

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des Sciences de l'Ingénieur
Département d'Informatique

Année : 2008/2009

THESE

présentée en vue de l'obtention du diplôme de DOCTORAT
de l'université Badji-Mokhtar d'Annaba

Construction d'un Environnement Pédagogique Adaptatif basé sur les Modèles et Techniques du Web Sémantique

Option
Informatique

Par
Amel YESSAD

Devant le Jury

Pr. Mahmoud BOUFAIDA (Professeur, LIRE, Université de Constantine)
Pr. Mohamed Tayeb LASKRI (Professeur, LRI, Université d'Annaba)
Pr. Abdelkader BENYETTOU (Professeur, Université USTOran)
Dr. Tahar BENSEBAA (Maître de conférences, LRI, Université d'Annaba)
Dr. Amar BALLA (Maître de conférences, INI, Alger)
Dr. Catherine FARON-ZUCKER (Maître de conférences, I3S, Sophia Antipolis)

Président
Directeur de thèse
Examineur
Examineur
Examineur
Invitée

**A mes parents,
A ma famille,
A Réda.**

Remerciements

Je tiens tout d'abord à remercier les membres du jury :

Med Tayeb Laskri, Professeur d'Informatique à l'Université Badji Mokhtar- Annaba, pour avoir accepté d'être mon directeur de recherche, pour l'intérêt qu'il a accordé à cette recherche, pour ses encouragements et son soutien,

Mahmoud Boufaïda, Professeur d'Informatique à l'Université Mentouri-Constantine, pour avoir accepté de présider mon jury de thèse,

Abdelkader Denyettou, Professeur d'Informatique à l'Université USTOran, Tahar Bensebaa, Maître de Conférences à l'Université Badji Mokhtar-Annaba et Amar Balla, Maître de conférences à l'INI-Alger qui ont accepté d'être les examinateurs de ma thèse et pour l'intérêt qu'ils ont porté à ce travail,

Je remercie tout particulièrement Catherine Faron, Maître de conférences à l'Université de Nice-Sophia Antipolis et membre associée de l'équipe Edelweiss à l'INRIA-Sophia Antipolis pour son soutien infatigable, son écoute permanente et ses précieuses relectures du manuscrit de thèse et des différents papiers de recherches.

Je remercie et je dédie ce mémoire à notre regrettée Rose Dieng, une grande dame de la recherche, pour l'accueil qu'elle m'avait réservé dans l'équipe Edelweiss-INRIA, pour son enthousiasme pour cette recherche et ses encouragements.

Je remercie Lamia Yessad pour son aide et son expertise lors de la phase expérimentation.

Je remercie Réda Nouacer pour ses conseils et son aide qui m'ont apporté la visibilité nécessaire à mon travail.

Je remercie tous les membres de l'équipe Edelweiss pour leur accueil, leurs orientations constructives et encouragements.

Je remercie les professeurs Danièle Hérin, responsable de l'équipe « Web sémantique et e-Learning » du laboratoire LIRMM de Montpellier-France, et Imad Saleh, responsable de l'équipe « Hypermédia » du laboratoire Paragraphe de l'université Paris 8, pour leur accueil pendant mes stages de recherches.

Je remercie aussi :

Les relecteurs d'articles, qui par leurs remarques constructives m'ont faite avancer.

Les ingénieurs Fayçal Hamdi, Moufid Souilah et Faiz Ferhat pour leur précieuse aide dans le cadre du projet OrPAF.

Les étudiants des promotions d'ingénieurs 2008 et 2009 de l'Université du 20 Août 1955-Skikda qui ont participé activement aux expérimentations de ce travail.

ملخص:

حالياً، هناك طلب كبير من أجل توفير أنظمة إعلام تتيح للأشخاص التكوين المستمر مدى الحياة.

الهدف من هذه المذكرة هو مساعدة الشركات والجامعات من أجل توفير محيط إعلامي يسمح لها بتكوين يتناسب مع القدرات المختلفة، الأهداف البيداغوجية والقيود الزمنية للأشخاص وذلك بأثمان معقولة.
لتحقيق هذا الهدف، أنجزنا النظام الإعلامي OrPAF.

Résumé

Aujourd'hui, il existe une forte demande pour des systèmes qui permettent à des apprenants avertis et autonomes de se former tout au long de la vie. Notre travail de thèse concerne l'exploration des moyens techniques et pédagogiques permettant de monter efficacement des formations en ligne et personnalisées. Précisément, nous travaillons sur la construction d'un environnement pédagogique adaptable à différentes demandes de formation, c'est-à-dire, à différents objectifs et profils apprenants, et à différents domaines de formation et stratégies pédagogiques.

Sur le plan de la pédagogie, notre objectif est d'offrir un environnement constructiviste centré sur l'apprenant, qui assiste les apprenants dans la réalisation de leurs objectifs d'apprentissage tout en leur permettant d'être libres de décider de leur rythme et de la manière d'appréhender leurs activités pédagogiques.

Sur le plan technique, cet environnement se base sur (i) les modèles et les technologies du Web sémantique et particulièrement les ontologies pour la description formelle, partagée et explicite des connaissances relatives à l'apprentissage (connaissances sur le domaine, la pédagogie et l'apprenant). D'une part, les ontologies améliorent l'interopérabilité et l'adaptation des composants du système pédagogique. D'autre part, elles sont utilisées pour annoter les ressources pédagogiques du Web et faciliter leur réutilisation dans le contexte d'une formation donnée, (ii) la technologie des hypermédias adaptatifs pour la présentation des cours et la mise en œuvre d'une approche constructiviste de l'apprentissage.

Mots clés : Web sémantique, ontologie, hypermédia adaptatif, ressource pédagogique, annotation sémantique

Table des matières

| | |
|---|-----------|
| Introduction Générale..... | 10 |
| Problématique de la recherche | 10 |
| Domaine de la recherche..... | 11 |
| Cadre pratique de la recherche..... | 12 |
| Plan du mémoire..... | 13 |
| Partie I. Etat de l'art..... | 15 |
| 1. Apprentissage en ligne | 16 |
| 1. De la formation à distance aux systèmes pédagogiques en ligne..... | 17 |
| 1.1. Origines historiques..... | 17 |
| 1.2. Définition(s) de la formation à distance | 18 |
| 1.3. Apports de la formation à distance pour le processus d'apprentissage | 20 |
| 1.4. Classification des systèmes de formation à distance | 21 |
| 1.5. Positionnement de notre approche | 22 |
| 2. Ingénierie pédagogique pour les systèmes pédagogiques | 23 |
| 2.1. Design pédagogique | 23 |
| 2.2. Génie logiciel | 24 |
| 2.3. Ingénierie des connaissances (cognitive) | 24 |
| 3. Personnalisation dans les systèmes pédagogiques | 25 |
| 4. Approche constructivisme pour les systèmes pédagogiques personnalisés | 26 |
| 4.1. Principes du constructivisme..... | 26 |
| 4.2. Positionnement de notre approche | 28 |
| 2. Systèmes pédagogiques adaptatifs..... | 32 |
| 1. Caractérisation des systèmes pédagogiques adaptatifs | 33 |
| 1.1. Modèle de contenu | 35 |
| 1.1.1. Structure logique | 35 |
| 1.1.2. Structure sémantique..... | 35 |
| 1.2. Modèle de l'apprenant..... | 36 |
| 1.2.1. Nature des connaissances..... | 36 |
| 1.2.2. Gestion des connaissances | 37 |
| 1.3. Modèle pédagogique | 38 |
| 1.3.1. Niveaux d'abstraction dans les démarches pédagogiques..... | 38 |
| 1.3.2. Modèles de démarches pédagogiques | 39 |
| 1.4. Modèle d'adaptation..... | 40 |
| 1.4.1. Méthodes d'adaptation | 41 |
| 1.4.2. Techniques d'adaptation | 42 |
| 2. Etude de quelques systèmes pédagogiques adaptatifs..... | 45 |
| 2.1. Description des systèmes | 45 |
| 2.2. Analyse des systèmes..... | 45 |
| 2.3. Problèmes des systèmes pédagogiques adaptatifs actuels | 47 |
| 3. Web sémantique et e-Learning..... | 49 |
| 1. Web sémantique (WS)..... | 50 |
| 2. Ontologies..... | 51 |
| 2.1. Définition des ontologies | 51 |
| 2.2. Intérêt des ontologies | 52 |
| 3. Langages du Web sémantique | 53 |
| 3.1. Les langages du W3C..... | 53 |
| 3.2. Le langage RDF (Resource Description Framework) | 54 |
| 3.3. Le langage RDFS (RDF Schema) | 55 |
| 3.4. Le langage OWL | 57 |
| 4. Web sémantique & e-Learning | 57 |
| 5. Métadonnées et e-Learning..... | 59 |
| 6. Exemple de Standards de métadonnées : Learning Object Metadata (LOM)..... | 60 |
| 6.1 Descripteurs du modèles LOM..... | 61 |
| 6.2. Limitations du modèle LOM..... | 64 |
| Partie II : Un environnement pédagogique adaptatif: le système OrPAF .. | 66 |
| 1. Cadre d'analyse et méthodologie de développement..... | 67 |

| | |
|---|------------|
| 2. Problématique de recherche | 67 |
| 3. Scénarios d'utilisation de l'environnement pédagogique adaptatif..... | 69 |
| 4. Environnement pédagogique adaptatif: le système OrPAF | 70 |
| 4. Représentation des connaissances dans OrPAF..... | 71 |
| 1. Modélisation des connaissances dans OrPAF (Yessad et al., 2008b) | 72 |
| 2. Acteurs du système OrPAF..... | 73 |
| 3. Ontologies du système OrPAF..... | 75 |
| 3.1. Ontologie du e-Learning | 76 |
| 3.1.1. Définition informelle..... | 76 |
| 3.1.2. Définition formelle..... | 76 |
| 3.1.3. Langage de description de l'ontologie du e-Learning | 78 |
| 3.1.4. Exemples de sérialisation de l'ontologie du e-Learning dans OWL | 78 |
| 3.2. Modèles du e-Learning..... | 81 |
| 3.2.1. Modèle du domaine..... | 81 |
| 3.2.2. Modèle pédagogique | 82 |
| 3.2.3. Modèle apprenant..... | 83 |
| 3.2.4. Langage de description des différents modèles..... | 85 |
| 3.2.5. Exemples de sérialisation des modèles du e-Learning par des annotations RDF..... | 86 |
| 4. Développement des ontologies : Environnement KAON (Maedche et al, 2000)..... | 88 |
| 4.1. Acquisition des connaissances du domaine avec TextToOnto..... | 90 |
| 4.2. Edition des ontologies avec OntoEdit | 91 |
| 5. Génération de cours hypermédias adaptatifs dans OrPAF..... | 94 |
| 1. Structure adaptative du cours hypermédia | 96 |
| 1.1. Définition de la carte conceptuelle..... | 96 |
| 1.2. Construction de la carte conceptuelle..... | 96 |
| 1.3. Filtres du modèle du domaine | 97 |
| 1.4. Présentation adaptative de la carte conceptuelle | 100 |
| 2. Contenu adaptatif du cours hypermédia..... | 102 |
| 2.1. Recherche distante des ressources pédagogiques..... | 103 |
| 2.2. Annotation conceptuelle des ressources pédagogiques | 105 |
| 2.3. Recherche locale des ressources pédagogiques..... | 107 |
| 3. Evaluation de la pertinence des ressources pédagogiques | 109 |
| 3.1. Analyse Sémantique | 109 |
| 3.1.1. Démarche adoptée pour l'évaluation de la pertinence sémantique | 109 |
| 3.1.2. Calcul du poids d'une notion pédagogique | 110 |
| 3.1.3. Calcul de la pertinence sémantique (Semantic relevance ou SR) | 112 |
| 3.2. Analyse de personnalisation..... | 114 |
| 3.2.1. Pertinence pédagogique d'une ressource pédagogique (PR) :..... | 114 |
| 3.2.2. Adaptation d'une ressource pédagogique aux préférences de l'apprenant..... | 116 |
| 4. Annotation des tests | 117 |
| 6. Réalisation de l'environnement pédagogique adaptatif OrPAF | 120 |
| 1. Architecture de l'environnement OrPAF..... | 121 |
| 2. Environnement proposé à l'apprenant | 123 |
| 3. Extraction de données RDF | 124 |
| 3.1. Moteur de recherche sémantique Corese..... | 124 |
| 3.2. Utilisation de Corese dans le système OrPAF..... | 126 |
| 4. Session d'apprentissage d'un apprenant | 128 |
| 4.1. Construction d'une ACM..... | 129 |
| 4.2. Affichage d'une ACM..... | 131 |
| 4.3. Evolution d'une ACM..... | 132 |
| 4.4. Sauvegarde/Restauration d'une ACM..... | 133 |
| 5. Apprentissage centré sur l'apprenant..... | 137 |
| 6. Base de données du système OrPAF..... | 138 |
| 7. Recherche des ressources pédagogiques distantes..... | 138 |
| 7.1. Interrogation manuelle du corpus ARIADNE | 139 |
| 7.2. Interrogation automatique du corpus ARIADNE..... | 141 |
| 8. Recherche des ressources pédagogiques locales | 142 |
| 7. Evaluation et retour d'expérience | 145 |
| 1. Evaluation du système OrPAF..... | 146 |
| 1.1. Test d'utilisabilité de l'environnement par des apprenants | 146 |

| | |
|---|------------|
| 1.1.1. Principe et critères du test d'utilisabilité | 146 |
| 1.1.2. Protocole du test d'utilisabilité : | 147 |
| 1.1.3. Résultats et dépouillement | 149 |
| 1.2. Test sur le développement des aptitudes conceptuelles chez les apprenants et la pertinence des ressources pédagogiques ressentie par les apprenants..... | 153 |
| 2. Evaluation de la mesure de calcul de la pertinence sémantique d'une ressource pédagogique | 155 |
| 2.1. Comparaison avec les résultats des experts humains | 156 |
| 2.1.1. Données de l'expérience | 156 |
| 2.1.2. Déroulement de l'expérience | 156 |
| 2.1.3. Résultats et discussion | 158 |
| 2.2. Performances de la mesure SR d'une ressource pédagogique..... | 159 |
| 2.2.1. Données de l'expérience | 159 |
| 2.2.2. Résultats et discussion | 161 |
| Conclusion Générale | 166 |
| Bilan..... | 166 |
| Perspectives de recherche | 168 |
| Références bibliographiques..... | 169 |

