beaucoup de copains (copines).
 - (b) j'ai rarement eu plus de 2 copains (copines).
14. je préfère lire
 - (a) un livre qui présente des techniques de bricolage ou de nouvelles recettes de cuisine.
 - (b) un livre qui parle de l'évolution future et hypothétique de la planète.
15. Pendant un cours
 - (a) je préfère que mon professeur dessine de nombreux graphiques et schémas au tableau.
 - (b) je préfère que mon professeur passe beaucoup de temps à expliquer verbalement.
16. Lorsque je lis une histoire ou un roman
 - (a) je réfléchis aux intrigues et j'essaie de les regrouper afin de comprendre le thème général.
 - (b) je comprends le thème général seulement lorsque j'ai terminé la lecture, et je suis obligé de retourner au texte pour retrouver les intrigues qui l'illustrent.
17. Lorsque je dois assembler les pièces d'un modèle réduit
 - (a) j'ouvre la boîte et commence l'assemblage des pièces.
 - (b) j'ouvre la boîte et parcours la notice explicative de montage.
18. Dans un cours de sciences de la vie et de la terre sur les mammifères
 - (a) je préfère étudier leur anatomie en faisant des mesures et des expérimentations.
 - (b) je préfère étudier la théorie de leur évolution.

-
19. Je me souviens mieux d'un conte
(a) si je le vois en dessin animé.
(b) si on me le raconte.
20. Il est plus important pour moi qu'un enseignant
(a) présente une matière étape par étape.
(b) présente une vue d'ensemble de la matière et la relie à d'autres sujets.
21. Quand il faut faire un exposé en classe,
(a) je préfère le préparer avec des camarades.
(b) je préfère le préparer tout(e) seul(e).
22. Les enseignants disent de moi que je suis
(a) minutieux(minutieuse), appliqué(e) dans mon travail.
(b) malin(e), débrouillard(e) dans la façon de faire mon travail.
23. Lorsqu'on me donne des indications pour aller à un nouvel endroit, je préfère
(a) un plan.
(b) des instructions écrites.
24. J'apprends
(a) à un rythme assez régulier. Si je travaille dur, j'y arrive.
(b) par à-coups. Je peux être complètement embrouillé, et puis soudainement avoir un déclic.
25. La première fois qu'un(e) ami(e) m'invite chez lui
(a) l'heure venue, je pars de chez moi et essaie de trouver le chemin.
(b) quelques minutes avant le départ, je réfléchis au chemin en utilisant un plan.
26. Lorsque je lis pour mon plaisir, j'aime les auteurs
(a) qui expriment clairement, simplement, nettement ce qu'ils veulent dire.
(b) qui expriment les choses d'une façon créative, ingénieuse, intéressante.
27. Lorsque je vois un diagramme ou un croquis en classe, j'ai tendance à mieux me rappeler
(a) de l'image.
(b) du commentaire de l'enseignant.
28. Lorsque je dois faire la somme des pièces de monnaie contenues dans ma tirelire, je préfère
(a) commencer à faire des tas par type de pièces et calculer après.
(b) prendre les pièces les une après les autres, comme elles viennent, pour en faire la somme.
29. Au bout de quelques jours, je me souviens de la règle d'un jeu
(a) seulement si j'y ai joué.
(b) même si je n'ai fait que lire la notice explicative.
30. Lorsque je dois réaliser une tâche (exemples : faire un gâteau, assembler des legos...), je préfère
(a) maîtriser une façon de faire.
(b) inventer de nouvelles façons de faire.

31. Lorsque le professeur d'Histoire Géographie nous présente des informations sur les Capitales Européennes
- (a) je préfère voir des schémas ou des graphes.
 - (b) je préfère lire un texte qui résume ces informations.
32. Lorsque j'écris une histoire, j'ai tendance à
- (a) travailler (réfléchir ou écrire) à l'introduction puis développer les parties suivantes.
 - (b) travailler (réfléchir ou écrire) aux différentes parties de l'histoire puis essayer de les assembler.
33. Lorsque je dois préparer un exposé en groupe,
- (a) je préfère que chacun apporte ses idées dès le départ.
 - (b) je préfère réfléchir seul(e) et ensuite comparer les idées en groupe.
34. Je considère comme une qualité d'être
- (a) prudent(e), calme et appliqué(e).
 - (b) imaginatif (imaginative), vif (vive), extravagant(e).
35. De retour de vacances, lorsque je me souviens de mes nouveaux (nouvelles) amis (amies), j'ai tendance à me rappeler
- (a) de quoi ils (elles) avaient l'air.
 - (b) de ce qu'ils (elles) disaient d'eux-mêmes (elles-mêmes).
36. Lorsque je fais un devoir de Mathématiques,
- (a) je me concentre sur l'écriture des formules pour avoir tout juste.
 - (b) je pense aux règles de grammaire du français pour mieux expliquer ma démarche.
37. Lorsque je suis invité à un anniversaire
- (a) je vais facilement vers des personnes qui me sont inconnues.
 - (b) je reste souvent avec mes copains (copines).
38. Je préfère les cours qui mettent l'accent sur
- (a) l'expérimentation afin de mesurer et analyser des données.
 - (b) des activités intellectuelles.
39. Pour me divertir, je préfère
- (a) regarder la télévision.
 - (b) lire un livre.
40. Lorsque un professeur commence l'heure de cours par un plan de ce qu'il va nous présenter
- (a) cela m'aide peu à comprendre les différentes parties du sujet étudié.
 - (b) cela m'aide beaucoup à comprendre les différentes parties du sujet étudié.
41. Lorsque nous réalisons une expérience de physique en groupe
- (a) je préfère que tous les membres du groupe aient la même note.
 - (b) je préfère que chaque membre du groupe puisse avoir une note individuelle.

42. Lorsque je fais de longs calculs,
- (a) j'ai tendance à répéter toutes les étapes afin de vérifier soigneusement mes calculs.
 - (b) je trouve ennuyeux de devoir vérifier mes calculs et je dois me forcer à le faire.
43. J'ai tendance à me représenter le lieu de mes dernières vacances
- (a) avec facilité et exactitude.
 - (b) avec difficulté et peu de détail.
44. Si je construis des voitures en legos avec mes copains (copines),
- (a) je réfléchis à la manière de construire les voitures.
 - (b) je réfléchis à la manière de jouer avec les voitures construites.

Annexe 3

Questionnaire adapté par Christophe Piombo (Doctorant INPT) du ILS de Richard M. Felder

Consignes

Entourez la réponse "a" ou "b" pour chacune des questions. *Veillez choisir une seule réponse pour chaque question.*

Si "a" et "b" semblent toutes les deux s'appliquer à vous, choisissez la réponse qui s'applique à vous le plus souvent.

Répondez le plus naturellement possible. Il n'y a pas de bonnes ou mauvaises réponses.

45. Pour comprendre la règle d'un jeu de société
(a) Je dois faire une partie.
(b) La lecture de la notice explicative me suffit.
46. Si j'étais un(e) enseignant(e), j'enseignerais plutôt un cours
(a) sur les sciences de la vie et de la terre.
(b) sur la philosophie.
47. Lorsqu'un ami m'explique l'itinéraire pour se rendre chez lui
(a) je préfère qu'il me fasse un plan.
(b) je préfère qu'il me donne des instructions écrites ou verbales.
48. Lorsque je résous des problèmes de mathématiques
(a) j'ai l'habitude d'élaborer la solution étape par étape.
(b) je vois souvent d'emblée la solution, mais j'éprouve des difficultés à trouver les étapes qui me mèneraient à cette solution.
49. La première fois qu'un(e) ami(e) m'invite chez lui
(a) l'heure venue, je pars de chez moi et essaie de trouver le chemin.
(b) quelques minutes avant le départ, je réfléchis au chemin en utilisant un plan.
50. Dans un cours de sciences de la vie et de la terre sur les mammifères
(a) je préfère étudier leur anatomie en faisant des mesures et des expérimentations.
(b) je préfère étudier la théorie de leur évolution.
51. Dans une bande dessinée
(a) je me concentre sur l'image de la scène.
(b) je me concentre sur le texte des bulles.
52. J'apprends
(a) à un rythme assez régulier. Si je travaille dur, j'y arrive.
(b) par à-coups. Je peux être complètement embrouillé, et puis soudainement avoir un déclic.
53. Au bout de quelques jours, je me souviens de la règle d'un jeu
(a) seulement si j'y ai joué.
(b) même si je n'ai fait que lire la notice explicative.
54. Je préfère les cours qui mettent l'accent sur
(a) l'expérimentation afin de mesurer et analyser des données.
(b) des activités intellectuelles.
55. Lorsque le professeur d'Histoire Géographie nous présente des informations sur les Capitales Européennes

- (a) je préfère voir des schémas ou des graphes.
 (b) je préfère lire un texte qui résume ces informations.

56. Lorsque je fais un devoir de Mathématiques,

- (a) je me concentre sur l'écriture des formules pour avoir tout juste.
 (b) je pense aux règles de grammaire du français pour mieux expliquer ma démarche.

Dépouillement des résultats

Questions	Nombre de a	Nombre de b	Entourez sur l'axe la valeur calculée (Nombre de b - Nombre de a)
Q1 Q5 Q9			<p>Actif Réfléchi</p> <p>-3 -1 +1 +3</p>
Q2 Q6 Q10			<p>Inductif Déductif</p> <p>-3 -1 +1 +3</p>
Q3 Q7 Q11			<p>Visuel Verbal</p> <p>-3 -1 +1 +3</p>
Q4 Q8 Q12			<p>Séquentiel Global</p> <p>-3 -1 +1 +3</p>

Annexe 4



Avenue de Karben 31520 Ramonville Saint-Agne Tel : 0562881350 Fax : 0562881351 Email : 0311633g@ac-toulouse.fr

Toulouse le 06 avril 2005

Lettre d'information aux parents des classes de cinquième

Madame, Monsieur,

Le collège Malraux de Ramonville est impliqué dans plusieurs projets innovants sur l'utilisation des nouvelles technologies et du multimédia.

Dans ce cadre, une expérimentation pédagogique sur les classes de cinquième, organisée en collaboration avec le Laboratoire d'Informatique et de Mathématiques Appliquées (LIMA) de l'école nationale supérieure d'électrotechnique, d'électronique, d'Informatique et d'Hydraulique de Toulouse (ENSEEIH) aura lieu le 14 avril 2005 dans les locaux du collège. Vous trouverez, ci-joint le détail de cette expérimentation.

Merci de bien vouloir remplir et nous retourner le coupon réponse ci-dessous autorisant votre enfant à participer à cette expérimentation.