Table des illustrations

| | |
|--|-----|
| Figure 1 : Niveaux d'abstraction dans les démarches pédagogiques (Allert et al., 2002)..... | 39 |
| Figure 2 : Modèle d'activités pédagogiques « donner exercice » (Vassileva, 1997) | 40 |
| Figure 3 : Méthodes et techniques d'adaptation des systèmes pédagogiques adaptatifs | 44 |
| Figure 4 : Les couches du Web sémantique (W3C) | 54 |
| Figure 5 : Exemple d'un graphe RDF | 55 |
| Figure 6: Phases de développement d'un environnement pédagogique (Lester et al., 2005)..... | 67 |
| Figure 7: Les acteurs et l'organisation des connaissances dans OrPAF..... | 75 |
| Figure 8: Un extrait de l'ontologie du e-Learning..... | 76 |
| Figure 9 : Un extrait du modèle du domaine de l'algorithmique | 82 |
| Figure 10 : Un extrait du modèle pédagogique | 83 |
| Figure 11 : Un extrait du modèle apprenant..... | 84 |
| Figure 12: Extrait de l'intégration des connaissances ontologiques..... | 86 |
| Figure 13: Architecture de l'environnement TextToOnto..... | 89 |
| Figure 14 : Extraction de termes candidats dans KAON à partir d'un document textuel | 91 |
| Figure 15 : Construction de l'ontologie du e-Learning avec OntoEdit | 92 |
| Figure 16 : Moteur d'Adaptation du cours dans OrPAF | 95 |
| Figure 17 : Filtres du modèle du domaine : (a) une CM simple, (b) une CM hiérarchique et (c) une CM relationnel (m=2)..... | 99 |
| Figure 18 : une ACM du domaine de l'algorithmique | 101 |
| Figure 19 : Diagramme d'états représentant les changements d'état d'une notion de l'ACM..... | 102 |
| Figure 20: Extension de l'ontologie du e-Learning pour définir un schéma pour l'annotation des ressources pédagogiques..... | 106 |
| Figure 21 : Annotations conceptuelles de la ressource http://www-sop.inria.fr/.../Course1 | 107 |
| (Le domaine de l'Algèbre) | 107 |
| Figure 22 : Exercice de placement de notions de l'algèbre..... | 112 |
| Figure 23: Assignement de poids relatifs aux notions reliées à la notion OrderRelation dans le domaine de l'algèbre | 114 |
| Figure 24: Assignement de poids relatifs aux activités pédagogiques reliés à l'activité pédagogique <i>Exercice</i> dans le modèle pédagogique..... | 116 |
| Figure 25: Extension de l'ontologie du e-Learning pour définir un schéma pour l'annotation des tests pédagogiques..... | 118 |
| Figure 26 : Architecture de l'environnement OrPAF..... | 122 |
| Figure 27 : Scénarios d'utilisation du système OrPAF par l'apprenant | 123 |
| Figure 28 : Environnement graphique du système OrPAF | 124 |
| Figure 29: Architecture 3-tiers de Corese (Corby et al., 2006) | 126 |
| Figure 30: Le diagramme de séquence des scénarios d'utilisation d'une ACM. | 129 |
| Figure 31: Création d'une ACM relationnelle (un parcours étendu) avec la notion cible <i>Type</i> (Le contexte de formation affiché est celui de la notion cible <i>Procedure</i> (ACM Hiérarchique))..... | 130 |
| Figure 32 : Architecture de la bibliothèque JUNG..... | 131 |
| Figure 33 : Annotation graphique des notions utilisées dans une ACM | 132 |
| Figure 34 : une ACM | 135 |
| Figure 35 : Tests associés à la notion <i>Statement</i> du modèle du domaine de l'algorithmique..... | 137 |
| Figure 36 : Le GDF des sessions des apprenants | 138 |
| Figure 37 : KPS Client dans l'architecture d'ARIADNE | 139 |
| Figure 38 : Recherche avancée de ressources pédagogiques dans ARIADNE | 140 |
| Figure 39 : Recherche fédérée de ressources pédagogiques dans ARIADNE | 140 |
| Figure 40 : Classification visuelle des ressources pédagogiques par domaine d'appartenance dans ARIADNE | 141 |
| Figure 41: Visualisation d'une ressource pédagogique par l'apprenant..... | 143 |
| Figure 43: Un contexte de formation du domaine de l'algorithmique | 160 |
| Figure 44: Variation des performances de la mesure SR en fonction du paramètre QS | 163 |

Introduction Générale

Problématique de la recherche

Aujourd'hui, les entreprises ont besoin régulièrement de former leurs employés car elles évoluent dans un environnement socio-économique dont les mots clés sont réactivité, compétitivité et performance. Or, les entreprises hésitent à investir dans les environnements logiciels de formation car les processus de développement de ces environnements sont complexes, coûteux et sans garantie de résultats. En effet, le même effort de développement est déployé à chaque fois qu'un nouveau besoin de formation apparaît. Dans ce contexte, pourquoi n'existe-t-il pas des environnements pédagogiques qui fonctionnent selon le principe du « plug and play » avec un minimum d'effort dans l'adaptation des composants logiciels aux différents besoins de formation, des environnements qui permettent à des apprenants autonomes d'atteindre efficacement leurs objectifs de formation ?

Nous avons également relevé une problématique majeure dans le domaine des technologies de l'information et de la communication appliquées à l'éducation (TICE) et qui représente un frein à l'efficacité de l'apprentissage : la quasi absence d'intégration des modèles pédagogiques dans les environnements de formation. Nous notons qu'à travers les décennies un changement de paradigme s'est opéré dans les pratiques éducatives : on est passé d'une vision centrée sur l'enseignant à une vision centrée sur l'apprenant. On est passé des environnements de formation aux environnements d'apprentissage. Il ne s'agit plus de délivrer des contenus mais d'aider l'apprenant à acquérir des connaissances, l'apprenant est véritablement acteur de la construction de ses connaissances, et le rôle joué par ses connaissances antérieures est primordial. Cette vision constructiviste de l'apprentissage a donné naissance à de nouveaux types d'environnements, constructivistes eux aussi (Tardif 1998, Jonassen & Marra 2001): dans ce type d'environnement, on fournit à l'apprenant des outils cognitifs informatisés pour l'aider à construire ses connaissances, on le guide dans son apprentissage tout en lui donnant une part importante d'initiative.

L'objectif de ce travail de thèse est de proposer des approches de construction d'un environnement pédagogique léger qui permet la génération de cours personnalisés destinés à des apprenants avec des profils différents et qui répond aux besoins de formation des entreprises. Par environnement léger nous voulons dire que le montage de la formation se fait avec un minimum d'effort et de coût et dans un minimum de temps. Pour atteindre cet

objectif, nous avons choisi de travailler dans un contexte particulier avec trois lignes directrices :

D'abord, les contenus pédagogiques sont recherchés du Web, c'est-à-dire, nous réutilisons les ressources pédagogiques qui existent déjà dans d'autres systèmes de formation ou corpus du Web. Ce choix nous permet d'économiser le temps et l'effort nécessaires au développement de contenus pédagogiques. Toutefois, cette économie n'est possible que si les moyens qui favorisent la réutilisation des ressources pédagogiques du Web sont mis en œuvre. Dans cette perspective, nous limitons notre recherche aux ressources pédagogiques du Web annotées par le standard LOM-IEEE¹.

Ensuite, l'approche pédagogique que nous adoptons pour la transmission des connaissances de l'expert à l'apprenant est l'approche constructiviste. Notre choix s'est porté sur cette approche car elle est bien adaptée pour des apprenants autonomes et motivés, par exemple, des employés qui ont besoin de formations « just in time » pour accomplir des tâches au sein de l'entreprise. Notre approche favorise la transmission du savoir par la découverte et l'acquisition de nouvelles notions et leur assimilation en vue de leur utilisation.

Finalement, les connaissances de l'environnement pédagogique que nous proposons sont décrites de manière consensuelle grâce aux modèles et techniques du Web sémantique et ceci dans un souci de partage et de réutilisation de ces connaissances. Le remplacement d'un module par un autre ne remet pas en cause l'organisation globale de l'environnement pédagogique. Nous avons proposé un environnement pédagogique réutilisable pour différents domaines de formation, pour différentes stratégies pédagogiques et pour différents profils d'apprenants.

Domaine de la recherche

Notre travail s'inscrit dans le domaine très actif des technologies de l'information et de la communication appliquées à l'éducation (TICE) et plus précisément dans le domaine des environnements informatiques d'apprentissage humain (EIAH). Parmi les deux principales orientations se développant actuellement dans ce domaine, la formation ouverte et à distance et les hypermédias éducationnels adaptatifs, notre approche combine les deux et s'inscrit dans le domaine des hypermédias adaptatifs et ouverts. Nous proposons un système de génération

¹ Learning Object Metadata (http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf)

de cours personnalisés qui réutilise des ressources pédagogiques du Web en utilisant les modèles et les techniques du Web sémantique.

Nous avons orienté notre travail vers l'apprentissage d'un type particulier de connaissances: l'apprentissage de notions, dans des domaines de connaissances scientifiques et techniques. Il nous semble qu'il y a en effet actuellement, dans un contexte d'accroissement rapide des technologies, un fort besoin de revenir dans la formation aux notions essentielles qui sous-tendent les technologies.

Cadre pratique de la recherche

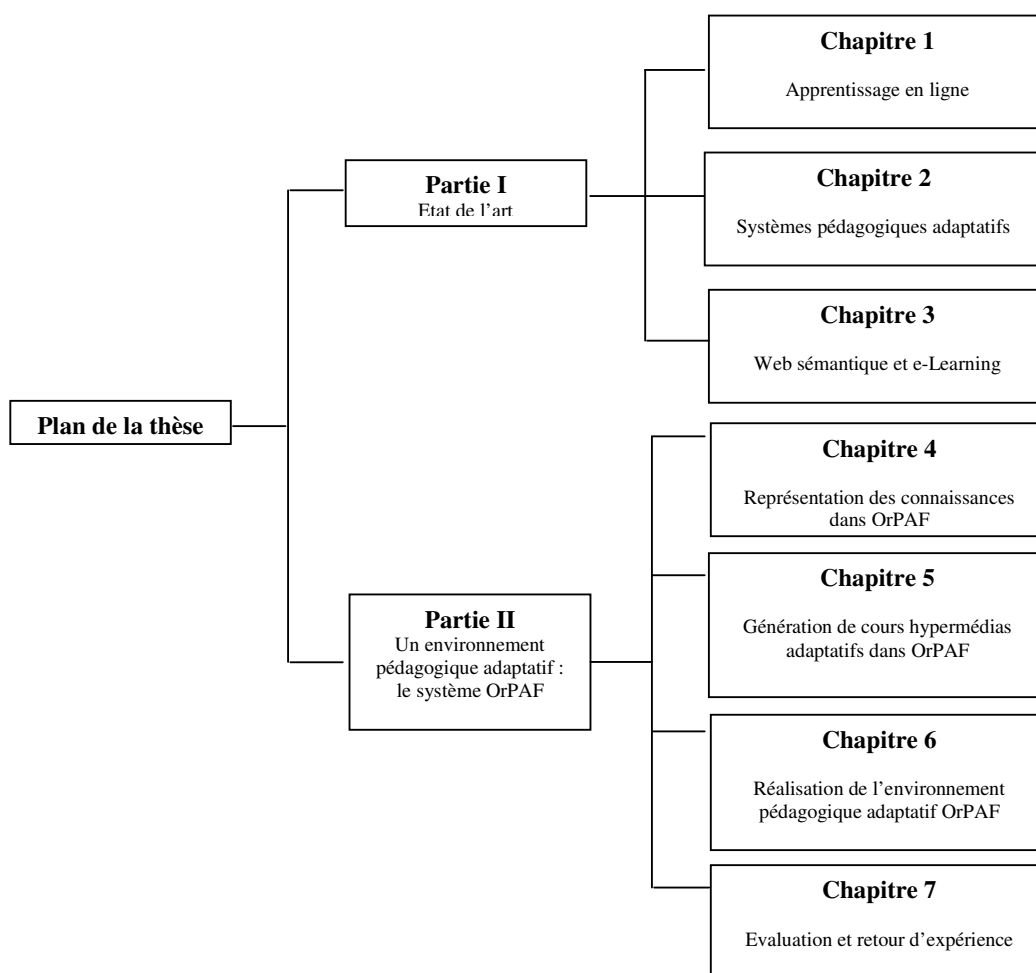
Plus concrètement, nous avons constaté une grande dynamique dans les besoins de formation d'une entreprise: d'une part les profils des employés sont différents et changeant et d'autre part, les domaines de compétences évoluent et changent en fonction de l'environnement socio-économique de l'entreprise. En d'autres termes, les employés ont régulièrement besoin de compléter, de faire évoluer et de valider leurs compétences. Particulièrement, dans les domaines scientifiques et techniques, l'appropriation conceptuelle de ces compétences est importante car elle leur permet de s'adapter efficacement aux nouvelles situations professionnelles.

Pour répondre à ce contexte réel, nous avons proposé une approche de construction d'environnements pédagogiques qui permet de :

1. représenter de manière standard et partagée les connaissances relatives à l'apprentissage, à savoir, les connaissances sur l'apprenant, les connaissances sur le domaine de la formation et les connaissances sur la pédagogie. Cette représentation permet la réutilisation avec un minimum de coût et d'effort du système pédagogique à chaque fois qu'un nouveau besoin de formation apparaît ;
2. annoter les ressources pédagogiques recherchées du Web et proposer des mesures de similarité qui permettent de calculer leurs pertinences sémantique et pédagogique pour un contexte de formation donné ;
3. générer en utilisant ces connaissances et ces mesures de similarité des cours hypermédias adaptatifs où :
 - La structure de l'hypermédia est un graphe de notions appartenant au domaine de la formation
 - Le contenu de l'hypermédia est un ensemble de ressources pédagogiques du Web, considérées comme pertinentes par le système.

Afin de tester nos approches et choix conceptuels, nous avons développé le système OrPAF, un Organisateur de Parcours Adaptatifs de Formation. Le domaine que nous avons retenu pour expérimenter OrPAF est le domaine de l’algorithmique et des langages de programmation. Nous avons travaillé lors des expérimentations du système OrPAF avec des experts du domaine de l’algorithmique et des langages de programmation et des étudiants initiés dans ce domaine qui préparent le diplôme d’ingénieurs d’état en Informatique à l’université du 20 Août 1955 (Algérie).

Plan du mémoire



Cette thèse est composée de deux parties.

La première partie présente un état de l’art relatif à notre problématique de recherche. Elle est structurée en trois chapitres :

Le premier chapitre traite de la formation à distance et de son évolution vers des systèmes pédagogiques en ligne. Nous présentons également certains aspects de l'ingénierie pédagogique et de la théorie constructiviste.

Le deuxième chapitre porte sur les systèmes pédagogiques adaptatifs. Nous délimitons un cadre de référence pour ce type de systèmes et présentons les méthodes et techniques utilisées pour mettre en œuvre l'adaptativité. Pour mieux positionner notre approche, nous avons étudié quelques systèmes pédagogiques adaptatifs.

Le troisième chapitre décrit l'apport du Web sémantique et des standards de l'éducation (particulièrement le standard LOM) pour l'adaptation des systèmes pédagogiques et la réutilisation des ressources pédagogiques.

La deuxième partie présente de manière exhaustive le cœur de notre travail de thèse : l'environnement pédagogique adaptatif OrPAF. Cette partie est composée de quatre chapitres.

Le quatrième chapitre décrit l'utilisation des modèles et technologies du Web sémantique pour la représentation des connaissances dans OrPAF.

Le cinquième chapitre explique comment les connaissances représentées dans OrPAF sont exploitées pour générer des cours hypermédias adaptatifs et améliorer la réutilisation des ressources pédagogiques du Web.

Le sixième chapitre décrit la réalisation de l'environnement OrPAF.

Le septième chapitre décrit l'évaluation de notre approche auprès d'experts dans le domaine de l'algorithmique et des langages de programmation et auprès d'étudiants en troisième et quatrième année de la formation d'ingénieurs d'état en Informatique de l'université 20 Août 1955. Nous détaillons les protocoles d'expérimentation, les retours d'expériences et les résultats statistiques obtenus.

Partie I. Etat de l'art

1. Apprentissage en ligne

L'apprentissage en ligne (ou le e-Learning) est la tendance actuelle des systèmes de formation à distance. La formation à distance est un concept général qui inclut les cours par correspondance et tout autre moyen d'enseignement en temps et lieu asynchrone. L'apprentissage en ligne doit son apparition à l'émergence et au succès que connaissent les Technologies de l'Information et de la Communication (TIC), le multimédia et l'Internet. Dans notre travail, nous le considérons comme une modalité pédagogique et technologique qui concerne l'enseignement supérieur, la formation continue ainsi que la formation en entreprise, c'est-à-dire, qui s'adresse à des apprenants avertis ayant une certaine autonomie dans l'organisation de leur processus d'apprentissage.

Depuis quelques années, les systèmes pédagogiques (désignés également par les termes : systèmes de e-Learning ou environnements informatiques d'apprentissage humain (EIAH)) investissent de plus en plus les institutions éducatives et les entreprises. Dans un système pédagogique le rôle de l'enseignant est marginal voire même inexistant et l'apprenant est mis au « centre » du système. Etant donné que les apprenants ont des caractéristiques différentes (niveau de connaissances, style cognitif, plateforme technologique, etc.) et dans un souci d'efficacité, les systèmes pédagogiques doivent être personnalisés, c'est-à-dire, adaptés aux caractéristiques de l'apprenant. Dans la suite de ce chapitre, nous abordons les notions liées à la formation à distance et l'évolution de celle-ci vers les systèmes pédagogiques actuels.

1. De la formation à distance aux systèmes pédagogiques en ligne

1.1. Origines historiques

Le concept de formation à distance est apparu vers le milieu du dix-neuvième siècle, faisant référence aux études par correspondance. Ce mode d'enseignement est encore en vigueur actuellement. La synonymie entre ces deux termes continue encore actuellement de prêter à confusion. Peraya (Peraya & McCluskey, 1996) présente quatre principes fondateurs de l'enseignement par correspondance et qui, à notre sens, restent valables pour les systèmes pédagogiques actuels :

- *vaincre la distance géographique*: donner la possibilité à des personnes éloignées géographiquement des établissements scolaires la possibilité de recevoir une formation ;
- *pallier les conséquences de circonstances exceptionnelles* : ces circonstances peuvent être des hospitalisations prolongées, des incarcérations, des exodes, etc. ;
- *pallier les insuffisances du réseau d'enseignement public et les inégalités d'accès à l'instruction* au moment de l'extension de l'enseignement universitaire à un plus large public dès le milieu du siècle dernier. La naissance des universités ouvertes s'inscrit dans la continuité de ce mouvement ;
- *suppléer à l'offre de formation souvent limitée*, notamment dans les matières techniques en constante évolution ;

On s'aperçoit à la lecture de ces principes que l'enseignement par correspondance est resté confiner à certaines fonctions particulières, alors que la formation à distance actuelle, et particulièrement les systèmes pédagogiques en ligne, sont appelés à jouer et jouent déjà une fonction déterminante et centrale dans la formation continue et s'insère même au sein de dispositifs de formation traditionnellement structurés en mode présentiel.

Rajoutons que c'est à la suite du développement des moyens de communication et de diffusion de l'information (dans ce cas présent, la poste) que l'enseignement par correspondance a vu le jour. De même qu'aujourd'hui, la formation à distance connaît un essor spectaculaire grâce au développement des technologies de l'information et de la communication (téléphone, radio, télévision, moyens informatisés et Internet).

Pour Henri et Kaye (Henri et Kaye, 1985), cet essor s'explique par la conjugaison de trois phénomènes: « l'évolution de la demande pour ce genre de formation, la nécessité économique de réduire de façon générale les coûts de l'éducation et finalement la pénétration des technologies de communication dans tous les secteurs d'activités, y compris celui de l'éducation ».

1.2. Définition(s) de la formation à distance

Afin de bien comprendre les systèmes pédagogiques en ligne, leur organisation, leurs apports et leurs limites, il paraît indispensable de bien comprendre le principe de la formation à distance, ancêtre direct de ce type de systèmes. Nous pensons qu'à l'avenir la formation à distance sera exclusivement en ligne.

Intuitivement, la formation à distance se comprend comme une situation de formation dans laquelle l'apprenant est séparé de l'enseignant et des autres apprenants dans l'espace et le temps. Cette notion englobe l'ensemble du processus éducatif, c'est-à-dire le pôle émetteur (enseignement par l'enseignant) ainsi que le pôle récepteur (apprentissage par l'apprenant). La formation à distance couvrirait alors l'ensemble des dispositifs techniques et des modèles d'organisation qui ont pour but de fournir un enseignement ou un apprentissage à des individus qui sont distants de l'organisme de formation prestataire du service.

Se satisfaire de cette définition serait trop simple. En effet, et presque tous les auteurs s'accordent sur ce point, il serait ambitieux de vouloir donner une définition de la formation à distance, « car ce système de formation a connu de nombreuses évolutions depuis sa naissance au milieu du siècle dernier et tout particulièrement durant ces dernières décennies » (Peraya et al., 1996). De plus, chaque pays, voire même chaque région, développe des modèles adaptés à son contexte. Les termes et leurs significations peuvent différer d'un pays à l'autre et d'une langue à l'autre. Tout ceci explique la difficulté à cerner cette notion. Nous retenons un ensemble de définitions qui se complètent pour caractériser la notion de formation à distance.

MOORE M. G. (1973)

« L'enseignement à distance peut être défini comme l'ensemble des méthodes pédagogiques par lesquelles l'acte d'enseignement est séparé de l'acte d'apprentissage incluant toutefois les méthodes réalisées en présence de l'étudiant de telle sorte que la communication entre le professeur et l'étudiant sera facilitée par l'usage du matériel imprimé, mécanique, électronique ou autres ».

PETERS O. (1973)

« L'enseignement à distance ou la formation à distance est une méthode de développement des connaissances, des habiletés et des attitudes qui est rationalisée par l'application des principes organisationnels de la division du travail aussi bien que par l'utilisation extensive des moyens techniques, spécialement dans le but de produire du matériel éducatif de grande qualité qui rend possible l'instruction d'un grand nombre d'étudiants au même moment sans

contingence géographique. C'est une forme industrialisée d'enseignement et d'apprentissage ».

HOLMBERG B. (1977)

« Le terme formation à distance recouvre les diverses formes d'étude à tous les niveaux qui ne sont pas sous la supervision immédiate et continue d'un tuteur en salle de classe ni/ou dans le même endroit mais qui, néanmoins, profitent de la planification de l'assistance d'une organisation ».

HENRI F. (1985)

« La formation à distance est le produit de l'organisation d'activités et de ressources pédagogiques dont se sert l'apprenant, de façon autonome et selon ses propres désirs, sans qu'il lui soit imposé de se soumettre aux contraintes spatio-temporelles ni aux relations d'autorité de la formation traditionnelle. Plus spécifiquement, elle se définirait comme une formule pédagogique au potentiel accru, qui permet à l'étudiant de redéfinir son rapport au savoir et d'utiliser, dans un modèle autodidactique, les ressources didactiques et d'encadrement ».

Il ressort de ces définitions que la formation à distance apparaît comme une sorte de « formule pédagogique au potentiel accru », pour reprendre les termes de Henri F. (1985). Parler de « formule pédagogique » nous paraît plus pertinent que de concevoir la formation à distance simplement comme un moyen d'enseignement. En effet, les implications de la gestion de la distance entre l'enseignant et l'apprenant – qui ne doit plus être seulement comprise en terme de distance géographique, mais bien plutôt en terme de délocalisation spatio-temporelle – sont trop conséquentes pour qu'elles n'atteignent le système lui-même. Holmberg B. (1977) mentionne cela en utilisant le terme de « planification de l'assistance d'une organisation » ; Henri F. (1985) parle elle de « l'organisation d'activités et de ressources pédagogiques » ; mais c'est Peters O. (1973) qui va le plus loin en parlant de « forme industrialisée d'enseignement et d'apprentissage ». Le choix d'une telle « formule pédagogique » provoque donc un certain nombre de changements au niveau organisationnel et pédagogique. Pour Henri et Kaye (Henri et al., 1985), en effet, « apprendre à distance, enseigner à distance, exigent une transformation radicale des pratiques et des moyens pédagogiques afin de surmonter la distance et l'isolement de l'étudiant. Il en résulte une mutation profonde dans le rôle et les tâches réservées aux enseignants ainsi que dans l'approche pédagogique désormais conçue en fonction de l'auto-apprentissage ». D'ailleurs, dans ce travail de thèse, nous concevons de nouveaux rôles et de nouvelles tâches qui incombent aux enseignants (ou aux experts d'un domaine) et nous

proposons des fonctions intégrées au système pédagogique pour assister ou aider les apprenants dans leur processus d'auto-apprentissage.

La formation à distance représente donc bien un système spécifique et a été décrite comme tel par Moore et Kearsley (Moore et Kearsley, 1996). Ce système doit être composé des cinq éléments suivants : les *sources* (servant à l'élaboration du matériel pédagogique, y compris les experts), le *design pédagogique* du matériel et de l'encadrement, la *diffusion*, *l'interaction* (entre enseignants, apprenants et personnel administratif) et *l'environnement d'apprentissage* (la classe, chez soi, le lieu de travail, le centre d'enseignement).

1.3. Apports de la formation à distance pour le processus d'apprentissage

Les caractéristiques fondatrices de la formation à distance pour permettre l'apprentissage peuvent être résumées en cinq points (Deschênes et al., 1996) :

- *L'accessibilité*

La formation à distance permet d'apprendre en proposant des situations d'apprentissage-enseignement qui tiennent compte des contraintes individuelles de chaque apprenant, c'est-à-dire, des distances spatiale, temporelle, technologique, psychosociale et socio-économique (Jacquinot, 1993) qui peuvent rendre le savoir inaccessible.

- *La contextualisation*

La formation à distance permet à l'individu d'apprendre dans son contexte immédiat, celui où habituellement les apprentissages devront être utilisés (par exemple le lieu de travail). Elle maintient ainsi un contact direct, immédiat et permanent avec les différentes composantes de l'environnement, facilitant l'intégration des savoirs scientifiques aux savoirs pratiques (Pépin, 1994) et le transfert des connaissances.

- *La flexibilité*

La formation à distance utilise des approches qui permettent à l'apprenant de planifier dans le temps et dans l'espace ses activités d'étude et son rythme d'apprentissage. De plus, elle peut proposer des activités offrant à l'apprenant des choix dans les contenus, les méthodes et les interactions et ainsi prendre en compte les caractéristiques individuelles de chacun (Deschênes, 1991 ; Moore, 1977).

- *La diversification des interactions*

En rapprochant le savoir des apprenants, la formation à distance reconnaît que l'apprentissage ne résulte pas essentiellement de l'interaction entre le professeur et l'apprenant ou entre ce dernier et d'autres apprenants, mais aussi entre l'apprenant et l'ensemble des individus qui l'entourent (famille, communauté, travail, etc) (Wagner et McCombs, 1995).

- *La désaffectivation des savoirs*

Not (cité par Tochon, 1992) écrit au sujet de l'enseignement en présence d'un professeur (ou de la relation pédagogique) : « cette médiation magistrale inaugurale...laisse l'élève à la remorque du maître biaisant ainsi les besoins du sujet aux prises avec l'objet à apprendre : les besoins qu'elle fait apparaître sont, en effet, ceux d'une adaptation aux représentations, à la pensée et aux démarches de l'enseignant et non ceux qui procèdent d'une recherche d'adaptation directe à l'objet ou à la situation qu'il doit maîtriser. ». Dans son travail sur la planification de l'enseignement, Tochon (Tochon, 1992) a constaté que « Les contenus sont formalisés de sorte à transmettre des connaissances, des cognitions, et une connaissance affective qui semble imbriquée dans la situation elle-même... En d'autres mots, la formalisation du sens paraît empreinte de la nature des échanges et de leur finalité conceptuelle ». Si on reconnaît que tout contexte a un impact sur la construction des connaissances chez l'apprenant, on peut s'interroger sur un impact affectif aussi grand provenant de l'enseignant et on peut comprendre pourquoi si souvent les apprenants n'arrivent pas à transférer les connaissances, celles-ci étant trop liées à un contrat didactique implicite entre l'enseignant et l'élève (Schubauer-Leoni & Ntamakiliro, 1994). La formation à distance, pour rapprocher le savoir des apprenants, élimine, dans la majorité des cas, la relation directe enseignant-apprenant. Elle permet donc, grâce aux technologies, la mise en œuvre de situations d'apprentissage-enseignement n'imposant pas cette médiation magistrale qui crée l'affectivation des contenus.

1.4. Classification des systèmes de formation à distance

Selon Paquette (Paquette, 2002), le terme de "formation à distance" recouvre plusieurs réalités techno-pédagogiques très différentes. Il propose six modèles qui intéressent notre recherche dans le sens où ils permettent de classifier les tentatives de médiatisation.

- *La classe technologique ouverte* : Paquette (Paquette, 2002) la décrit comme « une classe traditionnelle où un ensemble de technologies sont installées et utilisées de façon permanente ». Ces salles ont souvent été qualifiées de "laboratoire multimédia".

- *La classe technologique répartie* : c'est une classe technologique virtuelle qui est répartie en plusieurs lieux distincts. Les événements d'apprentissage se déroulent comme suit : un enseignant présente l'information en direct et les apprenants interagissent simultanément dans les différents lieux avec la salle principale dans laquelle se trouve l'enseignant.