Monsieur Bégué

Principal du Collège Malraux

Coupon Réponse

(rayer la mention inutile)

- J'autorise
 Je n'autorise pas

ma fille - mon fils

_____ classe _____ à participer à l'expérimentation
 pédagogique qui aura lieu le 14 avril 2005 entre 8h30 et 14h30 dans les salles
 d'ordinateurs du Collège Malraux.

Les parents

Date et signatures

Annexe 5

Code OWL de la méta-ontologie du modèle de contenus

```

<?xml version="1.0"?>

<!DOCTYPE rdf:RDF [
  <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
  <!ENTITY owl11 "http://www.w3.org/2006/12/owl11#" >
  <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
  <!ENTITY owl11xml "http://www.w3.org/2006/12/owl11-xml#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <!ENTITY contentmodel
"http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#" >
]>

<rdf:RDF
xmlns="http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#"
  xml:base="http://www.up2uml.org/ontologies/2007/7/contentmodel.owl"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl11="http://www.w3.org/2006/12/owl11#"
  xmlns:owl11xml="http://www.w3.org/2006/12/owl11-xml#"

  xmlns:contentmodel="http://www.up2uml.org/ontologies/2007/7/contentmodel.
owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <owl:Ontology rdf:about=""/>

  <!--

  //////////////////////////////////////
  //////////////////////////////////////
  //
  // Object Properties
  //

  //////////////////////////////////////
  //////////////////////////////////////
  -->

  <!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#estAssocieA -->

  <owl:ObjectProperty rdf:about="#estAssocieA">
    <rdfs:range>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#estAssocieA"/>
        <owl11:onClass rdf:resource="#Sequence"/>
        <owl:minCardinality
rdf:datatype="&xsd;nonNegativeInteger">1</owl:minCardinality>
        </owl:Restriction>

```

```

    </rdfs:range>
    <rdfs:range rdf:resource="#Sequence"/>
    <rdfs:domain rdf:resource="#ObjectifFinal"/>
  </owl:ObjectProperty>

  <!--
  http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#estComposeDe -->

  <owl:ObjectProperty rdf:about="#estComposeDe">
    <rdfs:range rdf:resource="#Activite"/>
    <rdfs:range rdf:resource="#Tache"/>
    <rdfs:domain rdf:resource="#Tache"/>
    <rdfs:domain rdf:resource="#Sequence"/>
    <rdfs:range>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#estComposeDe"/>
        <owl11:onClass rdf:resource="#ObjectifFinal"/>
        <owl:minCardinality
rdf:datatype="&xsd;nonNegativeInteger">1</owl:minCardinality>
        </owl:Restriction>
      </rdfs:range>
    <rdfs:range>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#estComposeDe"/>
        <owl11:onClass rdf:resource="#Activite"/>
        <owl:minCardinality
rdf:datatype="&xsd;nonNegativeInteger">1</owl:minCardinality>
        </owl:Restriction>
      </rdfs:range>
    <rdfs:range rdf:resource="#ObjectifFinal"/>
    <rdfs:range>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#estComposeDe"/>
        <owl11:onClass rdf:resource="#ObjectifGlobal"/>
        <owl:minCardinality
rdf:datatype="&xsd;nonNegativeInteger">1</owl:minCardinality>
        </owl:Restriction>
      </rdfs:range>
    <rdfs:range rdf:resource="#ObjectifGlobal"/>
    <rdfs:domain rdf:resource="#ObjectifGlobal"/>
  </owl:ObjectProperty>

  <!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#utilise
  -->

  <owl:ObjectProperty rdf:about="#utilise">
    <rdfs:range>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#utilise"/>
        <owl11:onClass rdf:resource="#Ressource"/>
        <owl:minCardinality
rdf:datatype="&xsd;nonNegativeInteger">1</owl:minCardinality>
        </owl:Restriction>
      </rdfs:range>
    <rdfs:range rdf:resource="#Ressource"/>
    <rdfs:domain rdf:resource="#Activite"/>
  </owl:ObjectProperty>

```



```

<!--
////////////////////////////////////
////////////////////////////////////
//
// Data properties
//
////////////////////////////////////
////////////////////////////////////
-->

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#acteur
-->

<owl:DatatypeProperty rdf:about="#acteur">
  <rdfs:domain rdf:resource="#Activite"/>
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Tache"/>
</owl:DatatypeProperty>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#bloom -
->

<owl:DatatypeProperty rdf:about="#bloom">
  <rdfs:domain rdf:resource="#ObjectifFinal"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#description -->

<owl:DatatypeProperty rdf:about="#description">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Activite"/>
</owl:DatatypeProperty>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#duree -
->

<owl:DatatypeProperty rdf:about="#duree">
  <rdfs:domain rdf:resource="#Activite"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#feedback -->

```

```
<owl:DatatypeProperty rdf:about="#feedback">
  <rdfs:domain rdf:resource="#Activite"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#methode
-->

<owl:DatatypeProperty rdf:about="#methode">
  <rdfs:domain rdf:resource="#Activite"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#numero
-->

<owl:DatatypeProperty rdf:about="#numero">
  <rdfs:domain rdf:resource="#Concept"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#ordonnancement -
->

<owl:DatatypeProperty rdf:about="#ordonnancement">
  <rdfs:domain rdf:resource="#Sequence"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#outil -
->

<owl:DatatypeProperty rdf:about="#outil">
  <rdfs:domain rdf:resource="#Activite"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#representation -
->

<owl:DatatypeProperty rdf:about="#representation">
  <rdfs:domain rdf:resource="#Ressource"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>
```

```
<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#scenario -->

<owl:DatatypeProperty rdf:about="#scenario">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Activite"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#structure -->

<owl:DatatypeProperty rdf:about="#structure">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Ressource"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#synchronisation
-->

<owl:DatatypeProperty rdf:about="#synchronisation">
  <rdfs:domain rdf:resource="#Activite"/>
  <rdfs:range rdf:resource="&xsd:string"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#technique -->

<owl:DatatypeProperty rdf:about="#technique">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Ressource"/>
</owl:DatatypeProperty>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#titre -
->

<owl:DatatypeProperty rdf:about="#titre">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Concept"/>
</owl:DatatypeProperty>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#typeConnaissance
-->

<owl:DatatypeProperty rdf:about="#typeConnaissance">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#ObjectifFinal"/>
</owl:DatatypeProperty>
```

```
<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#url -->

<owl:DatatypeProperty rdf:about="#url">
  <rdfs:range rdf:resource="&xsd:string"/>
  <rdfs:domain rdf:resource="#Ressource"/>
</owl:DatatypeProperty>

<!--

////////////////////////////////////
////////////////////////////////////
//
// Classes
//

////////////////////////////////////
////////////////////////////////////
-->

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#Activite -->

<owl:Class rdf:about="#Activite">
  <rdfs:subClassOf rdf:resource="#Concept"/>
</owl:Class>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#Concept
-->

<owl:Class rdf:about="#Concept">
  <rdfs:subClassOf rdf:resource="&owl;Thing"/>
</owl:Class>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#ObjectifFinal --
>

<owl:Class rdf:about="#ObjectifFinal">
  <rdfs:subClassOf rdf:resource="#Concept"/>
</owl:Class>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#ObjectifGlobal -
->

<owl:Class rdf:about="#ObjectifGlobal">
  <rdfs:subClassOf rdf:resource="#Concept"/>
</owl:Class>
```

```
<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#Ressource -->

<owl:Class rdf:about="#Ressource">
  <rdfs:subClassOf rdf:resource="#Concept"/>
</owl:Class>

<!--
http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#Sequence -->

<owl:Class rdf:about="#Sequence">
  <rdfs:subClassOf rdf:resource="#Concept"/>
</owl:Class>

<!-- http://www.up2uml.org/ontologies/2007/7/contentmodel.owl#Tache -
->

<owl:Class rdf:about="#Tache">
  <rdfs:subClassOf rdf:resource="#Concept"/>
</owl:Class>

<!-- http://www.w3.org/2002/07/owl#Thing -->

<owl:Class rdf:about="&owl;Thing"/>
</rdf:RDF>
```


Annexe 6

Artefacts obtenus dans le cadre du projet Européen UP2UML

1. Recueil des besoins de formation du métier

Relations entre Acteurs, Tâche et Artefacts

Nous avons exprimé que l'acteur « system analyst » était responsable de deux tâches (« requirements and analysis »). Chaque tâche produisait un artefact RUP (ex : la tâche « requirements » produit l'artefact « use case model »). Nous avons aussi exprimé la relation entre les artefacts RUP et les diagrammes UML (Figure 5-6).

(o) = Optional

RUP Actors	Tasks	RUP Artefacts	UML Diagrams
Systems Analyst	Requirements	Use Case Model	Use Case Diagram Sequence Diagram Statechart Diagram (o) Activity Diagram (o)
	Analysis	Analysis Model	Class Diagram Object Diagram Collaboration Diagram Sequence Diagram (o) Statechart Diagram (o) Activity Diagram (o)

Figure 5-6 : relations entre Acteurs, Tâches et Artefacts

Décomposition structurelle d'une tâche

Par exemple, la Figure 5-7 montre la décomposition structurelle de la tâche « requirements » de la méthode de conception logicielle RUP.

- T0378 : Requirements
 - T0084 : Define the system
 - T0550 : Detect Actors
 - T0720 : Detect use cases
 - T0018 : Detect how actors and use cases interact
 - T0617 : Package use cases and actors
 - T0762 : Detect Business Classes
 - T0721 : Refine the system definition
 - T0499 : Determine flow of events in use cases
 - T0220 : Structure use cases
 - T0034 : Structure actors

Figure 5-7 : décomposition structurelle de la tâche « requirements »

Dépendances des tâches métiers

La Figure 5-8 montre les dépendances entre les tâches métier de la phase de « requirements » de la méthode RUP. On y découvre que la tâche T0018 dépend des tâches

T0550 et T0720. Par contre aucune dépendance n'est exprimée entre les tâches T0550 et T0720. L'élève pourra donc commencer son étude par l'une ou l'autre des deux tâches.

Tasks	Dependency
T0018 : Detect how actors and use cases interact	<ul style="list-style-type: none"> • T0550:Detect Actors • T0720:Detect use cases
T0034 : Structure actors	<ul style="list-style-type: none"> • T0617:Package use cases and actors
T0220 : Structure use cases	<ul style="list-style-type: none"> • T0617:Package use cases and actors
T0499 : Determine flow of events in use cases	<ul style="list-style-type: none"> • T0018:Detect how actors and use cases interact
T0617 : Package use cases and actors	<ul style="list-style-type: none"> • T0550:Detect Actors • T0720:Detect use cases

Figure 5-8 : dépendance entre les tâches métiers

Compétences métier

La Figure 5-9 montre la décomposition structurelle en objectifs globaux de la tâche « Detect Actors ».