- *L'autoformation web-hypermédia* : l'apprentissage est individualisé et réalisé par un apprenant seul qui accède à des contenus préfabriqués multimédiatisés sur Internet ou sur un

CD-ROM. Le terme "hypermédia" fait référence au matériel d'apprentissage qui utilise le principe des hyperliens pour la navigation. L'apprentissage se fait en général sans l'intervention d'un formateur et sans nécessairement de collaboration entre les apprenants. Ce modèle n'impose donc pas de contraintes de lieu ni de temps.

- *L'enseignement en ligne* : cette forme d'enseignement est accessible sur l'Internet. Un formateur le gère en effectuant des présentations et en coordonnant des interactions (échanges) de manière asynchrone avec un groupe d'apprenants. Le mode asynchrone permet aux apprenants de progresser à leur rythme. Ils peuvent aussi interagir les uns avec les autres. C'est le professeur qui gère l'avancée des activités. Selon Paquette, ce modèle qui est transmis par exemple par les universités totalement à distance devient également un modèle que l'on trouve dans les universités campus où il permet de proposer une alternative à la formation en classe de type présentielle.

- *La communauté de pratique* : elle utilise les mêmes outils de communication asynchrone que l'enseignement en ligne mais également des outils de discussion synchrone. La principale caractéristique de ce modèle est la discussion entre un groupe de spécialistes. L'enseignant est ici plutôt un "animateur télématique", car il dispose en général de moins de connaissances que les participants mais il sait comment rendre l'échange fructueux entre les participants. L'apprentissage se fait en échangeant des informations et en comparant des pratiques par l'étude de cas.

- *Le soutien à la performance* : la formation est ici individuelle. Elle se déroule en liaison étroite avec les activités de travail, soit pendant ces activités lorsque l'apprenant a besoin de formation pour avancer dans la tâche, soit après, parce que l'apprenant veut approfondir des questions qu'il s'est posé dans l'exercice d'une tâche, ou encore avant l'activité en question. Ce sont les fameuses formations "just in time" décidées en fonction des tâches à résoudre.

1.5. Positionnement de notre approche

La recherche que nous menons dans le cadre de notre thèse s'intéresse à l'autoformation web-hypermédia avec l'objectif supplémentaire de proposer un système de soutien à la performance. En effet, la formation à distance que nous mettons en œuvre est caractérisée par :

- le système pédagogique proposé est ouvert et réutilise les ressources pédagogiques du Web ;
- l'absence de la supervision d'un enseignant et l'apprenant est libre dans sa progression pédagogique tout en étant assisté par le système pédagogique. Ce dernier est conçu pour être adaptatif grâce à des méthodes d'ingénierie pédagogique ;

- le matériel pédagogique proposé à l'apprenant est un système hypermédia dans lequel la structure de navigation et le contenu sont adaptés au profil de l'apprenant ;
- le contenu pédagogique est recherché sur le Web ;
- le développement qui sous-tend notre approche s'appuie sur les standards dans le domaine de l'éducation et les modèles et les techniques de la gestion des connaissances et du Web sémantique. Ces standards, modèles et techniques permettent une adaptation et une réutilisation efficaces de notre approche à d'autres besoins de formation qui apparaissent dynamiquement (formations "just in time").

2. Ingénierie pédagogique pour les systèmes pédagogiques

La mise en place d'un processus d'apprentissage en ligne implique de nombreux investissements à tous les niveaux. Aussi, la mise en œuvre d'une stratégie organisationnelle permet d'accompagner au mieux cette démarche.

De nombreux auteurs s'accordent à dire qu'une démarche de médiatisation d'un système de formation doit être basée sur l'ingénierie pédagogique. Pour définir ce concept, nous avons eu recours à l'argumentation de Paquette (Paquette, 2002). Selon lui, l'ingénierie pédagogique est tout d'abord une « méthode soutenant l'analyse, la conception, la réalisation et la planification de la diffusion des systèmes pédagogiques, intégrant les concepts, les processus et les principes du design pédagogique, du génie logiciel et de l'ingénierie cognitive ».

Le besoin d'employer une méthodologie découle, selon Paquette (Paquette, 2002), du fait que les avancées technologiques permettent des possibilités pédagogiques beaucoup plus variées et que les systèmes pédagogiques actuels sont beaucoup plus conséquents, en termes d'outils logiciels, de documents numérisés et de services de communication, que par le passé. Ainsi, nous soutenons l'idée que développer de tels systèmes pédagogiques ne peut donc plus se faire de manière artisanale.

2.1. Design pédagogique

Pour Paquette (Paquette, 2002), le design pédagogique est « l'ensemble des théories et des modèles permettant de comprendre, d'améliorer et d'appliquer des méthodes d'enseignement favorisant l'apprentissage. Par rapport aux théories élaborées en psychologie de l'apprentissage, le design pédagogique peut être vu comme une forme d'ingénierie visant à améliorer les pratiques éducatives ». Le résultat du design pédagogique correspond à un ensemble de plans ou de devis décrivant le déroulement des activités d'apprentissage et d'enseignement.

2.2. Génie logiciel

Le génie logiciel se consacre à la conception et à la réalisation des logiciels ainsi que des systèmes d'information. De plus en plus, la taille, la complexité, le caractère critique des systèmes d'information dépassent la capacité de compréhension d'un seul programmeur. Le but premier du génie logiciel est donc de fournir des processus, des méthodes, des techniques et des outils permettant de produire des systèmes satisfaisants, efficaces, extensibles, flexibles, modifiables, portables, fiables, réutilisables, validables et compréhensibles. Au-delà de la nature complexe des systèmes pédagogiques, nous notons également leur caractère pluridisciplinaire (ils sont au croisement de l'informatique, de la psychologie, de la didactique, des sciences sociales, etc.)

Les méthodes d'ingénierie des systèmes d'information prévoient un découpage méthodologique en différentes phases correspondant à des processus dont l'exécution n'est pas forcément linéaire :

- observation critique, état des lieux ou diagnostic,
- orientations générales, établissement des principes de fonctionnement,
- définition préliminaire et appel d'offres,
- architecture du système d'information,
- programmation du système,
- simulation et qualification du système,
- implantation du système,
- exploitation du système.

L'ingénierie pédagogique s'inspire de ces principes de conception des systèmes d'information en les adaptant à la conception de systèmes pédagogiques.

2.3. Ingénierie des connaissances (cognitive)

L'ingénierie des connaissances s'est développée à la suite des systèmes experts et de l'intelligence artificielle. L'objectif des systèmes experts était de créer des systèmes informatiques permettant de représenter et de reproduire l'expertise humaine. Dans le domaine de l'éducation, l'ingénierie des connaissances a été appliquée pour le diagnostic des lacunes d'un apprenant et la proposition de stratégies permettant d'y remédier.

L'ingénierie des connaissances implique des opérations comme la détermination des connaissances, leur explicitation, leur représentation et leur formalisation dans un langage symbolique ou graphique qui facilite leur utilisation subséquente. De façon typique, l'ingénieur de la connaissance procède à des discussions avec un ou plusieurs experts de

contenu qui détiennent l'expertise qu'il désire modéliser. A l'aide de méthodes d'interview systématiques ayant pour but l'acquisition de connaissances d'un domaine, l'ingénieur raffine de plus en plus sa représentation du domaine par étapes successives jusqu'à la capter dans une forme synthétique qui sera intégrée dans un système informatique.

L'ingénierie des connaissances est donc utile à la méthode d'ingénierie pédagogique dans le sens qu'elle permet d'obtenir une vue explicite et structurée des connaissances.

Il devient ainsi plus aisé de définir le contenu, les scénarios pédagogiques et les matériels pédagogiques.

3. Personnalisation dans les systèmes pédagogiques

L'importance primordiale qu'occupe la connaissance dans la vie de chacun, au travail et en société, a produit inévitablement l'émergence d'une nouvelle façon de se former et d'apprendre : l'apprentissage tout au long de la vie ou la formation continue. En effet, la qualité de la formation des individus réduit les risques d'exclusion sociale et professionnelle de ces mêmes individus. Par ailleurs, l'évolution rapide des technologies de l'information et de la communication, du multimédia et de l'Internet ont rendu possible l'apprentissage tout au long de la vie grâce à l'émergence d'un nouveau paradigme : l'apprentissage en ligne (ou e-Learning). Un des plus grands défis des systèmes pédagogiques en ligne est la personnalisation de l'apprentissage, c'est-à-dire, rendre les programmes de formation adaptés aux objectifs des apprenants, à leur niveau de connaissances, à leurs préférences, etc.

Dans ce contexte, au-delà de l'aspect asynchrone en temps et en lieu des systèmes pédagogiques, ils doivent également satisfaire aux besoins changeants des apprenants, à leurs connaissances, à leurs préférences, etc. L'apprentissage personnalisé représente une révolution radicale dans les théories et les technologies de l'éducation : nous sommes passés des systèmes pédagogiques interactifs traditionnels aux systèmes pédagogiques personnalisés. Nous pouvons résumer cette révolution en trois points : (i) tandis que les systèmes pédagogiques traditionnels sont de type un-à-plusieurs, les systèmes pédagogiques personnalisés sont de types un-à-un ou plusieurs-à-un (c'est-à-dire un ou plusieurs tuteurs pour un seul apprenant) ; (ii) les systèmes pédagogiques traditionnels sont conçus pour un « apprenant moyen » alors que dans les systèmes pédagogiques personnalisés l'enchaînement des unités du cours, le style d'apprentissage, etc. dépendent du profil de l'apprenant ; (iii) dans les systèmes pédagogiques traditionnels, les décisions pédagogiques, les unités d'apprentissage, etc. sont définis à priori par le tuteur alors que dans les systèmes

pédagogiques personnalisés, le système se base uniquement sur le profil de l'apprenant et ses interactions.

4. Approche constructivisme pour les systèmes pédagogiques personnalisés

Concevoir un système pédagogique personnalisé, revient à mettre l'apprenant au centre du processus d'éducation. En effet, plusieurs des pratiques dans les systèmes pédagogiques exigent d'une part, chez les apprenants, la mise en œuvre autonome d'habiletés et de stratégies visant des objectifs reliés au contenu d'un cours (Deschênes, 1991; Garrison, 1993) et d'autre part, chez les concepteurs, la planification d'un support pédagogique adéquat favorisant un apprentissage indépendant (Bolanos-Mora et al., 1992).

Les modèles théoriques classiques dans le domaine de l'éducation ont été considérés comme inadéquats pour répondre aux attentes actuelles des chercheurs et des praticiens ainsi qu'aux demandes des institutions qui désirent mettre en place un système pédagogique personnalisé. Seul le modèle constructivisme s'est révélé adapté à l'apprentissage autonome d'apprenants en leur permettant de construire de manière flexible leurs connaissances indépendamment de l'enseignant.

4.1. Principes du constructivisme

Parmi les modèles théoriques en éducation, les mécanismes constructivistes ont été, depuis longtemps, identifiés comme les assises du processus d'élaboration des connaissances (en particulier par Piaget). Par contre, leurs applications ont été marginales (Larochelle et Bednarz, 1994) et ce n'est que récemment qu'on les a repris pour structurer une solution alternative (Lebow, 1993) au behaviorisme persistant (Garrison, 1993) ou au cognitivisme trop souvent rigide (Kintsch, 1988) qui sous-tendent les pratiques pédagogiques.

Le constructivisme est un processus de « construction continue » des connaissances (Piaget, 1970). L'idée première qui sous-tend les travaux sur le constructivisme est que la connaissance n'est pas la reproduction d'une réalité qu'on appelle le savoir scientifique mais la construction de cette réalité. Trois aspects inhérents au constructivisme peuvent servir dans le contexte de l'apprentissage personnalisé (Deschênes et al., 1996) : 1) les connaissances sont construites, 2) l'apprenant est au centre du processus et, 3) le contexte d'apprentissage joue un rôle déterminant. Ces aspects montrent clairement l'intérêt que peut représenter la théorie constructiviste pour le développement des systèmes pédagogiques personnalisés.

Les connaissances sont construites

Pour un constructiviste, la connaissance est activement construite par celui qui apprend dans chacune des situations où elle est utilisée ou expérimentée. La fonction de la cognition est l'adaptation et sert à l'organisation du monde qu'on expérimente et non à la découverte de la réalité (Jegede, 1992). La connaissance résulte donc de l'activité de l'apprenant et est construite en relation avec son action et son expérience du monde (Clancey, 1991). Dans ce cas l'apprentissage n'est plus une simple transposition d'un savoir mais l'appropriation de celui-ci par l'apprenant.

L'apprenant est au centre du processus

Dans une approche académique, la connaissance est considérée comme une accumulation de données que l'on emmagasine en mémoire (Lave, 1990). Cette conception suppose que la connaissance est transmise selon un processus hiérarchique d'enseignement et est mesurée par un test : on croit qu'il n'y a pas d'apprentissage sans enseignement. Dans une approche constructiviste, au contraire, la connaissance est expérimentée à travers une activité cognitive de création de sens par l'apprenant (Jonassen et al., 1994). Le rôle joué par ce dernier devient donc primordial : l'apprentissage se réalise grâce à l'interaction que l'apprenant établit entre les diverses composantes de son environnement qui comprend les informations disponibles (savoirs scientifiques et savoirs pratiques). La nature et le type d'interactions mises en œuvre dépendent de la perception qu'a l'individu des diverses composantes.

Au plan du transfert et de l'application des connaissances, une approche constructiviste préconise l'utilisation de tâches authentiques pour l'apprenant, c'est-à-dire reliées directement au contexte concret et réel de l'objet d'apprentissage traité. Elle considère que l'objectif fondamental est la réalisation de la tâche globale dans toute sa complexité et non la réalisation décontextualisée de petites étapes visant qu'à la fin, la tâche complète soit réalisée puis reprise dans un contexte concret ou réel (Lave, 1990; Perkins, 1991). Pour y arriver, les apprenants doivent jouer un rôle plus important dans la gestion et le contrôle de leur apprentissage (Lebow, 1993). Cela peut se réaliser grâce à la mise en place de situations assurant que les apprenants utilisent leurs connaissances, en s'appuyant sur des processus de résolution de problèmes et en leur fournissant des activités de type métacognitif (Lebow, 1993). Mais surtout, cela peut être facilité en maintenant les apprenants dans leur contexte immédiat au moment de l'apprentissage.