<ul style="list-style-type: none"> ▪ T0550 : Detect Actors <ul style="list-style-type: none"> G0124 : Distinguish user business roles G0704 : Apply actors creation process G0798 : Understand actors creation process G0253 : Know actors creation process G0150 : Know use case diagram G0560 : Know actors G0789 : Know business roles

Figure 5-9 : décomposition structurelle des compétences métier

La Figure 5-10 montre une vue partielle des dépendances pour les compétences G0560, G0308 et G0205. On observe par exemple que la compétence G0560 ne pourra être abordée que lorsque les compétences G0150 et G0789 seront acquises par l'élève.

Global Objectives	Dependency
G0560 : Know actors	<ul style="list-style-type: none"> • G0150 : Know use case diagram • G0789 : Know business roles
G0308 : Know use case	<ul style="list-style-type: none"> • G0150 : Know use case diagram • G0566 : Know business activity
G0205 : Know relationship between use cases and actors	<ul style="list-style-type: none"> • G0150 : Know use case diagram • G0317 : Know relationships between business activities and business roles

Figure 5-10 : dépendances des compétences métier

2. Analyse pédagogique

Objectifs pédagogiques

Pour exprimer un objectif pédagogique pour l'apprentissage d'un *concept*, nous avons utilisé les verbes suivants :

- au niveau *connaissance* les trois verbes suivants :
 - Le verbe « **define** » est utilisé pour indiquer que l'élève doit être capable de choisir parmi une liste de définition, celle qui correspond au concept associé.
 - Le verbe « **list** » est utilisé pour exprimer qu'un élève doit être capable de faire une liste des instances du concept associé.
 - Le verbe « **identify** » est utilisé pour exprimer qu'un élève doit être capable de retrouver dans une liste, l'instance du concept associé.
- au niveau *compréhension* les deux verbes suivants :
 - Le verbe « **explain** » est utilisé pour indiquer que l'élève doit être capable de donner une définition du concept avec son propre vocabulaire.
 - Le verbe « **associate** » est utilisé pour exprimer qu'un élève doit être capable de relier des éléments un à un entre deux listes.
- au niveau *application* le verbe suivant :
 - Le verbe « **create** » est utilisé pour indiquer que l'élève doit être capable de créer une instance du concept associé.

- au niveau *analyse* le verbe suivant :
 - Le verbe « **compare** » est utilisé pour indiquer que l'élève doit être capable de comparer plusieurs instances d'un concept, et de choisir celle qui est la plus réaliste.
- au niveau *synthèse* le verbe suivant :
 - Le verbe « **deduce** » est utilisé pour indiquer que l'élève doit être capable de combiner plusieurs instances d'un concept afin d'en faire un tout cohérent.

La Figure 5-11 montre la décomposition structurelle des objectifs globaux G0150, G0308 et G0566 en objectifs pédagogiques.

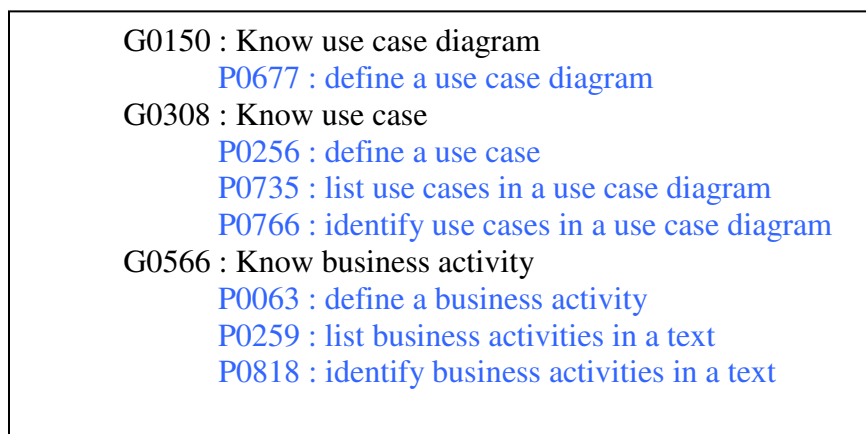


Figure 5-11 : décomposition structurelle des objectifs pédagogiques

La Figure 5-12 montre un exemple de dépendances entre objectifs pédagogiques.

Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
<p>P0677 : define a use case diagram</p>	<ul style="list-style-type: none"> • P0003 : identify relationship between business roles and business activities in a use case diagram • P0285 : identify relationship between business roles and business activities in a text • P0521 : list relationship between business roles and business activities in a use case diagram • P0699 : define relationship between business roles and business activities (between actors and use case) • P0739 : list relationship between business roles and business activities in a text

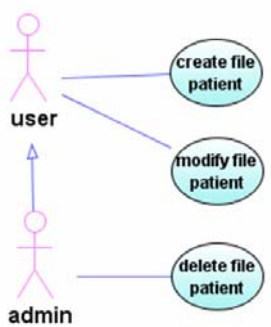
Figure 5-12 : dépendances des objectifs pédagogiques

3. Conception pédagogique

La Figure 5-13 représente la maquette de l'écran proposé pour travailler l'objectif pédagogique « *Identify an actor in a use case diagram* » par une activité d'apprentissage active.

■ Please select from the list below those **actors** that are capable to create and modify a patient file in the system

- Create file patient
- User
- Delete file patient
- Admin
- Modify file patient



Evaluate
Reset
Help
Solution

Figure 5-13 : maquette écran du verbe « Identify »

Les ressources sont des fichiers HTML dans lesquels sont inclus d'autres composants tels que des animations FLASH ou bien tout simplement des images, des tableaux, du texte ou des formulaires.

Le Tableau 5-5 présente le type de ressources que nous avons utilisées pour instancier certains types d'activités pour chaque type de connaissances (concept, fait, processus, principe, procédé). Il est à noter que les activités de type « active » et « interrogative » ne dépendent pas du type de connaissances. Seules les activités de type expositive varient en fonction du type de connaissances [Cisco System, 2000].

Type d'activité			Type de ressource
Expositive	Fait	Exemple	- Animation FLASH
		Contre exemple	- Texte HTML avec représentations graphiques (histogramme, courbe, camembert)
	Concept	Définition	- Animation FLASH
		Exemple	- Texte HTML avec représentations graphiques (images, photos, dessin)
		Contre exemple	
	Procédure	Définition	- Animation FLASH
		Exemple	- Texte HTML avec représentations graphiques (blocs diagramme, images, photos, dessin)
	Principe	Définition	
		Exemple	
	Active	Exercice	deplacédepose
completeblanc			Animation FLASH Formulaire HTML
Experience		- Animation FLASH - Texte HTML avec représentations graphiques (images, photos, dessin)	
Interrogative	formative	testVF	Formulaire HTML
		QCM1	
		QCMN	
	sommative	testVF	
		QCM1	
		QCMN	

Tableau 5-5 : type de ressources en fonction du type d'activités

Annexe 7

Ontologie du domaine du projet UP2UML

Process Breakdown Structure (PBS)

- T0378 : Requirements
 - T0084 : Define the system
 - T0550 : Detect Actors
 - T0720 : Detect use cases
 - T0018 : Detect how actors and use cases interact
 - T0617 : Package use cases and actors
 - T0762 : Detect Business Classes
 - T0721 : Refine the system definition
 - T0499 : Determine flow of events in use cases
 - T0220 : Structure use cases
 - T0034 : Structure actors

- T0776 : Analysis
 - T0769 : Establish an initial architecture
 - T0608 : Analyse software architecture
 - T0096 : Structure Business Classes
 - T0473 : Analyse Behaviour
 - T0651 : Establish the screen navigation paths
 - T0688 : Establish Class responsibility

- T0568 : Design
 - T0213 : Refine the architecture
 - T0579 : Design Use Case
 - T0687 : Use Design Pattern
 - T0395 : Detail the design

- T0503 : Implementation
 - T0459 : Structure components
 - T0244 : Implement components

- T0367 : Tests
 - T0201 : Realize Unit Test
 - T0530 : Realize Integration Test
 - T0791 : Realize Validation Test

- T0481 : Planning and Monitoring
 - T0061 : Establish software development process
 - T0087 : Establish Iteration plan

Tasks Dependency (TD)

Tasks	Dependency
T0018 : Detect how actors and use cases interact	<ul style="list-style-type: none"> • T0550:Detect Actors • T0720:Detect use cases
T0034 : Structure actors	<ul style="list-style-type: none"> • T0617:Package use cases and actors
T0096 : Structure Business Classes	<ul style="list-style-type: none"> • T0762:Detect Business Classes
T0220 : Structure use cases	<ul style="list-style-type: none"> • T0617:Package use cases and actors
T0244 : Implement components	<ul style="list-style-type: none"> • T0459:Structure components
T0395 : Detail the design	<ul style="list-style-type: none"> • T0579:Design Use Case
T0459 : Structure components	<ul style="list-style-type: none"> • T0395:Detail the design
T0499 : Determine flow of events in use cases	<ul style="list-style-type: none"> • T0018:Detect how actors and use cases interact
T0579 : Design Use Case	<ul style="list-style-type: none"> • T0034:Structure actors • T0220:Structure use cases • T0499:Determine flow of events in use cases • T0608:Analyse software architecture • T0651:Establish the screen navigation paths • T0687:Use Design Pattern • T0688:Establish Class responsibility
T0617 : Package use cases and actors	<ul style="list-style-type: none"> • T0550:Detect Actors • T0720:Detect use cases
T0688 : Establish Class responsibility	<ul style="list-style-type: none"> • T0096:Structure Business Classes

Tasks Breakdown Structure (TBS)

Tasks	Business Objectives
T0018 : Detect how actors and use cases interact	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases

	<ul style="list-style-type: none"> <ul style="list-style-type: none"> ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0635 : Distinguish the relationship between business roles and business activities <ul style="list-style-type: none"> ○ B0488 : start practice on how actors and use cases interact ○ G0787 : Apply relationship process creation <ul style="list-style-type: none"> ▪ G0347 : Understand relationship process creation <ul style="list-style-type: none"> ▪ G0585 : Know relationship process creation <ul style="list-style-type: none"> ▪ G0150 : Know use case diagram ▪ G0317 : Know relationship between business activities and business roles ▪ G0205 : Know relationship between use cases and actors ○ B0343 : execute practice on how actors and use cases interact ○ B0723 : terminate practice on how actors and use cases interact
T0034 : Structure actors	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0756 : Analyse relationship between userâ€™s business roles
T0061 : Establish software development process	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases

	<ul style="list-style-type: none"> ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0323 : Understand Analysis model • G0533 : Understand Design Model • G0569 : Understand Implementation Model • G0021 : Understand Deployment Model • G0380 : Design RUP based software method for the project
T0087 : Establish Iteration plan	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0480 : Understand Use case model • G0323 : Understand Analysis model • G0533 : Understand Design Model • G0569 : Understand Implementation Model • G0021 : Understand Deployment Model • G0306 : Plan RUP iteration for the project
T0096 : Structure Business Classes	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0232 : Analyse relationship between business terms
T0201 : Realize Unit Test	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle

	<ul style="list-style-type: none"> ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0533 : Understand Design Model • G0569 : Understand Implementation Model • G0668 : Apply unit testing
T0220 : Structure use cases	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0207 : Analyse relationship between business activities
T0244 : Implement components	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0164 : Analyse components collaboration
T0395 : Detail the design	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities)

	<ul style="list-style-type: none"> ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0161 : Design Class
T0459 : Structure components	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0125 : Organize design classes into components
T0499 : Determine flow of events in use cases	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0470 : Analyse steps of business activity
T0530 : Realize Integration Test	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0533 : Understand Design Model • G0569 : Understand Implementation Model • G0021 : Understand Deployment Model

	<ul style="list-style-type: none"> • G0763 : Apply integration testing
T0550 : Detect Actors	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0124 : Distinguish user's business roles <ul style="list-style-type: none"> ○ B0095 : start practice on actor detection ○ G0704 : Apply actors process creation <ul style="list-style-type: none"> ▪ G0798 : Understand actors process creation <ul style="list-style-type: none"> ▪ G0253 : Know actors process creation <ul style="list-style-type: none"> ▪ G0150 : Know use case diagram ▪ G0560 : Know actors ▪ G0789 : Know business roles ○ B0341 : execute practice on actor detection ○ B0206 : terminate practice on actor detection
T0579 : Design Use Case	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0480 : Understand Use case model • G0323 : Understand Analysis model • G0143 : Deduce design class from analysis class
T0608 : Analyse software architecture	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle

	<ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0524 : Distinguish actor interfaces Classes • G0559 : Distinguish controls classes • G0414 : Distinguish entities classes
T0617 : Package use cases and actors	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0262 : Organize use cases • G0752 : Organize actors
T0651 : Establish the screen navigation paths	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0221 : Analyse the screen flows
T0687 : Use Design Pattern	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities)

	<ul style="list-style-type: none"> ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0803 : Organize design class to comply with design pattern
T0688 : Establish Class responsibility	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0665 : Analyse business class collaboration
T0720 : Detect use cases	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0371 : Distinguish business activities <ul style="list-style-type: none"> ○ B0494 : start practice on use cases detection ○ G0700 : Apply use case process creation <ul style="list-style-type: none"> ▪ G0567 : Understand use case process creation <ul style="list-style-type: none"> ▪ G0643 : Know use case process creation <ul style="list-style-type: none"> ▪ G0150 : Know use case diagram ▪ G0308 : Know use case ▪ G0566 : Know business activity ○ B0078 : execute practice on use cases detection ○ B0730 : terminate practice on use cases detection
T0762 : Detect Business	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method

Classes	<ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0666 : Classify business terms
T0791 : Realize Validation Test	<ul style="list-style-type: none"> • G0505 : Understand RUP based software method <ul style="list-style-type: none"> ○ G0498 : Know software development method lifecycle ○ G0313 : Know RUP method lifecycle <ul style="list-style-type: none"> ▪ G0526 : Know RUP method <ul style="list-style-type: none"> ▪ G0774 : Know RUP phases ▪ G0299 : Know RUP workflows (activities) ▪ G0108 : Know RUP models ▪ G0175 : Know RUP roles ○ G0686 : Know UML language <ul style="list-style-type: none"> ▪ G0242 : Know UML diagrams • G0480 : Understand Use case model • G0323 : Understand Analysis model • G0657 : Apply validation testing

Annexe 8
Ontologie pédagogique du projet UP2UML

Global Objectives Breakdown Structure (GOBS)

Business Objective	Pedagogical Objective
G0108 : Know RUP models	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0479 : define RUP models (use case model, analysis model, design model, deployment model, implementation model, test model) • P0663 : list RUP models
Business Objective	Pedagogical Objective
G0124 : Distinguish user's business roles	<p>LVI 4 (Analysis)</p> <ul style="list-style-type: none"> • P0249 : compare actors comply with a text description of business roles
Business Objective	Pedagogical Objective
G0150 : Know use case diagram	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0677 : define term : use case diagram
Business Objective	Pedagogical Objective
G0175 : Know RUP roles	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0054 : list principal RUP roles • P0350 : define principal RUP roles (Systems analyst, Software architect, Implementer, Project Manager)
Business Objective	Pedagogical Objective
G0205 : Know relationship between use cases and actors	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0314 : define relationship between actors and use cases
Business Objective	Pedagogical Objective
G0242 : Know UML diagrams	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0016 : list UML Diagrams • P0628 : define UML Diagrams (use

	<p>case diagram, sequence diagram, statechart diagram, activity diagram, class diagram, object diagram, collaboration diagram, deployment diagram, component diagram)</p> <ul style="list-style-type: none"> • P0714 : identify UML Diagrams
Business Objective	Pedagogical Objective
G0299 : Know RUP workflows (activities)	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0329 : list RUP workflows (activities) • P0631 : define RUP workflows (requirements, analyse, design, implementation, test)
Business Objective	Pedagogical Objective
G0308 : Know use case	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0256 : define term : use case • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram
Business Objective	Pedagogical Objective
G0313 : Know RUP method lifecycle	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0001 : identify RUP workflows (activities) • P0006 : identify RUP models • P0792 : identify RUP phases
Business Objective	Pedagogical Objective
G0317 : Know relationship between business activities and business roles	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0003 : identify relationship between business roles and business activities in a use case diagram • P0285 : identify relationship between business roles and business activities in a text • P0521 : list relationship between business roles and business activities in a use case diagram • P0699 : define relationship between

	<p>business roles and business activities (between actors and use case)</p> <ul style="list-style-type: none"> • P0739 : list relationship between business roles and business activities in a text
Business Objective	Pedagogical Objective
G0347 : Understand relationship process creation	<p>LVI 2 (Comprehension)</p> <ul style="list-style-type: none"> • P0014 : explain relationship between actors and use case • P0128 : associate relationship description in a text with relation between actors and use case in a use case diagram
Business Objective	Pedagogical Objective
G0371 : Distinguish business activities	<p>LVI 4 (Analysis)</p> <ul style="list-style-type: none"> • P0592 : compare use cases comply with a text description of business activities
Business Objective	Pedagogical Objective
G0498 : Know software development method lifecycle	<p>LVI 1 (Knowledge)</p> <ul style="list-style-type: none"> • P0076 : list software development lifecycle phases • P0102 : define term : Lifecycle • P0139 : define term : Method • P0605 : list artefact's software development lifecycle phases
Business Objective	Pedagogical Objective
G0505 : Understand RUP based software method	<p>LVI 2 (Comprehension)</p> <ul style="list-style-type: none"> • P0394 : associate RUP roles with RUP models • P0465 : explain UML Diagrams • P0492 : associate RUP models with UML Diagrams • P0674 : associate RUP workflows with RUP models
Business Objective	Pedagogical Objective

G0526 : Know RUP method	LVI 1 (Knowledge) <ul style="list-style-type: none"> • P0126 : define RUP method
Business Objective	Pedagogical Objective
G0560 : Know actors	LVI 1 (Knowledge) <ul style="list-style-type: none"> • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0591 : list actors in use case diagram
Business Objective	Pedagogical Objective
G0566 : Know business activity	LVI 1 (Knowledge) <ul style="list-style-type: none"> • P0063 : define term : business activity • P0259 : list business activities in a text • P0818 : identify business activities in a text
Business Objective	Pedagogical Objective
G0567 : Understand use case process creation	LVI 2 (Comprehension) <ul style="list-style-type: none"> • P0042 : explain use case • P0088 : explain business activity • P0520 : explain use case diagram • P0689 : associate business activities description in a text with use case in use case diagram
Business Objective	Pedagogical Objective
G0635 : Distinguish the relationship between business roles and business activities	LVI 4 (Analysis) <ul style="list-style-type: none"> • P0743 : compare relationship comply with a text description of relationship between business roles and business activities
Business Objective	Pedagogical Objective
G0686 : Know UML language	LVI 1 (Knowledge)

	<ul style="list-style-type: none"> • P0390 : define UML language
Business Objective	Pedagogical Objective
G0700 : Apply use case process creation	LVI 3 (Application) <ul style="list-style-type: none"> • P0218 : create use cases according to a text description of business activities
Business Objective	Pedagogical Objective
G0704 : Apply actors process creation	LVI 3 (Application) <ul style="list-style-type: none"> • P0282 : create actors according to a text description of business roles
Business Objective	Pedagogical Objective
G0774 : Know RUP phases	LVI 1 (Knowledge) <ul style="list-style-type: none"> • P0111 : list RUP phases • P0611 : define RUP phases (inception, Elaboration, construction, transition)
Business Objective	Pedagogical Objective
G0787 : Apply relationship process creation	LVI 3 (Application) <ul style="list-style-type: none"> • P0059 : create relationship according to a text description of relationship between business roles and business activities
Business Objective	Pedagogical Objective
G0789 : Know business roles	LVI 1 (Knowledge) <ul style="list-style-type: none"> • P0296 : identify business roles in a text • P0438 : define term : business roles • P0557 : list business roles in a text
Business Objective	Pedagogical Objective
G0798 : Understand actors process creation	LVI 2 (Comprehension) <ul style="list-style-type: none"> • P0083 : associate business roles

	<p>description in a text with actors in use case diagram</p> <ul style="list-style-type: none"> • P0383 : explain actors • P0388 : explain business roles • P0520 : explain use case diagram
--	---

Pedagogical Objectives Dependency (POD)

Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
<p>P0003 : identify relationship between business roles and business activities in a use case diagram</p>	<ul style="list-style-type: none"> • P0063 : define term : business activity • P0256 : define term : use case • P0259 : list business activities in a text • P0296 : identify business roles in a text • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0438 : define term : business roles • P0557 : list business roles in a text • P0591 : list actors in use case diagram • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram • P0818 : identify business activities in a text
Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
<p>P0285 : identify relationship between business roles and business activities in a text</p>	<ul style="list-style-type: none"> • P0063 : define term : business activity • P0256 : define term : use case • P0259 : list business activities in a text • P0296 : identify business roles in a text • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0438 : define term : business roles • P0557 : list business roles in a text • P0591 : list actors in use case