Le contexte joue un rôle déterminant

L'activité de l'apprenant est toujours insérée dans un environnement qui en rend possible l'appropriation. Le processus d'apprentissage s'inscrit dans une réalité culturelle et

contextualisée où la compréhension des objets et des événements est directement reliée à la forme dans laquelle elle se produit. En décontextualisant l'apprentissage (enseignement se déroulant sur campus, par exemple), la connaissance devient inerte ou difficile à utiliser car elle se construit en interaction avec un environnement différent de celui où elle a été créée (en particulier pour les savoirs pratiques) ou de celui où elle devra être utilisée (Jonassen, Campbell et Davidson, 1994). Les constructivistes considèrent que tout processus d'acquisition de connaissances et de compréhension est intimement lié au contexte social et émotionnel dans lequel l'apprentissage se déroule (Lebow, 1993). L'enseignement de ce point de vue n'est pas considéré comme la mise en séquence des événements d'enseignement mais comme une application des principes permettant de répondre aux besoins et aux contraintes de la situation dans laquelle se trouve l'apprenant. La compréhension se réalise grâce à l'utilisation continue et contextualisée des connaissances et ainsi, la situation d'apprentissage devrait promouvoir la manipulation des connaissances dans le contexte des pratiques ordinaires de la culture cible (Lebow, 1993). Certains des principes du constructivisme renvoient donc à la «cognition en situation» (situated cognition) (Brown, Collins et Duguid, 1989; Fredericksen et Donin, 1994). L'apprentissage contextualisé s'intéresse à l'interaction entre l'individu et son environnement (son contexte au sens large du terme) parce qu'elle détermine la perception et les comportements (Chiou, 1992).

4.2. Positionnement de notre approche

L'approche pédagogique que nous avons adoptée dans notre travail de thèse est l'approche constructiviste. En effet, nous avons voulu mettre l'apprenant au centre de l'apprentissage en lui permettant de construire son savoir au fur et à mesure de son évolution. L'apprenant est actif, autonome et le seul responsable de son évolution. Le système pédagogique qui réalise notre approche lui permet de gérer efficacement son parcours en le guidant vers la réalisation de ses objectifs d'apprentissage. Notre système pédagogique s'inscrit dans la catégorie des systèmes de formation à distance qui permet à l'apprenant une contextualisation de son apprentissage, c'est-à-dire, l'apprenant apprend à distance et dans son environnement immédiat (par exemple dans son lieu de travail). Pendant que l'apprenant entre en interaction avec le savoir scientifique proposé par le matériel qu'on lui fournit, il demeure en interaction avec les diverses composantes de son environnement qui lui offrent des informations pertinentes à la construction de son savoir. La désaffectation du savoir, possible dans un contexte de formation à distance grâce à l'utilisation des technologies pour transmettre des

informations, favorise aussi une construction plus personnalisée du savoir et moins influencée par l'enseignant.

L'utilisation des technologies pour déplacer le savoir des institutions scolaires vers le contexte immédiat des apprenants (accessibilité) facilite la contextualisation et le transfert des connaissances. Qu'il soit chez lui ou à son travail, l'apprenant peut, à tout moment, utiliser son environnement pour vérifier, confirmer, confronter ou ajuster les connaissances qu'il construit. Lorsque la formation se déroule à distance dans son milieu professionnel, l'apprenant peut être contraint d'appliquer directement ce qu'il est en train d'apprendre. La présence continue et immédiate du contexte peut stimuler les utilisations concrètes.

Dans notre approche, nous avons voulu expliciter la représentation des contenus et des démarches d'apprentissage pour permettre une grande flexibilité des modalités d'apprentissage. Il est donc possible de placer l'apprenant au centre du processus et de tenir compte de son lieu physique, de ses disponibilités temporelles, de son style d'apprentissage et de son environnement culturel (par exemple sa langue). Cela est possible en offrant des démarches pédagogiques qui peuvent être adaptées tant au plan des contenus, de la structure que des modalités d'interactions qui composent une activité d'apprentissage (Deschênes, 1991; Moore, 1977). L'utilisation de l'approche constructiviste nous permet de placer l'apprenant au centre du processus d'apprentissage et ainsi prendre en compte ses caractéristiques personnelles (Wagner et McCombs, 1995). En effet, l'apprentissage personnalisé constitue un champ d'application privilégié des concepts fondamentaux du constructivisme. Il fournit l'occasion de rompre avec certaines caractéristiques de l'enseignement traditionnel, en particulier, sa vision académique de la connaissance et la nécessité de retirer l'apprenant de son environnement immédiat pour l'insérer dans un établissement scolaire créant le plus souvent un contexte différent de celui dans lequel les connaissances seront utilisées. De plus, les individus perçoivent qu'apprendre à distance exige une plus grande responsabilité de leur part ce qui les rend plus disponibles à une telle approche (Wagner et McCombs, 1995).

En permettant à l'apprenant d'apprendre à distance et dans son environnement immédiat (par exemple dans son milieu professionnel), notre approche permet la construction d'une connaissance qui correspond bien à la vision des constructivistes. En effet, l'apprenant reste en contact continu avec son environnement ayant ainsi accès aux divers savoirs pratiques nécessaires à l'élaboration d'une connaissance plus complète du domaine qu'il étudie.

Les diverses personnes qui entourent l'apprenant constituent une composante importante du contexte. En enseignement présentiel, on réduit souvent l'aspect socioaffectif de

l'apprentissage aux seules interactions que l'apprenant peut avoir avec l'enseignant. Dans son milieu naturel, l'apprenant peut facilement initier et maintenir une plus grande diversité d'interactions en utilisant, selon ses besoins et ses préférences, des parents, des amis, des collègues de travail pour obtenir un support plus adapté. L'utilisation des technologies permet aussi d'avoir, en direct ou en différé, des communications avec les ressources et l'institution qui diffuse l'activité de formation à distance à laquelle il est inscrit ou avec les autres apprenants qui y participent.

Conclusion : Dans ce chapitre, nous avons présenté un panorama historique de l'apprentissage en ligne qui trouve son origine dans la formation à distance. Nous avons montré que l'apparition des systèmes pédagogiques en ligne, forme moderne de la formation à distance, est liée au succès que connaît les Technologies de l'Information et de la Communication (TIC), le multimédia et l'Internet. La complexité de la médiatisation de l'apprentissage nécessite de recourir à un processus d'ingénierie pédagogique qui se situe à l'intersection du design pédagogique, du génie logiciel et de l'ingénierie des connaissances. Nous avons essayé de faire apparaître tout au long de ce chapitre ce qui distingue les systèmes pédagogiques en ligne, de l'enseignement traditionnel en présentiel, à savoir, l'accessibilité, la contextualisation, la flexibilité, la diversification des interactions et la désaffectation des savoirs. Enfin, nous avons présenté les principes de l'approche constructiviste et son adéquation avec un processus d'apprentissage personnalisé. En effet, l'absence d'aide (fournie habituellement par l'enseignant) et la liberté offerte à l'apprenant dans la construction de son savoir peuvent constituer un frein à l'évolution de celui-ci. Le recours à la personnalisation de l'apprentissage où l'apprenant est mis au centre du système en favorisant une approche constructiviste dans l'acquisition du savoir constitue une réponse viable à ce problème.

Les éléments présentés dans ce chapitre nous permettent de mieux positionner notre travail dans le domaine du e-Learning. Il s'agit d'offrir un environnement pédagogique par découverte et qui est mis en œuvre par un système hypermédia adaptatif. La construction du savoir et l'acquisition des connaissances par l'apprenant se font à son rythme : l'apprenant est au centre du système. L'assise technologique de notre système est l'utilisation des technologies Web sémantiques et des standards de l'éducation.

2. Systèmes pédagogiques adaptatifs

Le processus d'apprentissage traditionnel peut être caractérisé par une autorité centralisée (le contenu est sélectionné par l'enseignant), une forte livraison (les enseignants transmettent la connaissance aux étudiants), un manque de personnalisation (le contenu doit satisfaire les besoins de plusieurs ou le principe du « one size fits all ») et un processus d'apprentissage statique/linéaire (un contenu inchangé). Cependant, à notre époque, l'environnement socio-économique des entreprises évolue dynamiquement et fait émerger d'autres types de défis sur les processus d'apprentissage: efficacité, réutilisabilité et pertinence de l'apprentissage (dépendant du problème). Pour répondre à ces attentes, un système pédagogique doit être orienté apprenant et muni d'un processus d'apprentissage non linéaire et adaptatif.

Il est vrai que de nombreuses solutions ont été mises en pratique dans le cadre des systèmes pédagogiques. Or, la majorité de ces solutions privilégie une approche centrée sur la mise à disposition de ressources pédagogiques de qualité, souvent au détriment de leur exploitation au sein de « scénarios pédagogiques » diversifiés. Face à ce constat, de nouvelles approches émergent. La qualité du service pédagogique rendu dépend, dès lors, de la capacité de ces nouvelles approches à fournir aux apprenants, d'une part, des contenus pédagogiques adaptés à leur profil et d'autre part, des processus qui les guident véritablement dans leur

apprentissage. Les systèmes pédagogiques adaptatifs ont pour objectif de répondre à ce besoin.

Dans ce domaine de l'ingénierie pédagogique, une première problématique de recherche concerne la représentation des contenus pédagogiques. Un résultat essentiel a consisté à séparer deux niveaux de description : le niveau logique et le niveau sémantique. Le niveau sémantique vise à exprimer les caractéristiques pédagogiques et cognitives des ressources. La logique de partage et de large diffusion des ressources pédagogiques a fait émerger un autre besoin, celui de personnaliser la forme, le contenu et l'assemblage des ressources pédagogiques. Ce nouveau besoin a conduit à introduire dans les systèmes pédagogiques une nouvelle dimension, celle de l'apprenant. Une troisième dimension doit être prise en compte dans l'ingénierie des systèmes pédagogiques, il s'agit des processus pédagogiques. Leur formalisation et leur gestion dans les systèmes adaptatifs deviennent essentielles si l'on veut proposer non seulement des contenus pédagogiques mais aussi des méthodes pédagogiques adaptées aux objectifs pédagogiques de chaque apprenant.

A la lumière de ce qui a été fait ces dernières années dans ce domaine, nous considérons qu'un système pédagogique adaptatif est basé sur quatre composants : un modèle de contenu qui contient, à différents niveaux d'abstraction, la description des ressources pédagogiques proposées aux apprenants ; un modèle de l'apprenant qui décrit les différents types de connaissances relatives aux apprenants ; et un modèle pédagogique qui définit les différentes méthodes et activités pédagogiques mises en œuvre pour atteindre des objectifs. Les méthodes et les techniques d'adaptation assurent la construction de parcours individualisés en utilisant les trois modèles.

Toutes les approches que nous avons étudiées n'utilisent pas obligatoirement ces quatre éléments. Par exemple, les approches issues du domaine des systèmes tuteurs intelligents décrivent les processus pédagogiques et peu le modèle de l'apprenant, alors que les approches issues du domaine des systèmes hypermédias adaptatifs ne prennent pas (ou peu) en compte le processus pédagogique et se concentrent sur le modèle de l'apprenant. Par ailleurs, les modèles de contenu peuvent être très différents dans leur contenu et dans leur forme d'une approche à une autre.

1. Caractérisation des systèmes pédagogiques adaptatifs

Aujourd'hui, les recherches et développements en ingénierie des systèmes pédagogiques adaptatifs s'appuient essentiellement sur :

i) des travaux autour des systèmes hypermédias adaptatifs, ii) des théories sur les méthodes d'apprentissage, iii) des recherches sur le Web sémantique adaptatif.

Dans les systèmes hypermédias adaptatifs (Brusilovsky, 2001), les connaissances sur l'apprenant sont essentielles, il s'agit de son expérience, de son niveau de connaissance du domaine (De Bra et al., 1998), de ses préférences en matière de méthodes pédagogiques, de formats de ressources, etc. (Höök, 1997). Les recherches dans ce domaine ont abouti d'une part, à des modèles apprenant intégrant différentes formes de connaissances, et d'autre part à des techniques d'adaptation du contenu et/ou de la structure de navigation de l'hypermédia. Dans ce courant de recherche, nous pouvons citer comme exemples le système KBS Hyperbook (Henze, 2000), le système AHA ! (De Bra et al, 2007), le système InterBook (Brusilovsky et al., 1998) et le système ELM-ART (Brusilovsky, 1996).

De nombreuses recherches théoriques dans les domaines de la Psychologie, de l'Intelligence Artificielle et de l'Education ont influencé la conception des systèmes pédagogiques (Schneider, 2003 ; Rhéaume, 1991). Ces recherches ont permis essentiellement de formaliser les méthodes d'apprentissage. De nombreux auteurs considèrent l'apprentissage comme un véritable processus dans lequel l'apprenant cherche à satisfaire des objectifs ou à résoudre des problèmes. L'activité pédagogique d'un domaine est alors considérée comme une activité guidée par des objectifs.

L'arrivée du Web et des technologies associées permet d'envisager l'utilisation de ressources pédagogiques qui sont distribuées et accessibles par Internet. Le Web Sémantique est une extension du Web qui vise à enrichir les ressources disponibles sur le Web avec des descriptions sémantiques de leur contenu dans le but de les rendre exploitables par des programmes (Berners-Lee et al., 2001). Dans le domaine des objets pédagogiques, plusieurs modèles de description sont aujourd'hui proposés: SCORM (ADL, 2001), LOM (LOM, 2002) ou IMS (IMS, 2002). Il y a aujourd'hui un changement de centre d'intérêt des objets pédagogiques vers des services pédagogiques. Cette évolution montre la nécessité d'identifier de nouvelles formes de métadonnées pour caractériser non seulement le contenu des ressources pédagogiques mais aussi les processus qui permettent de les utiliser dans différents contextes d'apprentissage.

Afin de présenter une vision complète et unifiée des recherches existantes dans le domaine de la conception des systèmes pédagogiques adaptatifs, nous proposons un cadre de référence pour caractériser ce domaine. Aucune des approches et des systèmes existants ne couvre la totalité de ce cadre de référence, en revanche chaque approche ou système peut être intégré dans ce cadre et peut être situé par rapport à d'autres approches. Ce cadre de référence est

défini par un modèle de contenu, un modèle de l'apprenant, un modèle pédagogique et un modèle d'adaptation.

1.1. Modèle de contenu

Un système pédagogique met à la disposition des apprenants des ressources pédagogiques qui traitent de notions appartenant à un domaine de connaissances. La description de ces ressources peut être faite sur deux niveaux d'abstraction distincts i) le niveau logique et ii) le niveau sémantique. Tous les systèmes ne font pas forcément la différence entre ces deux niveaux. Par ailleurs, il existe toujours un niveau physique qui est celui des contenus exprimés dans un certain langage sous forme de fichiers informatiques.