	<p>diagram</p> <ul style="list-style-type: none"> • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram • P0818 : identify business activities in a text
Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
<p>P0521 : list relationship between business roles and business activities in a use case diagram</p>	<ul style="list-style-type: none"> • P0063 : define term : business activity • P0256 : define term : use case • P0259 : list business activities in a text • P0296 : identify business roles in a text • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0438 : define term : business roles • P0557 : list business roles in a text • P0591 : list actors in use case diagram • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram • P0818 : identify business activities in a text
Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
<p>P0677 : define term : use case diagram</p>	<ul style="list-style-type: none"> • P0003 : identify relationship between business roles and business activities in a use case diagram • P0285 : identify relationship between business roles and business activities in a text • P0521 : list relationship between business roles and business activities in a use case diagram • P0699 : define relationship between business roles and business activities (between actors and use case) • P0739 : list relationship between business roles and business activities

	in a text
Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
P0699 : define relationship between business roles and business activities (between actors and use case)	<ul style="list-style-type: none"> • P0063 : define term : business activity • P0256 : define term : use case • P0259 : list business activities in a text • P0296 : identify business roles in a text • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0438 : define term : business roles • P0557 : list business roles in a text • P0591 : list actors in use case diagram • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram • P0818 : identify business activities in a text
Pedagogical Objectives : LVI 1 (Knowledge)	Dependency
P0739 : list relationship between business roles and business activities in a text	<ul style="list-style-type: none"> • P0063 : define term : business activity • P0256 : define term : use case • P0259 : list business activities in a text • P0296 : identify business roles in a text • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0438 : define term : business roles • P0557 : list business roles in a text • P0591 : list actors in use case diagram • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram • P0818 : identify business activities in a text

Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0014 : explain relationship between actors and use case	<ul style="list-style-type: none"> • P0003 : identify relationship between business roles and business activities in a use case diagram • P0285 : identify relationship between business roles and business activities in a text • P0521 : list relationship between business roles and business activities in a use case diagram • P0699 : define relationship between business roles and business activities (between actors and use case) • P0739 : list relationship between business roles and business activities in a text • P0042 : explain use case • P0088 : explain business activity • P0383 : explain actors • P0388 : explain business roles • P0520 : explain use case diagram
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0042 : explain use case	<ul style="list-style-type: none"> • P0256 : define term : use case • P0735 : list use case in use case diagram • P0766 : identify use case in use case diagram
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0083 : associate business roles description in a text with actors in use case diagram	<ul style="list-style-type: none"> • P0383 : explain actors • P0388 : explain business roles
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0088 : explain business activity	<ul style="list-style-type: none"> • P0063 : define term : business activity • P0259 : list business activities in a text • P0818 : identify business activities in a text
Pedagogical Objectives : LVI 2	Dependency

(Comprehension)	
P0128 : associate relationship description in a text with relation between actors and use case in a use case diagram	<ul style="list-style-type: none"> • P0014 : explain relationship between actors and use case
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0383 : explain actors	<ul style="list-style-type: none"> • P0332 : define term : actors • P0398 : identify actors in use case diagram • P0591 : list actors in use case diagram
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0388 : explain business roles	<ul style="list-style-type: none"> • P0296 : identify business roles in a text • P0438 : define term : business roles • P0557 : list business roles in a text
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0394 : associate RUP roles with RUP models	<ul style="list-style-type: none"> • P0001 : identify RUP workflows (activities) • P0006 : identify RUP models • P0111 : list RUP phases • P0329 : list RUP workflows (activities) • P0479 : define RUP models (use case model, analysis model, design model, deployment model, implementation model, test model) • P0611 : define RUP phases (inception, Elaboration, construction, transition) • P0631 : define RUP workflows (requirements, analyse, design, implementation, test) • P0663 : list RUP models • P0792 : identify RUP phases
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0465 : explain UML Diagrams	<ul style="list-style-type: none"> • P0016 : list UML Diagrams • P0628 : define UML Diagrams (use case diagram, sequence diagram, statechart diagram, activity diagram,

	<p>class diagram, object diagram, collaboration diagram, deployment diagram, component diagram)</p> <ul style="list-style-type: none"> • P0677 : define term : use case diagram • P0714 : identify UML Diagrams
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0492 : associate RUP models with UML Diagrams	<ul style="list-style-type: none"> • P0394 : associate RUP roles with RUP models • P0465 : explain UML Diagrams
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0520 : explain use case diagram	<ul style="list-style-type: none"> • P0042 : explain use case • P0088 : explain business activity • P0383 : explain actors • P0388 : explain business roles • P0465 : explain UML Diagrams
Pedagogical Objectives : LVI 2 (Comprehension)	Dependency
P0689 : associate business activities description in a text with use case in use case diagram	<ul style="list-style-type: none"> • P0042 : explain use case • P0088 : explain business activity
Pedagogical Objectives : LVI 3 (Application)	Dependency
P0059 : create relationship according to a text description of relationship between business roles and business activities	<ul style="list-style-type: none"> • P0128 : associate relationship description in a text with relation between actors and use case in a use case diagram
Pedagogical Objectives : LVI 3 (Application)	Dependency
P0218 : create use cases according to a text description of business activities	<ul style="list-style-type: none"> • P0689 : associate business activities description in a text with use case in use case diagram
Pedagogical Objectives : LVI 3 (Application)	Dependency
P0282 : create actors according to a text description of business roles	<ul style="list-style-type: none"> • P0083 : associate business roles description in a text with actors in use case diagram

	<ul style="list-style-type: none"> • P0520 : explain use case diagram
Pedagogical Objectives : LVI 4 (Analysis)	Dependency
P0249 : compare actors comply with a text description of business roles	<ul style="list-style-type: none"> • P0282 : create actors according to a text description of business roles
Pedagogical Objectives : LVI 4 (Analysis)	Dependency
P0592 : compare use cases comply with a text description of business activities	<ul style="list-style-type: none"> • P0520 : explain use case diagram • P0218 : create use cases according to a text description of business activities
Pedagogical Objectives : LVI 4 (Analysis)	Dependency
P0743 : compare relationship comply with a text description of relationship between business roles and business activities	<ul style="list-style-type: none"> • P0059 : create relationship according to a text description of relationship between business roles and business activities

Exemple de description de sequence pour l'objectif pédagogique « P0001 : identify RUP workflows »

Title	Description	Scenario	Actors / Mode	Feedback	Delivery Mode / Synchronization	Nature / Duration / Narrative	Required Tools / Complementary Tools	Id Activity
Sequence Number : S0467 This sequence is meant for objectives of BLOOM level = 1. It concentrates on Identification. It composes of : - exercises - animations - definitions - examples - formative evaluations - summative evaluation								
Introduce sequence	A simple screen or verbal introduction to explain the objective, content and activities of the sequence.	- Learner starts the sequence - System introduces the aim of the sequence - Learner reads (and/or listens to) the aim of the sequence - System invites to start the sequence	System, Individual /		Online /	Self-paced / 1 minute /	/	ACIntroduction6
Study definition	The system presents textual and/or verbal definition of lifecycle phases diagram (without text for workflows). Illustration are used whenever possible to help memories the definition.	- System introduces the aim of the activity - System presents a definition of lifecycle phases diagram (without text for workflows) - Learner reads (and/or listens to) the definition - System presents a summary of the activity	System, Individual /		Online /	Self-paced / 3 minutes /	/	ACIdentify306
Study example	A series of examples of lifecycle phases diagram (without text for workflows) are presented to the learner. The examples situate the concept within its context of usage by displaying related artefacts.	- System introduces the aim of the activity - System presents examples of artefacts featuring lifecycle phases diagram (without text for workflows) - Learner reads (and/or listens to) examples - System presents a summary of the activity	System, Individual /		Online /	Self-paced / 3 minutes /	/	ACIdentify406
Play animation	These animations	- System introduces the	System, Individual /		Online /	Self-paced /	/	ACIdentify206

	<p>are meant to show how to identify words that match the concept of lifecycle phases diagram (without text for workflows). The system displays a lifecycle phases diagram (without text for workflows) and a list of words. It moves each word onto the lifecycle phases diagram (without text for workflows) displaying an explanation.</p>	<p>aim of the activity - System displays a list of animations - Learner selects an animation - System presents lifecycle phases diagram (without text for workflows) and several terms - System places (one by one) the terms onto lifecycle phases diagram (without text for workflows) with a short explanation - Learner can replay the same animation or select another one - System presents a summary of the activity</p>				3 minutes /		
Perform exercise	<p>This exercise consists of identifying words that match the concept of lifecycle phases diagram (without text for workflows). The system displays a lifecycle phases diagram (without text for workflows) and a list of words. The learner identifies the relevant words and moves them onto the lifecycle phases diagram (without text for workflows).</p>	<p>- System introduces the aim of the activity - System presents lifecycle phases diagram (without text for workflows) and several terms - Learner selects terms and place them on the lifecycle phases diagram (without text for workflows) - System displays the errors - System suggests another exercise - System presents a summary of the activity</p>	System, Individual /	System	Online /	Self-paced / 5 minutes /		ACIdentify106

Perform formative evaluation	The assessment consists of identifying words that match the concept of lifecycle phases diagram (without text for workflows). The system displays a lifecycle phases diagram (without text for workflows) and a list of words. The learner identify the relevant words and moves them onto the lifecycle phases diagram (without text for workflows). The system evaluates the answers, displays errors and recommends links to relevant material.	- System introduces the aim of the activity - System presents lifecycle phases diagram (without text for workflows) and several terms - Learner selects terms and place it in lifecycle phases diagram (without text for workflows) - System evaluates answers and shows errors - System recommends links to material related to errors - Learner browses the recommended material or start moves to another exercise within the evaluation. - System presents a summary of the activity	System,Individual /	System	Online /	Self-paced / 5 minutes /		ACIdentify506
Perform summative evaluation	The system displays a lifecycle phases diagram (without text for workflows) and a list of words. The learner identify the relevant words and moves them onto the lifecycle phases diagram (without text for workflows). The system evaluates the answers, and displays performance.	- System introduces the aim of the activity - System presents lifecycle phases diagram (without text for workflows) and several terms - Learner selects terms and moves them onto lifecycle phases diagram (without text for workflows) - System evaluates the answer of the learner - System	System,Individual /	System	Online /	Self-paced / 5 minutes /		ACIdentify606

		indicates the percentage of good answers - System presents a summary of the activity						
Summarize sequence	A conclusion is presented featuring a short summary of the sequence, its activities and objective.	- Learner terminates the sequence - System presents a summary of the sequence - Learner reads (and/or listens to) the summary - The system suggests to close the sequence	System, Individual /		Online /	Self-paced / 1 minutes /		ACSummary6

Annexe 9



Éducation et culture

Leonardo da Vinci

Projets pilotes



UML Training Questionnaire

This study investigates the use of blended learning through a product created by the Up2UML consortium. Up2UML (www.up2uml.org) is a European project funded by the Leonardo Da Vinci programme. It aims to enhance UML skills for software development in SMEs by developing new content and training material for UML2. The ultimate goal of the project is to create innovative learning materials for work-based UML training.