1.1.1. Structure logique

Au niveau logique, un contenu pédagogique est un ensemble de ressources pédagogiques reliées entre elles qui peuvent agréger un ou plusieurs contenus. Une ressource pédagogique est identifiée par un code unique (typiquement une URL) permettant sa localisation par une machine ou un homme. Plusieurs modèles de structuration logique ont été proposés dans le domaine des systèmes hypermédias. Le modèle DEXTER (Halasz, 1994) constitue un modèle de référence dans lequel la structure logique (appelée « Storage Layer ») est constituée de composants qui peuvent être des éléments atomiques, des liens ou des éléments composites. Au niveau logique, il est usuel d'utiliser une typologie des ressources pédagogiques basée sur la forme de leur contenu. (Crozat 2002) propose 9 formes de contenus: texte écrit, texte parlé, bruitage, musique, image, etc. Dans (Vassileva 1997 ; Dehors et al. 2006), les ressources pédagogiques (appelées teaching materials) sont caractérisées par le type de média qu'elles utilisent mais aussi par un type pédagogique, ainsi sont distinguées les unités d'information de type « introduction à une notion », « exercice sur une notion », « explication d'une notion » ou « test d'une notion ».

1.1.2. Structure sémantique

La structure sémantique est une description sémantique des ressources pédagogiques. Dans les systèmes actuels, la structure sémantique peut prendre différentes formes. Il peut s'agir de métadonnées associées aux ressources (ex., LOM, Dublin Core, etc.) ou bien d'un véritable modèle de domaine décrivant les ressources pédagogiques à un niveau conceptuel. L'approche métadonnée a essentiellement pour objectif de standardiser la description des ressources dans un souci de partage et de réutilisation. Plusieurs travaux de normalisation sur les métadonnées pour les ressources pédagogiques ont été menés (Duval, 2002). Dans l'autre

approche qui consiste à décrire la structure sémantique comme un modèle de domaine. Le domaine est celui du sujet d'enseignement. En général, ce modèle est un réseau de notions et de liens entre ces notions. Il peut exister différents types de liens entre les notions, par exemple les liens de type « pré-requis », « analogie », « généralisation », etc. Le modèle proposé dans (Vassileva, 1997) permet de représenter les notions et les liens. Quatre types de liens sont proposés : agrégation, généralisation, analogie et implication. La structure sémantique est un graphe ET/OU dans lequel il est possible d'exprimer que l'apprentissage d'une notion nécessite préalablement celui de plusieurs autres ou bien qu'il nécessite alternativement une notion ou une autre.

L'approche modèle de domaine apporte une aide dans la démarche pédagogique puisqu'elle définit un ordre dans l'apprentissage des notions, conduisant à des parcours bien précis à travers les ressources pédagogiques. En revanche, dans l'approche métadonnées, la structure sémantique est relativement indépendante de toute activité pédagogique.

1.2. Modèle de l'apprenant

Dans les systèmes pédagogiques adaptatifs, le modèle de l'apprenant est primordial pour la génération de parcours personnalisés. Il s'agit de considérer les propriétés de l'apprenant auxquelles le système doit s'adapter. Ces propriétés peuvent être différentes d'un apprenant à un autre, et pour un même apprenant, elles peuvent varier au cours du temps. Au niveau des modèles de l'apprenant, il est important de distinguer un modèle adaptable d'un modèle adaptatif. Dans le premier cas, le modèle est modifié uniquement par l'apprenant, dans le second cas, le modèle est modifié automatiquement par le système au fur et à mesure des interactions avec l'apprenant. Enfin, les modèles de l'apprenant peuvent être construits selon deux approches ; une approche « stéréotype », dans laquelle on considère des groupes d'apprenants ayant des caractéristiques communes, et une approche « individuelle » dans laquelle chaque apprenant est décrit avec ses propres caractéristiques. Le critère essentiel qui permet de classer les modèles de l'apprenant concerne la nature des connaissances qu'ils permettent d'exprimer. Le deuxième critère qui est utilisé porte sur les techniques d'initialisation et de gestion de ces modèles.

1.2.1. Nature des connaissances

Ce critère permet de caractériser les types de connaissances qui sont représentées dans le modèle de l'apprenant. Il est usuel de considérer qu'un modèle de l'apprenant est composé de quatre sous-modèles, chaque sous-modèle correspondant à un type particulier de

connaissance. Le sous-modèle de la connaissance du domaine : il définit les connaissances de l'apprenant relativement à la matière enseignée. Dans (Benyon 1993), trois niveaux de connaissance sont distingués, le niveau tâche, le niveau logique et le niveau physique. Le niveau tâche exprime les objectifs réalisés par l'apprenant sur la matière enseignée. Le niveau logique définit les connaissances de l'apprenant sur les notions de la matière enseignée et le niveau physique correspond à la connaissance effective de l'apprenant. Le sous-modèle de l'expérience regroupe des connaissances de l'apprenant qui sont indépendantes du domaine. Ces connaissances sont relatives aux sujets d'intérêt d'un apprenant, à sa culture, à son expérience... Le sous-modèle des préférences exprime les préférences de l'apprenant relatives à la forme de la présentation du cours mais aussi aux styles et aux méthodes d'enseignement. Par ailleurs, ce sous-modèle intègre en général aussi « les traits de personnalité » de l'apprenant : sa motivation, son degré de concentration... L'ensemble de ces connaissances ne peut pas être généré par le système mais doit être fourni par l'apprenant. Ainsi, ces connaissances jouent un rôle plus important dans les systèmes adaptables que dans les systèmes adaptatifs. Le sous-modèle des tâches capture la connaissance relative aux activités effectivement mises en œuvre par l'apprenant, au cours de sessions d'enseignement. Cette connaissance contextuelle est essentielle, elle explique pourquoi l'apprenant a réalisé certaines actions.

1.2.2. Gestion des connaissances

Cet aspect permet de caractériser les techniques d'acquisition et de mise à jour du modèle de l'apprenant. Il existe trois approches essentielles. Lorsque l'apprenant ou l'administrateur renseigne les différentes rubriques du modèle apprenant en utilisant par exemple un formulaire ou un questionnaire, il s'agit d'une approche statique. Ce type d'acquisition est utilisé pour les données qui n'évoluent pas en fonction des actions de l'apprenant. L'acquisition dynamique consiste à observer l'apprenant et à tracer ses activités et ses résultats. Dans ce cas, l'apprenant n'est pas sollicité et le modèle est mis à jour automatiquement en fonction de l'observation de l'apprenant. Le système METADYNE (Delestre et al., 2000) par exemple, utilise des mécanismes d'inférence pour générer des connaissances sur l'apprenant. D'une manière générale, les systèmes tirent partie des deux approches. Une approche mixte permet de bâtir un premier modèle avec par exemple, des questionnaires à remplir lors de la première connexion, le modèle est ensuite mis à jour par le système en inférant de nouvelles valeurs en fonction du comportement de l'apprenant.

1.3. Modèle pédagogique

Dans les systèmes pédagogiques adaptatifs, les recherches se sont surtout centrées sur les contenus pédagogiques et peu sur les démarches pédagogiques. Pourtant, il existe plusieurs manières d'apprendre, et un système pédagogique adaptatif doit permettre d'adapter la démarche aux besoins de l'apprenant. Dans ces conditions la modélisation et la gestion des démarches pédagogiques est essentielle. La modélisation des démarches pédagogiques peut se faire à différents niveaux (stratégique, opérationnel...) et la manière de modéliser une démarche pédagogique peut être très différente d'une approche à une autre.

1.3.1. Niveaux d'abstraction dans les démarches pédagogiques

Nous considérons qu'une démarche pédagogique peut être modélisée à différents niveaux d'abstraction. En référence à (Allert, Dhraief & Nejdil, 2002), quatre niveaux d'abstraction peuvent être identifiés (voir Figure 1). Le quatrième niveau définit les processus d'apprentissage du point de vue des stratégies d'enseignement. A ce niveau, il est fréquent de distinguer les courants d'apprentissage tels que le courant « behavioriste », le courant « cognitiviste » et le courant « constructiviste ». Le troisième niveau définit des approches pédagogiques. Dans la pratique, les approches les plus utilisées sont l'apprentissage basé sur la résolution de problèmes, le raisonnement à base de cas, l'approche collaborative, l'approche par l'exposé, par l'analogie... Le deuxième niveau définit les modèles opérationnels d'apprentissage. Il s'agit d'une structuration des processus pédagogiques. Cette structuration conduit à une description du processus en plusieurs phases, mettant en évidence des cycles et des ordonnancements d'activités pédagogiques. Le premier niveau définit les activités pédagogiques.

Ce cadre de modélisation permet de montrer d'une part, qu'il existe différents niveaux d'abstraction dans la modélisation des processus pédagogiques et d'autre part, que les différents niveaux sont interdépendants. En effet, un choix, par exemple effectué en termes de stratégie pédagogique, contraint les niveaux inférieurs.

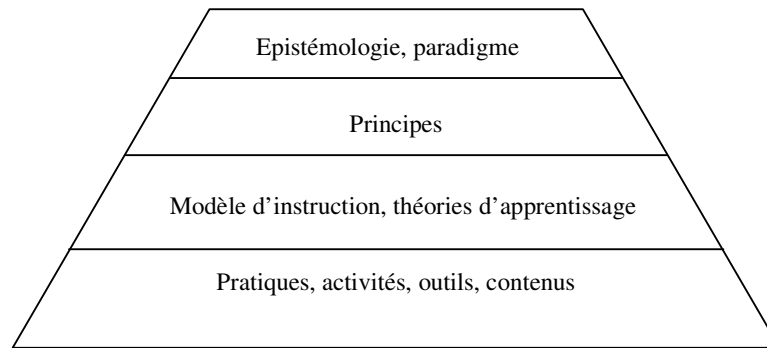


Figure 1 : Niveaux d'abstraction dans les démarches pédagogiques (Allert et al., 2002)

1.3.2. Modèles de démarches pédagogiques

On distingue trois approches de modélisation d'une démarche pédagogique.

L'approche orientée activités consiste à décrire la démarche sous la forme d'un processus composé d'activités ou tâches qui peuvent être simples ou complexes. Dans les systèmes actuels, cette approche de modélisation est la plus répandue. Le processus est alors souvent représenté par un graphe ET/OU dans lequel les tâches pédagogiques sont identifiées et décomposées. Par exemple, la figure 2 décrit une partie d'un modèle de tâches dans lequel est détaillée la tâche « donner un exercice ».

L'approche orientée ressources considère la démarche pédagogique comme un processus de sélection, d'assemblage et de présentation de contenus. Cette approche met l'accent davantage sur la modélisation des contenus que sur celle des processus. Dans les systèmes qui adoptent cette approche, il n'y a pas en général de modélisation explicite des démarches pédagogiques.

L'approche orientée objectifs considère la démarche pédagogique comme un processus de satisfaction d'objectifs pédagogiques. Chaque objectif est associé à un ou plusieurs scénarios permettant de le satisfaire. L'intérêt de cette approche est double, d'abord elle permet la personnalisation puisqu'elle autorise différentes manières de satisfaire un même objectif, ensuite la modélisation orientée objectif favorise une pédagogie centrée sur les besoins des apprenants.

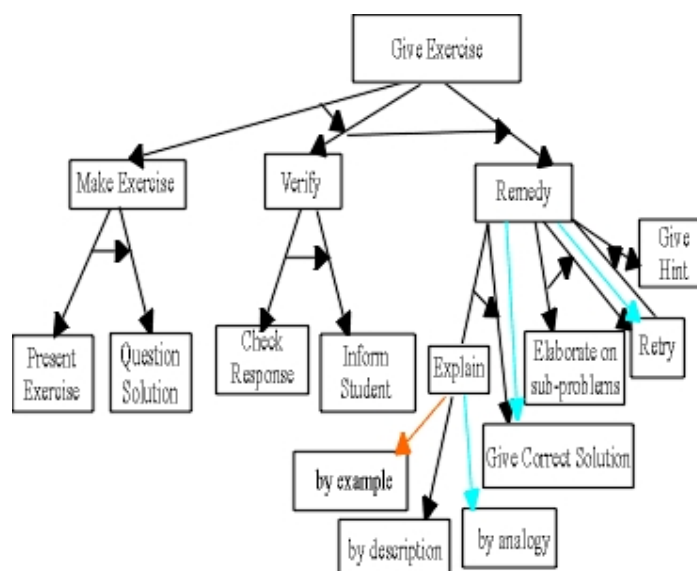


Figure 2 : Modèle d'activités pédagogiques « donner exercice » (Vassileva, 1997)

1.4. Modèle d'adaptation

Un système pédagogique adaptatif met en œuvre des mécanismes qui exploitent la connaissance du domaine, la connaissance sur les apprenants et la connaissance sur les processus pédagogiques pour offrir des parcours et des contenus pédagogiques personnalisés. Dans les systèmes pédagogiques adaptatifs, l'adaptation proposée peut être plus ou moins « puissante ». Nous proposons de caractériser l'adaptation par un mode d'adaptation et une portée (l'objet de l'adaptation).

Nous distinguons trois modes d'adaptation :

- Parcours type : ce sont des parcours prédéfinis à l'avance par le concepteur. Ces parcours types peuvent être définis selon différents critères, en fonction des connaissances du domaine des apprenants ou en fonction des stratégies pédagogiques souhaitées par les apprenants. Une des difficultés de cette approche est d'identifier les classes d'apprenants, ces classes étant nécessaires pour élaborer les parcours types.
- Parcours adaptatif : ce sont des parcours individuels qui s'adaptent à chaque apprenant en fonction de ce qu'il est et non pas de ce qu'il fait. En général, ces parcours sont construits en début de session d'apprentissage à partir des connaissances contenues dans le modèle de l'apprenant.
- Parcours dynamique : ce sont des parcours construits dynamiquement, qui s'ajustent à l'apprenant en fonction des actions qu'il accomplit tout au long de son processus

d'apprentissage. De tels parcours sont générés et mis à jour au cours de la session d'apprentissage.

Du point de vue "portée" de l'adaptation, nous distinguons trois formes d'adaptation : l'adaptation de la navigation, l'adaptation du contenu et l'adaptation de la présentation.

L'adaptation de la navigation permet l'adaptation dans l'ordonnement des notions pédagogiques qui vont être présentées à l'apprenant.

L'adaptation du contenu est basée sur la sélection de contenus en fonction des connaissances de l'apprenant sur le domaine.

L'adaptation de la présentation traite de la disposition et de l'aspect visuel. Cette forme d'adaptation exploite certaines préférences de l'apprenant.

Nous notons que l'adaptation du contenu et l'adaptation de la présentation sont souvent confondues.

Les méthodes et les techniques d'adaptation varient considérablement d'un système à un autre (voir figure 3).