The UML (Unified Modelling Language) was originally developed in the mid 1990s as to enable object-oriented software developers to model computer applications. It is a general purpose modelling language that can be used with existing methodologies to create abstract models of a system. It has now become a standard modelling language ISO/IEC 19501:2005. UML is programming-language independent and can easily be adapted by any company.

Would you like the opportunity to learn the UML? This course is conducted in a Blended Learning Mode: Blended learning means that there will be a short initial meeting where topics and fellow learners are introduced, you work through the material at your own time and attend a concluding meeting. During the learning experience, you will have access to your fellow students and an instructor through both discussion forums and e-mail. People will be assigned to each group on a random basis.

At the end of the course, you will be asked to complete two questionnaires, one about whether or not you liked the course, the other will measure how you have met the learning objectives.

If you are interested in participating in this study, please leave your contact details below:

Yes I am interested in participating in this study.

Name	
e-mail address	
Contact number	

Please note: Personal details will not be divulged to any third party and will only be used for the purpose of this study

No, I am not interested in participating in the study.

Thank you for taking the time to fill out this questionnaire. All answers will be treated in confidence.

**Please answer all questions by marking the appropriate answers.
Please choose one answer only unless otherwise stated.**

Section 1: Your role in the organisation

1. How much experience do you have in software development?

- less than one year 1 to 5 years 5 to 10 years more than 10 years

2. How many employees do you have in your company?

- fewer than 10 10 to 50 50 to 100 100 to 500 more than 500

3. What is the average number of people involved in a typical software project in your organisation? people

4. What is your role in software development? (you may choose more than one)

- | | | |
|---|---|--|
| <input type="checkbox"/> Manager | <input type="checkbox"/> Software project manager | <input type="checkbox"/> Software designer |
| <input type="checkbox"/> Software analyst | <input type="checkbox"/> Software maintainer | <input type="checkbox"/> Software tester |
| <input type="checkbox"/> Programmer | <input type="checkbox"/> GUI programmer | <input type="checkbox"/> Quality engineer |
| <input type="checkbox"/> Other (please specify) | | |

5. What is the main application domain of your software development?

- | | | | | |
|---|---------------------------------------|-------------------------------------|------------------------------------|---|
| <input type="checkbox"/> Administration | <input type="checkbox"/> Aeronautics | <input type="checkbox"/> Automotive | <input type="checkbox"/> Chemistry | <input type="checkbox"/> Communication |
| <input type="checkbox"/> Construction | <input type="checkbox"/> Distribution | <input type="checkbox"/> Education | <input type="checkbox"/> Energy | <input type="checkbox"/> Finance |
| <input type="checkbox"/> Food | <input type="checkbox"/> Health | <input type="checkbox"/> Military | <input type="checkbox"/> Tourism | <input type="checkbox"/> Transportation |
| <input type="checkbox"/> Other (please specify) | | | | |

Section 2: Lifelong learning (*you may choose more than one answer in this section*)

6. What forms of training opportunities are provided in your company?

- | | |
|--|---|
| <input type="checkbox"/> Internal Training day | <input type="checkbox"/> Sent on short external courses |
| <input type="checkbox"/> 1-to-1 training at desk | <input type="checkbox"/> e-Learning courses |
| <input type="checkbox"/> Reimbursed for external certified courses | <input type="checkbox"/> Coaching/mentoring |
| <input type="checkbox"/> Other (please specify) | |

7. What is your preferred method of learning?

- | | |
|--|--|
| <input type="checkbox"/> Attend classroom course | <input type="checkbox"/> Complete on-line course |
| <input type="checkbox"/> Read manual/book | <input type="checkbox"/> Attend talk/seminar |
| <input type="checkbox"/> Discussing with colleagues/networking | <input type="checkbox"/> Coaching/mentoring |
| <input type="checkbox"/> Searching articles/FAQs on web | <input type="checkbox"/> Discussion forums |
| <input type="checkbox"/> Other (please specify) | |

Section 3: Experience of UML

8. Please indicate the level of your experience with UML? (tick all that apply)

- Unfamiliar with UML – never heard of it
- Heard of UML, but unfamiliar with it
- Familiar with UML, but feel it is not relevant to my job
- Familiar with UML, but find it too hard to apply
- I don't use UML, but others within my organisation do
- I use some aspects of UML, sometimes
- I use some aspects of UML, always
- I'm a UML guru
- Other (please specify)

8a. Comments (optional)**9. If you are familiar with UML, how did you hear about it?**

- It was part of my higher education course
- I attended training course
- I completed on-line/e-learning course
- I read a book/manual
- I was coached/shown by colleagues
- I searched on-line for relevant resources
- Other (please specify)

Please send completed questionnaire to training@up2uml.org

Thank you for taking the time to complete this questionnaire. All responses will be treated with confidence.

Annexe 10

End of Course Survey

Thank you for taking the time to complete this questionnaire.
Please indicate if you Strongly Agree (SA), Agree(A), are Unsure (U), Disagree (D) or Strongly Disagree (SD) with each of the following statements,

Legend:

SA Strongly Agree
A Agree
U Unsure
D Disagree
SD Strongly Disagree

Section 1: Course Participation

Statement	SA	A	U	D	SD
1. I would participate in a similar type course again	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. I found myself eagerly anticipating the next session	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. I would have liked to attend face-to-face sessions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. The discussion forums helped me to understand the content	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. I would prefer attend classes rather than planning my own learning time	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. The discussion forums helped me to understand the content	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Discussing topics with others helps me to learn better	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. I felt that I was learning in isolation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Interacting with people through the forums did not enhance my learning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. I do not think that discussion forums are essential to on-line courses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. I would like to be able to discuss issues face-to-face with other learners	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. I liked being able to complete the lessons in my own time	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. I would have liked a more structured course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. I would have learned as much if I had completed a blended learning version of the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. The amount of structure in the course was just right	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Statement	SA	A	U	D	SD
16. Developing a relationship with my fellow learners is important to me	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. I don't see the need to get to know my fellow learners	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. I would not be keen to participate in an e-learning course again	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. I would have liked to meet other people doing the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Discussion forums enhance learning in on-line courses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. I like to work out issues myself, without discussion with other learners	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. The availability of discussion forums did not enhance my learning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. The level of interactivity was appropriate for the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Attending class would have had no impact on my learning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. The course was too structured	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Attending a class would have motivated me to learn more	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. I would have liked more interaction opportunities during the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. I felt that I was left on my own without any help if needed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. I am glad I attended the on-line only version of the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. I found it hard to motivate myself to complete the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. I learn best when I go through material myself, reflecting internally on it	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. I felt part of a learning community	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. I would have preferred to complete the course in a traditional classroom setting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34. I felt that the facilitator was available to answer queries at any time	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35. By taking responsibility for my own learning, I retained much more than if I had attended classes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36. Meeting other learners would not impact on my enjoyment of the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 2: Lessons and Resources

Statement	SA	A	U	D	SD
37. I did not think that the project enhanced learning on the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38. The course content was difficult to follow	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39. There was a good range of activities within the lessons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
40. I felt comfortable with the learning environment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
41. The project helped me to see how UML can be applied in real life	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
42. Lessons were boring	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
43. The animations on the course helped me to learn the material	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
44. The resources provided, other than lessons, enhanced the learning experience	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
45. The links enabled me to learn the material better	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
46. Exercises helped reinforce the materials of the lessons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47. I would have preferred more opportunities for reassurance that I was progressing well	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
48. The project allowed me to demonstrate what I learned on the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
49. The class exercises were not useful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
50. It took me a while to get to grips with the course environment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
51. The content was presented at a level which could be easily understood	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
52. I found the links provided useful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
53. There was no need for additional resources, outside the lessons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
54. The course environment was easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
55. The animations facilitated my learning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
56. The project brief was not realistic.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 3: Overall experience

Statement	SA	A	U	D	SD
57. This course will help improve my job performance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
58. The course satisfied my needs and expectations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
59. This course is one of the most difficult I have undertaken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
60. The podcasts did not provide added value to the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
61. The welcome video made me feel part of a class	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
62. This course will have no impact on how I perform at work	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
63. I would not recommend this course to others	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
64. I did not see the point of including the welcome video	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
65. Compared to what I knew before the course, I learned a lot	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
66. I learned very little on this course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
67. I would like additional training before feeling confident enough to use UML on the job	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
68. I see how I can apply what I learned in this course in the workplace	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
69. I found it hard to motivate myself to complete the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
70. I do not feel that I learned a lot on this course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
71. I was disappointed overall in the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
72. I would like to complete additional lessons from this course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
73. I enjoyed completing this course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
74. I thought this was an excellent course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
75. The course did not meet my expectations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
76. The availability of podcasts enhanced my learning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
77. I would not be interested in completing other lessons from the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
78. I have made a big improvement in my skills and knowledge of UML	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
79. I found this course very easy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
80. I would recommend this course to my friends/colleagues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Section 4: Additional feedback

81. Was there anything you particularly liked/did not like about the course?

82. What aspects of the course did you find most/least relevant to your working life?

83. How could the course be improved?

84. Additional comments:

Annexe 11

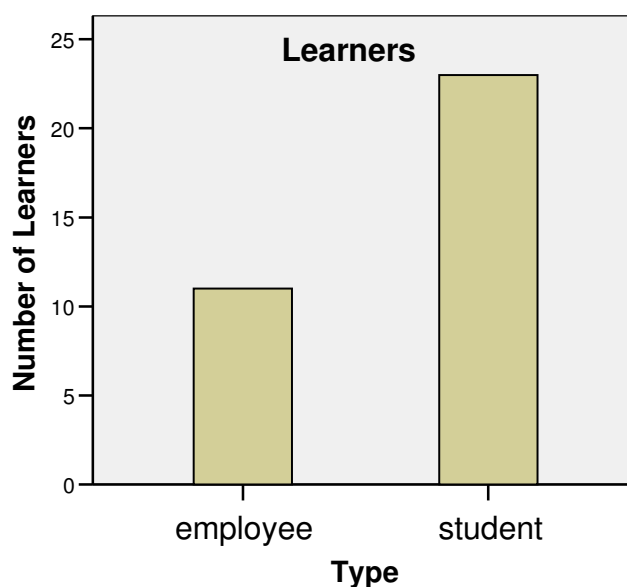
Evaluation projet UP2UML

Learning Environments

For the purpose of this evaluation the lessons were embedded in courses by adding an introduction, table of contents and an overview of the content. The lessons were provided through two Learning Content Management Systems (LCMS), namely Moodle and Intuitext. Through these systems, additional functionality was available such as log-in, synchronous and asynchronous communication, tutor contact etc. However, the evaluation study was designed to focus on the quality of the lessons, i.e., the material developed, only.