1.4.1. Méthodes d'adaptation

Méthodes d'adaptation du contenu :

On distingue quatre méthodes d'adaptation du contenu :

- Explications par les pré-requis : avant de présenter une explication sur une notion, le système insère des explications sur tous ses notions pré-requis qui ne sont pas connus par l'apprenant.
- Explications comparatives : si une notion similaire à la notion présentée est connue, l'apprenant obtient des explications comparatives qui insistent sur les différences et similitudes entre les deux notions. Cette méthode est très utilisée pour enseigner les langages de programmation.
- Explications variées : le système sauvegarde plusieurs variantes d'une même partie de la page et fournit à l'apprenant la partie qui correspond à son modèle.
- Tri : les différents fragments d'information au sujet d'une notion peuvent être triés par ordre décroissant de pertinence en fonction du background et des connaissances de l'apprenant.

Méthodes d'adaptation de la navigation :

On distingue cinq méthodes d'adaptation de la navigation:

- Guidage global : l'objectif global de l'apprenant étant connu par le système, il s'agit de trouver le chemin le plus court pour atteindre l'information pertinente. Dans chaque page, le

système fournit à l'apprenant le lien à suivre, c'est-à-dire, un guidage direct à l'aide par exemple du bouton « suivant ».

- Guidage local : il s'agit de suggérer à l'apprenant les liens les plus pertinents de la page en cours en fonction de son niveau de connaissances, de ses préférences et de son expérience.

L'objectif global n'est pas connu dans le guidage local.

- Orientation locale : il s'agit d'aider l'apprenant à comprendre sa position relative dans l'hyperespace (1) en fournissant des informations sur les pages en sortie de la page en cours et (2) en diminuant le nombre d'hyperliens sortant d'une page pour limiter la surcharge cognitive de l'apprenant.

- Orientation globale : il s'agit d'aider l'apprenant à comprendre la structure globale de l'hyperespace et sa position absolue dans cet hyperespace.

- Gestion de vues personnalisées : il s'agit d'offrir à l'apprenant la possibilité d'organiser une partie de l'hyperespace en fonction d'objectifs de travail ponctuels. Ceci évite à l'apprenant d'être confrontés à un hyperespace trop complexe avec des risques de désorientation et de surcharge cognitive. Cette partie de l'hyperespace est un ensemble d'hyperliens qui pointent vers des hyperdocuments pertinents et représente une vue personnalisée pour l'apprenant. Les navigateurs Web proposent par exemple des bookmarks et des hotlists pour construire des vues personnalisées.

1.4.2. Techniques d'adaptation

L'adaptation nécessite des mécanismes internes et externes très spécifiques. Au niveau externe, le domaine des systèmes hypermédias adaptatifs propose de nombreuses techniques d'adaptation permettant de présenter un hyperespace de manière personnalisée. Ces techniques d'adaptation largement décrites dans (Brusilovsky, 1996 ; Brusilovsky, 2001) distinguent l'adaptation du contenu de l'adaptation de la navigation :

Technique d'adaptation du contenu :

Parmi les techniques d'adaptation du contenu, il y a :

- Texte conditionnel : toutes les informations sur une notion sont divisées en plusieurs parties. Chaque partie est associée à une condition sur le niveau de connaissance de l'apprenant. Quand le système présente des informations à l'apprenant, il ne présente que celles dont la condition est vraie.

- « StretchText » est une technique de « plus haut niveau » et correspond à un type particulier d'hypertextes. Tandis que dans un hypertexte classique, l'activation d'un hyperlien ou « hot word » signifie l'ouverture d'une nouvelle page, dans un StretchText le « hot word » est

simplement remplacé par le texte correspondant et étend ainsi la page courante. L'opération inverse est possible. L'idée de l'adaptation des «Stretch Text» est de «déplier» les «hot word» pertinents et de laisser les autres «pliés».

- Les pages variées: permet d'implanter des explications variées. On conçoit plusieurs présentations pour une même page en fonction du profil de l'apprenant
- Les fragments variés permettent d'implanter les explications variées où une page contient (1) des explications sur plusieurs notions, c'est-à-dire, un fragment par notion ou plusieurs variantes d'une même notion ; (2) différentes explications structurelles de la même notion en fonction du niveau de connaissances de l'apprenant.
- les frames : les informations sur une notion particulière sont représentées par un frame. Les slots du frame peuvent contenir des explications différentes de la notion, des liens avec d'autres frames, des exemples. Des règles de présentations particulières sont utilisées pour décider quels slots doivent être présentés à chaque apprenant.

Techniques d'adaptation de la navigation :

Parmi les techniques d'adaptation de la navigation, il y a :

- Guidage direct : en fonction des objectifs d'un apprenant, le système définit l'hyperlien le plus approprié et le visualise explicitement en changeant par exemple sa couleur.
- Tri des liens : cette technique consiste à ordonner les hyperliens en fonction de leur pertinence pour un apprenant donné. Les liens susceptibles d'intéresser l'apprenant sont situés en haut de la page pour favoriser leur découverte.
- Annotation des liens : le principe de cette technique est de rajouter aux différents hyperliens des commentaires pour indiquer à l'apprenant le contenu de la page qui est accessible par le lien.
- Masquage des liens: cette technique permet de masquer les hyperliens les moins pertinents pour l'apprenant.
- Adaptation de carte : c'est une carte de navigation indiquant la structure hiérarchique ou en réseau de l'hyperespace pour permettre à l'apprenant de choisir les pages à consulter.

Au niveau interne, les techniques d'adaptation mises en œuvre par les systèmes sont des règles qui exploitent les trois types de connaissances : les connaissances de domaine, les connaissances de l'apprenant et les connaissances pédagogiques. Dans (Vassileva, 1997), on distingue quatre types de règles d'apprentissage. Les « règles d'enseignement » déterminent un parcours d'apprentissage dans l'ensemble des notions du domaine. Les «règles de sélection de stratégies » définissent une stratégie pédagogique avant de commencer l'exécution du plan. Par exemple la stratégie « structurée » est proposée par le système comme une approche

d'enseignement dans laquelle l'apprenant est complètement guidé ; c'est le système qui détermine les notions à acquérir et les tâches associées. Les « règles de sélection de méthodes » déterminent pour une notion, l'activité pédagogique et une décomposition pour l'apprendre. Ces règles utilisent les tâches déjà effectuées ainsi que les ressources pédagogiques déjà utilisées. Les « règles de sélection des ressources pédagogiques » déterminent la ressource pédagogique à présenter en fonction des préférences de l'apprenant en matière de média. Dans les approches Web sémantique adaptatif, les règles exploitent les métadonnées définies sur les ressources pédagogiques. Par exemple, dans le système Personal Reader (Dolog, Henze, Nedjl, & Sintek, 2004) les règles, exprimées dans le langage TRIPLE (Sintek & Decker, 2002) permettent de raisonner sur les métadonnées annotant les ressources pédagogiques. Les règles sont utilisées pour générer des liens personnalisés.

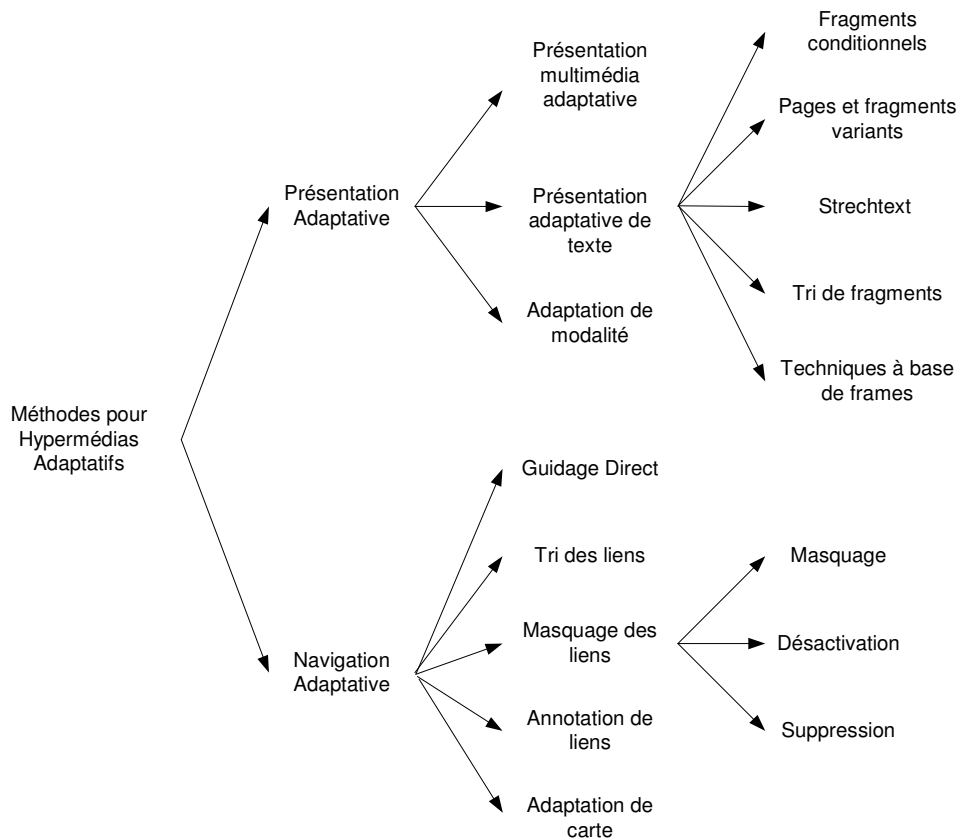


Figure 3 : Méthodes et techniques d'adaptation des systèmes pédagogiques adaptatifs

2. Etude de quelques systèmes pédagogiques adaptatifs

Dans le cadre de cette étude, notre choix s'est porté sur des systèmes qui gèrent (plus ou moins) les trois types de connaissances décrites ci-dessus et qui mettent en œuvre des règles d'adaptation. Nous avons retenu cinq systèmes que nous décrivons sommairement et ensuite nous essayons de dégager les caractéristiques de chaque système selon le cadre de référence présenté précédemment.

2.1. Description des systèmes

AHA ! (de Bra et al., 2007) est un système auteur pour construire des hypertextes adaptatifs. Il génère dynamiquement des pages adaptées à l'apprenant à partir de pages HTML extraites du web.

ELM-ART (Brusilovsky et al., 2001) est un hypertexte adaptatif. C'est un tuteur intelligent adapté au Web et dédié à l'apprentissage du langage Lisp. Le cours hypertexte est organisé hiérarchiquement en chapitres, sections, sous-sections, etc. et l'apprentissage se fait via des explications, des exemples, et la résolution de problèmes.

KBS Hyperbook (Henze et al., 1999) est un système hypermédia adaptatif qui permet aux apprenants de déterminer leurs objectifs et leur fournit une aide pour appréhender les différentes connaissances nécessaires pour atteindre leurs objectifs.

METADYNE (Delestre et al., 2000) est un hypermédia adaptatif et dynamique où l'hypermédia est virtuel et les cours sont générés dynamiquement en filtrant des briques élémentaires de cours en se basant sur le modèle de l'apprenant.

Personal Reader (Dolog et al., 2004) est basé sur le Web sémantique. Il exploite des métadonnées sur les ressources pédagogiques et les profils des apprenants pour générer des liens hypertextes personnalisés.

2.2. Analyse des systèmes

Les modèles de contenu des systèmes analysés reposent sur une description des contenus à deux niveaux : le niveau logique et le niveau sémantique. Dans la majorité des systèmes étudiés, la structure logique est un hyperespace composé de nœuds (pages) et de liens hypertextes. La structure sémantique reste « pauvre » dans la plupart des systèmes. Elle est limitée à une description des notions du domaine avec des liens de prérequis entre ces notions. Peu de systèmes associent dans la structure sémantique les activités permettant l'apprentissage de ces notions. La structure sémantique prend le plus souvent la forme d'une structure en réseau ou d'une structure en arbre. Dans le cas d'AHA ! (De Bra et al, 2007), le

niveau sémantique est représenté par une structure arborescente de notions. Cette hiérarchie permet une structuration du cours en chapitres, sections, sous sections, etc.

L'acquisition et la gestion du modèle de l'apprenant reposent dans la plupart des systèmes étudiés sur une approche individuelle de modélisation avec une mise à jour continue du modèle de l'apprenant au fur et à mesure de l'apprentissage. Du point de vue des connaissances prises en compte dans le modèle de l'apprenant, les systèmes sont de plus en plus « riches ». Cette évolution est probablement liée à l'utilisation de plus en plus intensive des ressources pédagogiques du Web. Dans ce nouveau contexte, le modèle de l'apprenant est essentiel pour mettre en œuvre des services d'individualisation. Malgré cette évolution, les préférences des apprenants en termes de stratégies et de méthodes d'enseignement sont peu prises en compte dans les modèles de l'apprenant ; ceci constitue une lacune importante des systèmes actuels. Le modèle pédagogique concerne essentiellement les démarches pédagogiques mises en œuvre. Peu de systèmes prennent en compte cette dimension ; la démarche pédagogique se limite à des parcours d'unités pédagogiques. Par exemple, le système METADYNE (Delestre et al., 2000) intègre un générateur de cours qui a pour fonction d'assembler et de présenter les cours à l'apprenant. Il applique, dès lors, une approche orientée ressources. Toutefois, certains des systèmes présentés font appel à une approche orientée activités qui explicite et décompose le processus pédagogique en activités simples ou complexes. Le système Personal Reader (Dolog et al., 2004) utilise cette approche. Les méthodes et techniques d'adaptation utilisées sont centrées essentiellement sur une adaptation de la présentation des contenus et de la navigation à travers ces contenus. Les règles d'adaptation sous-jacentes utilisent uniquement les connaissances sur la situation de l'apprenant. Il s'agit de règles de filtrage qui permettent de sélectionner les pages à afficher en fonction des pages déjà utilisées. La puissance de ces méthodes est directement dépendante de la richesse des trois modèles précédents. En particulier, l'existence de règles permettant la personnalisation des méthodes d'apprentissage dépend de la formalisation et de la richesse du modèle pédagogique.