Participants

The courses were made available to learners in three different countries, Germany, Romania and Ireland. Learners were recruited mainly within the partner organisations from people who are not involved in the project and without payment. A total of 34 learners provided formal feedback, 23 of these are students enrolled on a software systems course and 11 are employees with various backgrounds.



Data Collection and Analysis

Participants were asked to complete a questionnaire as described in the evaluation plan D4.2. The questionnaire consisted both of quantitative ratings and the opportunity to comment on a number of aspects in free form.

The analysis showed only minor differences between the different countries. Accordingly, the results below are reported as a whole. If there is a statistical significant between countries this will be highlighted.

Also the analysis showed only minor differences between students and employees. Again, if there is a statistical significant between students and employees this

will be highlighted but in general we can assume that the results reported hold for both target groups.

Résultats de l'étude

The items in the questionnaire covered four different main areas: Learning Process, Content, Exercises & Tests, and Navigation & Orientation.

Feedback on Learning Process

Overall, participants perceived the quality of the learning process as very positive. The average per item is consistently better than the mid-scale (i.e., between 1 and 3). The error bars in Figure 5-14 and all subsequent result Figures indicate a 95% confidence interval.

Learners seem to be satisfied with the learning process in general: The learning units are clearly indicated and learning goals are considered at all times.

The two weakest aspects seem to relate to feedback on progress and linearity of entry: Participants felt that the feedback of the course on their progress is too limited. Additional feedback mechanisms such as permanent status bar, number or percentage of tests completed, percentage of screens covered, overview of content successfully passed might be considered.

Moreover, many learners obviously had the impression that they had to go through the content in a linear way. While the prototype offers free navigation and topic selection, this might be caused by either the instruction or the structure of the content. Additional support for non-linear entry should be considered, e.g., sophisticated search, list of recommended resources for specific learning goals or problems, or prior-knowledge tests.

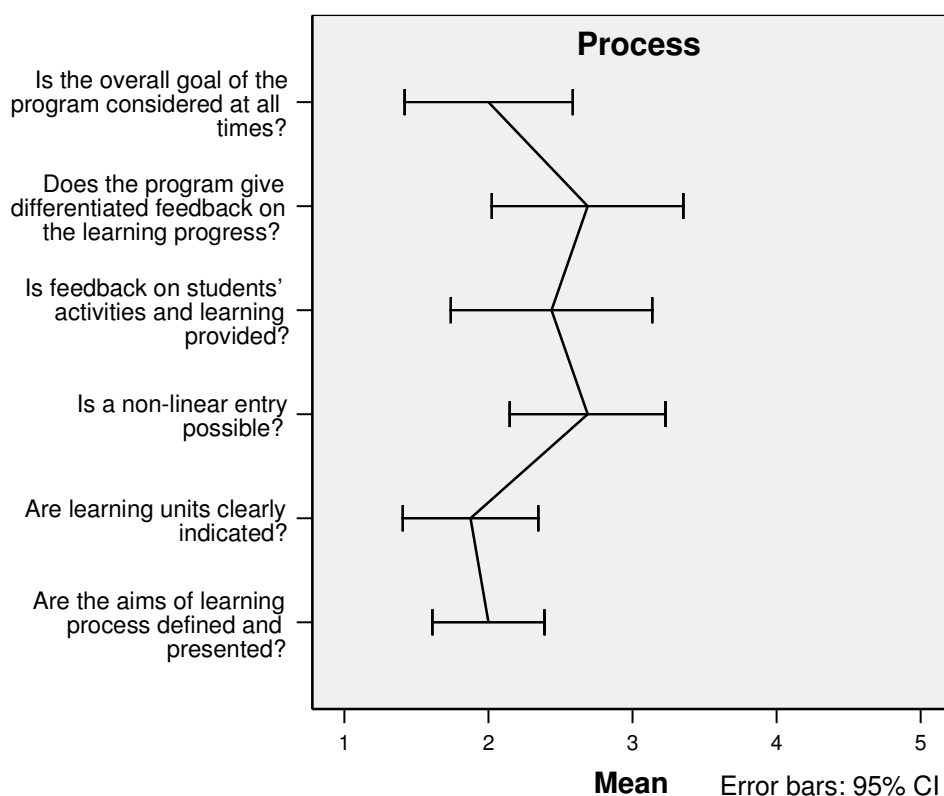


Figure 5-14: Results of Learning Process

The quantitative results above are also reflected in the open ended comments.

Participants commented that the abstract in the beginning of each section “clearly stated” the learning objectives. They felt they were “easy to understand” and “well communicated”.

They also felt the units were clearly separated into “intuitive levels”.

Regarding the feedback about learning progress they commented that “the only feedback provided would be the answer on the multiple choice exercises” which was “quite limited”. They were asking for “dynamic differentiated feedback”.

Regarding the linearity, participants appreciated the fact that the separation of units and the table of contents provides the opportunity to use units “independently”. However, “once within a unit, moving to another unit requires quite a bit of navigation”. “The lessons follow a very linear structure, the only non-linear aspect is the ability to backtrack through pages.

Feedback on Content

Overall, as depicted in Figure 5-15, the feedback on the content is very positive. Participants felt the content was, structured, concise and balanced. The language is clear and precise.

The five most negative items all refer to animations. Some of these negative ratings are based on the fact that a number of users had difficulties in accessing the animations due to technical problems. In debriefing sessions participants reported on unscalable windows. In other cases browser or proxy settings blocked loading the

animations. While these are obviously technical problems that can be overcome, it becomes clear from the comments that users are not satisfied with the animations themselves (see below).

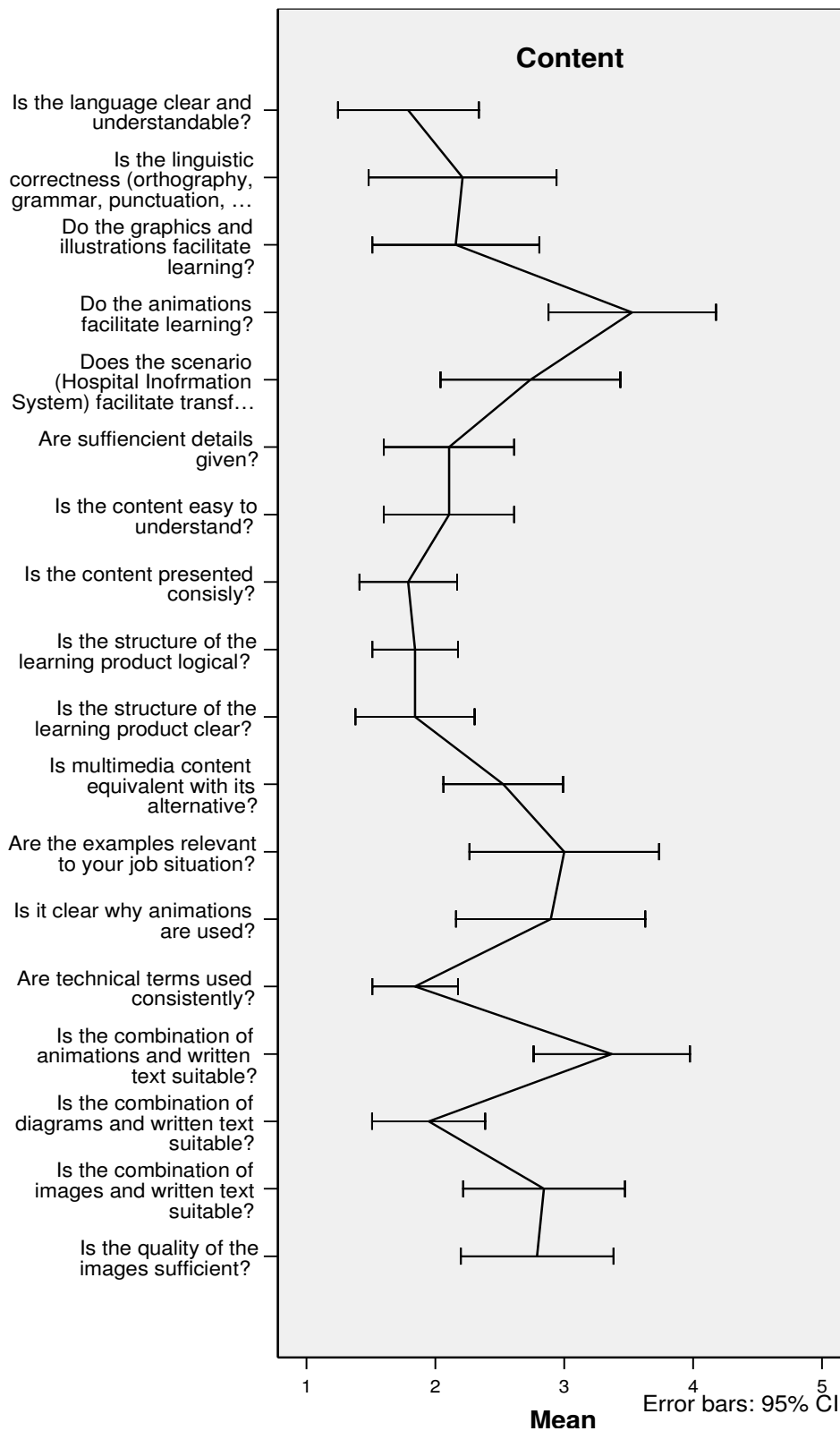


Figure 5-15: Results of Content

While the ratings are quite positive, some participants commented critically about a number of content aspects.

The language was perceived as being “used consistently”

While participants appreciated the “balanced ratio” and quality of images, they perceived images as “of little relevance”. Diagrams seem to “explain very well”, defining “UML terminologies and concepts ... clearly”.

However, a number of participants identified typological and grammatical errors. Irish participants would prefer British English as opposed to American spelling. The level of detail given resulted in mixed feedback. While some participants explicitly said sufficient details were given, others thought that the course “never really goes beyond basic definitions and repetitive applications”. Also the terminology was perceived as “foreign to a non-programmer” and should have “the layman’s term with it too”.

The animations were seen critically both from a content and a entertainment perspective. The fact that most animations consist of diagrams or figures that are displayed step by step, explained by audio was referred to as “boring”. Dynamic, interactive animation or videos were requested. Other users preferred to work with text only. The Hospital example seems to be useful for explanation purposes, but seems to have low relevance for the job situation of many participants. The scenario fails to illustrate the software life-cycle (as opposed to UML concepts). While some participants praised the consistent hospital examples, some felt it was too much repetition.

Feedback on Exercises and Tests

Participants appreciated very much the fact that there are exercises at the end of each unit to test the acquired knowledge. According the quantitative feedback, there seem to be sufficient exercises of appropriate difficulty level (Figure 5-16).

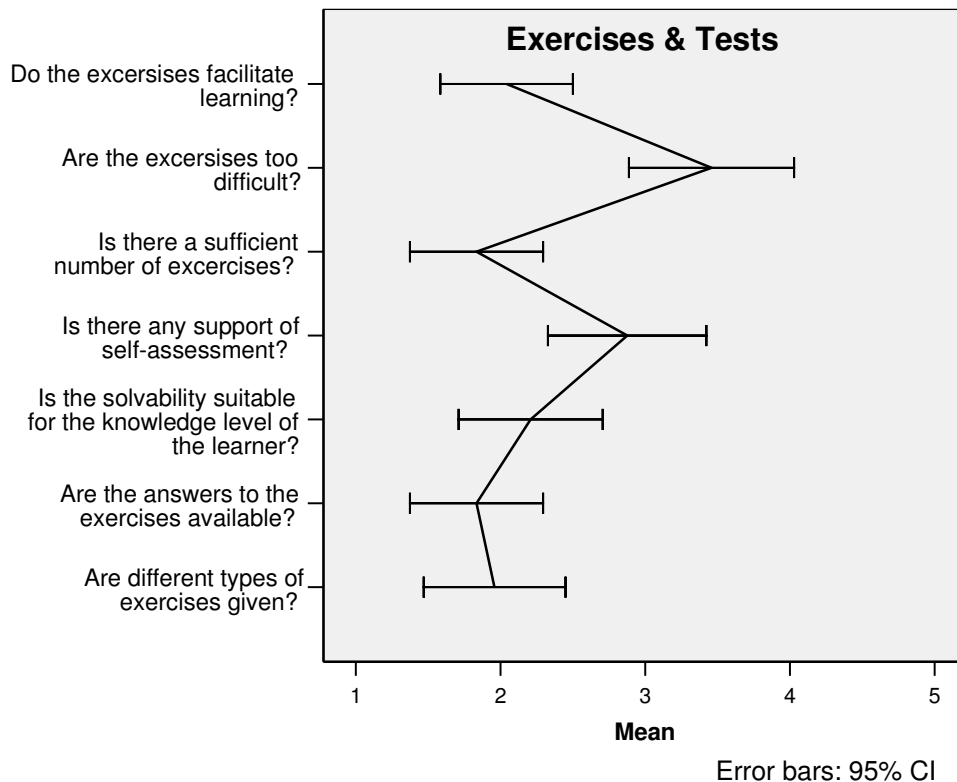


Figure 5-16: Results of Exercises and Tests

In contrary to the above figures, many participants commented in a negative way on the exercises. Again, due to technical difficulties, some exercises did not work for some users.

Some exercises were perceived as very difficult and advanced while others were too simplistic. This wide range could possibly be seen as a feature rather than a limitation. Some participants asked for a higher number of exercises in order to improve the opportunities for self-assessment. Some exercises could also be improved by providing more detailed feedback, explaining why an answer was incorrect and referring to relevant sections in the course.

Feedback on Navigation and Orientation

The prototype seems to support navigation and orientation properly. The index seems to be helpful and the clear structure (see above) supports orientation during learning.

The fact that the author of units was not always obvious can be regarded as minor problem, as long as all content originates from the same source. The low rating of the search function reflects again, that participants would have preferred a more flexible access to the content.

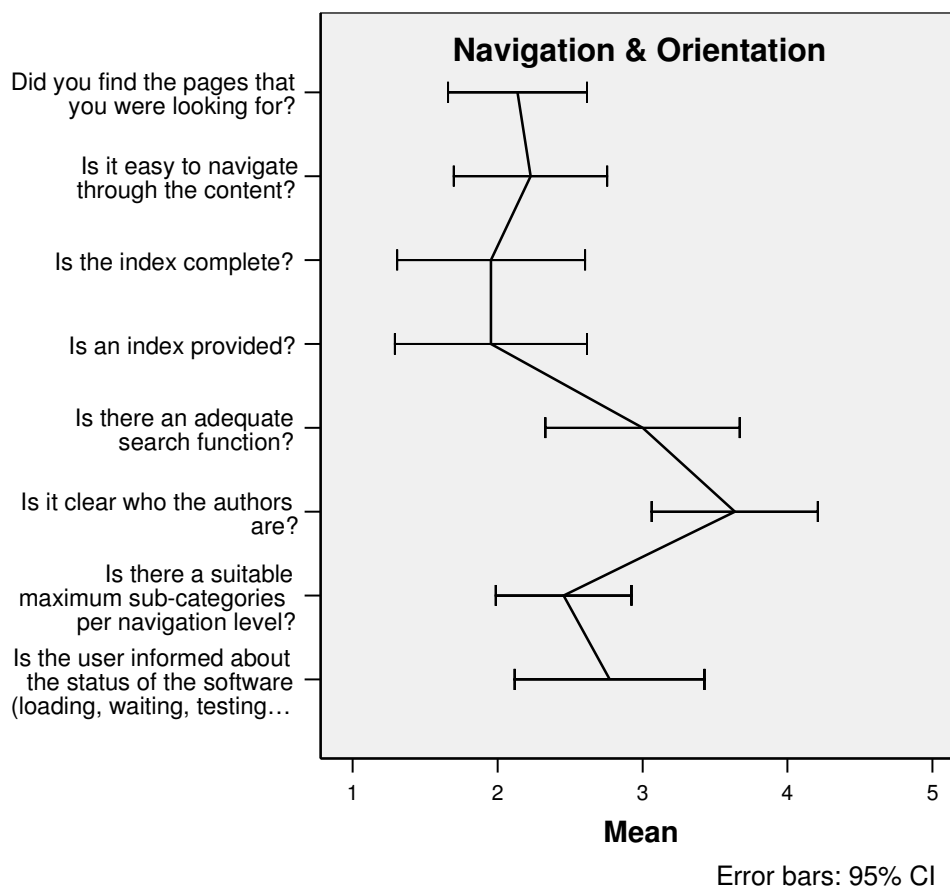


Figure 5-17: Results of Navigation & Orientation

Some participants found it “difficult to navigate back to the 'top level'”, and “navigation when at a granular low level can get difficult”.

Title: Probabilistic modelling of learning style and application to the adaptation of educational content indexed by ontology

Abstract

This thesis deals with adaptive teaching systems. The research question is how to adapt pedagogical activities to the preferred learning mode of a student. The scientific objectives are:

- modelisation of student's learning preferences
- modelisation of adaptive learning contents of a given domain
- establishing an adaptation strategy that maps preferences to contents in order to recommend an appropriate teaching method

In order to reach the first objective, the thesis studies the learning style of Felder. An empirical study to derive a dependency model between the style, the pedagogy, and the student behaviour has been conducted. Results led to creating a probabilistic preference model. A two-stage method to learn and reinforce the model is developed. Two implementations are proposed: a bayesian network and an SVM classifier. The content is represented using an ontology that combines the domain, the pedagogy, and the physical resources. An adaptation strategy centered around four dimensions is presented. This consists of searching the content to retrieve the most semantically pertinent pedagogical sequence given the student preferences. The search implements an original semantic similarity measure.

This work significantly impacted two European research projects. The production and structuration method designed in this thesis and based on SCORM has been applied in the Leonardo Da Vinci project called UP2UML. The student modeling approach serves currently our research on user profiling in the KPLAB project.

Keywords:

Probabilistic student model, Bayesian network, support vector machine, learning style, Adaptation, Knowledge model, Preferences model, Ontology

Auteur	Christophe Piombo
Titre	Modélisation probabiliste du style d'apprentissage et application à l'adaptation de contenus pédagogiques indexés par une ontologie
Directeur de thèse	Alain Ayache
Date de soutenance	30 octobre 2007 à Toulouse

Résumé

Cette thèse s'inscrit dans le cadre général des systèmes d'enseignement adaptatifs. La problématique traitée est l'adaptation de l'activité pédagogique au mode d'apprentissage préféré de l'élève. Les travaux réalisés ont eu pour objectifs de :

- modéliser les préférences d'apprentissage de l'élève
- modéliser les contenus pédagogiques du domaine à enseigner
- proposer une stratégie d'adaptation qui rapproche les préférences des contenus afin de proposer une méthode pédagogique appropriée.

Pour atteindre le premier objectif, la thèse étudie le style d'apprentissage de Felder. Une étude empirique pour établir un modèle de dépendance entre style, pédagogie, et comportement de l'élève a été réalisée. Les résultats ont permis d'établir un modèle de préférences probabiliste. Une méthode en deux étapes pour apprendre ce modèle puis le renforcer est développée. Deux implantations sont proposées : un réseau bayésien et une machine à vecteurs de support. Le contenu quant à lui est modélisé en utilisant une ontologie combinant le domaine, la pédagogie, ainsi que les ressources physiques. Une stratégie d'adaptation structurée sur quatre dimensions est présentée. Celle-ci consiste à rechercher dans le contenu la séquence pédagogique sémantiquement pertinente pour les préférences de l'élève. La recherche s'appuie sur une mesure de similarité sémantique qui est établie.

Ce travail a eu un impact sur deux projets européens. En effet, la méthode de production et structuration des contenus, basée sur SCORM, qui est proposée a servi pour le projet UP2UML. L'approche de modélisation de l'élève sert aux recherches sur le profilage dans le projet KPLAB.

Mots clés

Modèle probabiliste d'élève, Réseaux Bayésiens, machine à vecteurs de support, Styles d'apprentissage, Adaptation, Modèle de connaissances, Modèle de préférences, Ontologie