Table 1 : Etude comparative de systèmes pédagogiques adaptatifs

| | | AHA ! | ELM-ART | KBS Hyperbook | METADYNE | Personal Reader |
|-----------------------|-------------------------|---|---|---|---|---|
| Modèle de contenu | Structure logique | Hypertexte hiérarchique | Hypertexte organisé hiérarchiquement en chapitres, sections, sous-sections... | Hypermédia | Hypermédia | Hypermédia |
| | Structure sémantique | Modèle de domaine | Modèle de domaine | Métadonnées basés sur une ontologie | Modèle de domaine | Métadonnées |
| Modèle de l'apprenant | Nature | Modèle de recouvrement | Modèle de recouvrement | Réseau Bayésien | Modèle épistémique (connaissances) +Modèle comportemental (préférences, objectifs, capacités) | Intérêts, préférences, objectifs et prérequis |
| | Gestion | Mixte | Mixte | Mixte | Mixte | Statique |
| Modèle pédagogique | Niveau d'abstraction | | | Constructiviste | | |
| | Modèle de démarches | Orienté activités | | orienté objectifs | Orienté ressources | Orienté activités |
| Modèle d'adaptation | Type d'adaptation | Adaptation du contenu et de la navigation (guidage local) | Adaptation de la navigation (guidage local) | Adaptation de la navigation (guidage local) | Adaptation du contenu et de la navigation (guidage local) | Adaptation de la navigation (guidage local) |
| | Techniques d'adaptation | Texte conditionnel et annotation des liens | Annotation des liens et tri des liens | Annotation des liens | Règle de sélection de briques de cours | Règles de génération de liens |

2.3. Problèmes des systèmes pédagogiques adaptatifs actuels

Un travail d'exploration et d'analyse des systèmes pédagogiques adaptatifs existants nous a permis de constater certaines limites :

- Le manque de visibilité des parcours d'apprentissage par les apprenants (guidage local, mauvaise représentation de l'hyperespace, etc.) ;
- La dimension pédagogique n'est pas (ou très peu) prise en compte ;
- Les connaissances du modèle de l'apprenant sont considérées comme fiables et ne sont pas remises en cause ce qui implique des erreurs importantes d'adaptation (effet boule de neige) ;

- Seuls les liens de prérequis et/ou la hiérarchie des notions sont pris en compte dans l'adaptation.

Conclusion : Les systèmes pédagogiques adaptatifs constituent une réponse fiable aux besoins des entreprises en particulier et des sociétés humaines en général en termes de formation et de formation continue. Ces systèmes permettent de fournir un apprentissage efficace, adapté et pertinent, des mots clés, censés aujourd'hui intéresser tout demandeur de formation.

Dans ce chapitre, nous avons essayé de caractériser les systèmes pédagogiques adaptatifs en définissant un cadre de référence permettant de les comparer les uns avec les autres. Nous avons identifié quatre composants impliqués dans la réalisation des systèmes pédagogiques adaptatifs, à savoir, le modèle de contenu, le modèle de l'apprenant, le modèle pédagogique et le modèle d'adaptation. Nous avons mené une étude comparative de cinq systèmes dans le domaine mais aucun de ces systèmes ne couvre la totalité du cadre de référence.

Dans notre travail de thèse, nous proposons un système pédagogique adaptatif qui essaye de répondre aux lacunes des systèmes existants en combinant les systèmes hypermédias adaptatifs et les modèles du Web sémantique (présenté dans le chapitre 3) tout en introduisant la dimension pédagogique de l'apprentissage.

3. Web sémantique et e-Learning

Ce chapitre présente la notion d'ontologie dans le cadre du Web sémantique (WS) et l'intérêt de l'utilisation des ontologies et du standard des métadonnées LOM dans le domaine du e-Learning pour le développement de systèmes pédagogiques adaptatifs.

Le Web sémantique est une vision permettant de faciliter et d'améliorer la recherche d'information, faite par l'homme et la machine, grâce à la représentation sémantique du contenu des ressources manipulées. Il étend le Web actuel (Berners-Lee et al., 2001) dans le sens où le Web est strictement syntaxique, la structure des ressources accessibles est bien définie mais le contenu du Web reste quasiment inaccessible aux traitements machines.

Les ontologies jouent un rôle très important pour la réalisation du Web sémantique. Elles sont utilisées pour fournir le vocabulaire et la structure des métadonnées associées aux ressources du Web ou aux services Web et servent de représentation pivot pour l'intégration de sources de données hétérogènes.

Dans ce contexte, nous pensons que les ontologies peuvent apporter énormément aux systèmes pédagogiques adaptatifs et particulièrement à ceux qui s'appuient sur des technologies du Web (par exemple les hypermédias) et qui réutilisent des ressources pédagogiques du Web.

1. Web sémantique (WS)

L'expression Web sémantique, attribuée à Tim Berners-Lee (Berners-Lee et al., 2001) au sein du W3C², fait d'abord référence à la vision du Web de demain comme un vaste espace d'échange de ressources entre êtres humains et machines permettant une meilleure exploitation de grands volumes d'informations et de services variés. Le Web sémantique, à la différence du Web actuel, devrait décharger les utilisateurs d'une bonne partie de leurs tâches de recherche, de construction et de combinaison des résultats, grâce aux capacités accrues des machines à accéder aux contenus des ressources et à effectuer des raisonnements sur ceux-ci.

Le Web sémantique, concrètement, est d'abord une infrastructure de langages permettant l'utilisation de connaissances formalisées qui s'ajoutent au contenu informel actuel du Web, même si aucun consensus n'existe sur le degré de cette formalisation. Cette infrastructure doit :

- permettre de localiser, d'identifier et de transformer des ressources (des documents ou des services) de manière robuste tout en renforçant l'esprit d'ouverture du Web avec sa diversité d'utilisateurs.
- s'appuyer sur des consensus, par exemple, sur les langages de représentation ou sur les ontologies utilisées.
- contribuer à assurer, le plus automatiquement possible, l'interopérabilité et les transformations entre les différents formalismes et les différentes ontologies.
- faciliter la mise en œuvre de calculs et de raisonnements complexes tout en offrant des garanties sur leur validité.
- offrir des mécanismes de protection (droits d'accès, d'utilisation et de reproduction).

Les recherches actuelles s'appuient énormément sur des résultats de recherches en ingénierie des connaissances (Gandon, 2002). Or, l'utilisation et l'acceptation de ces résultats à l'échelle du (ou d'une partie du) Web posent de nouveaux problèmes et défis : changement d'échelle dû au contexte de déploiement, le Web et ses dérivés (intranet, extranet), nécessité d'un niveau élevé d'interopérabilité, ouverture, standardisation, diversités des usages, distribution bien sûr et aussi impossibilité d'assurer une cohérence globale. Comme l'écrit, en substance, Tim Berners-Lee, le Web sémantique est ce que nous obtiendrons si nous réalisons le même processus de globalisation sur la représentation des connaissances que celui que le Web fit initialement sur l'hypertexte.

² World Wide Web Consortium

2. Ontologies

2.1. Définition des ontologies

En Intelligence Artificielle, la définition communément admise d'une ontologie est énoncée par T. Gruber (Gruber, 1993) comme la « spécification explicite d'une conceptualisation ».

Cette définition fait suite à un premier essai en 1991 (Gruber, 1991) : « An ontology defines the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary ».

Cette définition a ensuite été précisée (Studer et al., 1998) pour devenir : « La spécification formelle et explicite d'une conceptualisation partagée ».

Dans cette définition, par "spécification explicite", les auteurs indiquent qu'une ontologie est un ensemble de concepts, de propriétés, d'axiomes, de fonctions et de contraintes explicitement définis.

L'adjectif "formelle" précise que cette conceptualisation doit pouvoir être comprise et interprétée par des agents logiciels. En effet, la formalisation est nécessaire pour que ces agents puissent être munis de capacités de raisonnement permettant de décharger les différents utilisateurs d'une partie de leur tâche d'exploitation et de combinaison des ressources du Web.

Le terme "partagé" précise l'aspect consensuel du vocabulaire employé. Ce terme rappelle que l'on doit assurer une réutilisation de la formalisation choisie pour permettre l'exploitation des ressources du Web par différentes applications ou agents logiciels.

Enfin, le terme "conceptualisation" implique également l'aspect intentionnel, lié à un objectif de réalisation.

En bref, une ontologie fournit un vocabulaire de termes et de relations pour modéliser les connaissances d'un domaine d'application.

Pour représenter les ontologies, le W3C a standardisé le langage OWL³. Le langage OWL s'appuie sur le langage DAML+OIL, produit de la combinaison de DAML⁴ et OIL⁵. Il est actuellement construit sur RDFS⁶, et apporte ainsi aux langages du Web sémantique, l'équivalent d'une logique de description tout en disposant d'une syntaxe XML.

³ Ontology Web Language

⁴ Darpa Agent Markup Language

⁵ Ontology Inference Layer

⁶ Resource Description Framework Schema

2.2. Intérêt des ontologies

Plusieurs recherches (Grüninger et Lee, 2002 ; Russ et al., 1999 ; Uschold et Grüninger, 1996 ; Mizoguchi et Bourdeau, 2000; Staab & Maedche, 2001 ; Psyché et al., 2003), ont dégagé l'intérêt d'utiliser des ontologies dans les systèmes à base de connaissances (SBC) et le Web sémantique :

La représentation des connaissances du domaine d'un SBC : les ontologies servent à représenter explicitement les connaissances du domaine d'un SBC. En particulier, elles servent de schéma à la représentation des connaissances du domaine dans la mesure où elles décrivent les objets, leurs propriétés et la façon dont ils peuvent être combinés pour constituer des connaissances explicites et complètes du domaine.

La communication : les ontologies peuvent intervenir dans la communication entre personnes, organisations et logiciels [Uschold et Grüninger, 1996]. En effet, les ontologies servent, par exemple, à créer au sein d'un groupe ou d'une organisation un "vocabulaire conceptuel commun". Dans ce cas, on est plutôt dans le cadre d'une ontologie informelle. Dans le cas de la communication entre personnes et systèmes, l'ontologie est formelle et effectue en général une tâche précise dans le SBC ou le système d'information. L'ontologie est un puissant moyen pour lever les ambiguïtés dans les échanges.

L'interopérabilité : selon (Uschold et Grüninger, 1996) « le développement et l'implantation d'une représentation explicite d'une compréhension partagée dans un domaine donné, peut améliorer la communication, qui à son tour permet une plus grande réutilisation, un partage plus large et une interopérabilité plus étendue ». L'interopérabilité est donc une spécialisation de la communication qui permet de répertorier les concepts que des applications peuvent s'échanger même si elles sont distantes et développées sur des plateformes différentes.

L'aide à la spécification de systèmes : (Mizoguchi & Ikeda., 1997) expliquent que « la plupart des logiciels conventionnels sont construits avec une conceptualisation implicite et que la nouvelle génération des systèmes utilisant les travaux en intelligence artificielle devrait être basée sur une conceptualisation explicitement représentée ». Dans ce contexte, l'ontologie fournit une description explicite des objets que doit manipuler un système.

L'indexation et la recherche d'information : dans le Web sémantique, les ontologies y sont utilisées pour déterminer les annotations sémantiques décrivant les ressources sur le Web.

3. Langages du Web sémantique

3.1. Les langages du W3C

La proposition du W3C s'appuie au départ sur une pyramide de langages dont seulement les couches basses sont aujourd'hui relativement stabilisées. La figure 4 montre l'organisation en couches des langages du WS proposée par le W3C. Deux types de bénéfices peuvent être attendus de cette organisation.

(1) Elle permet une approche graduelle dans les processus de standardisation et d'acceptation par les utilisateurs.

(2) Par ailleurs, si elle est bien conçue, elle doit permettre de disposer du langage au bon niveau de complexité, celle-ci étant fonction de l'application à réaliser.

Un aspect central de l'infrastructure du WS est sa capacité d'identification et de localisation des diverses ressources. Elle repose sur la notion d'URI (Uniform Resource Identifier) qui permet d'attribuer un identifiant unique à un ensemble de ressources, sur le Web bien sûr mais aussi dans d'autres domaines (documents, téléphones portables, personnes, etc.). Cette notion est à la base même des langages du W3C.

Une autre caractéristique de tous ces langages est d'être systématiquement exprimables et échangeables dans une syntaxe XML. Ceci permet de bénéficier de l'ensemble des technologies développées autour de XML : XML Schemas, outils d'exploitation des ressources XML (bibliothèques JAVA, etc.), bases de données gérant des fichiers XML, même si un langage de requêtes spécifique (par exemple SPARQL⁷, RDF⁸ Query) est nécessaire pour les langages construits au dessus de XML, tel que RDF.

⁷ Simple Protocol and RDF Query Language

⁸ Resource Description Framework

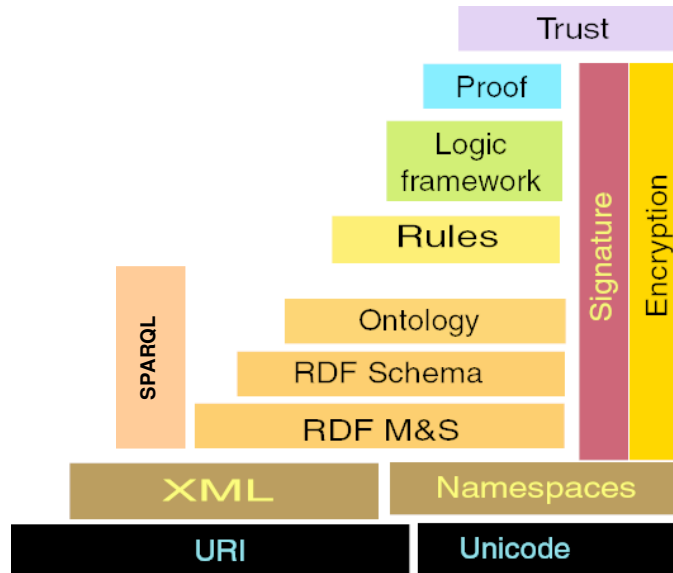


Figure 4 : Les couches du Web sémantique (W3C)

3.2. Le langage RDF (Resource Description Framework)

RDF est un modèle de graphe destiné à décrire formellement les ressources du Web et leurs métadonnées, de sorte à permettre le traitement automatique de telles descriptions. RDF se base sur des vocabulaires formels et précis, tels que ceux exprimés en RDFS ou en OWL. Il constitue le langage de base du Web sémantique. Une des syntaxes (sérialisation) de ce langage est RDF/XML (extrait de Wikipédia).

En annotant des documents non structurés et en servant d'interface pour des applications et des documents structurés (par exemple des bases de données), RDF permet une certaine interopérabilité entre des applications échangeant de l'information non formalisée et non structurée sur le Web.

Un document structuré en RDF est un ensemble de triplets de la forme $\{sujet, prédicat, objet\}$ où :

- Le *sujet* représente la ressource à décrire ;
- Le *prédicat* représente un type de propriété applicable à cette ressource ;
- L'*objet* représente une donnée (un littéral) ou une autre ressource ; c'est la valeur de la propriété.

Le sujet, et l'objet dans le cas où celui-ci est une ressource, peuvent être identifiés par une URI ou être des ressources anonymes. Le prédicat est nécessairement identifié par une URI. Un document RDF ainsi formé correspond à un multi-graphe orienté et étiqueté. Chaque

triplet correspond alors à un arc orienté dont le label est le prédicat, le nœud source est le sujet et le nœud cible est l'objet.

Par exemple, dans la figure 5, nous modélisons par un graphe RDF le fait qu'un cours de bases de données 'sem316' est enseigné par l'enseignante 'Rose Dieng' : (une ellipse représente une ressource, un rectangle représente un littéral et une flèche représente une propriété) :

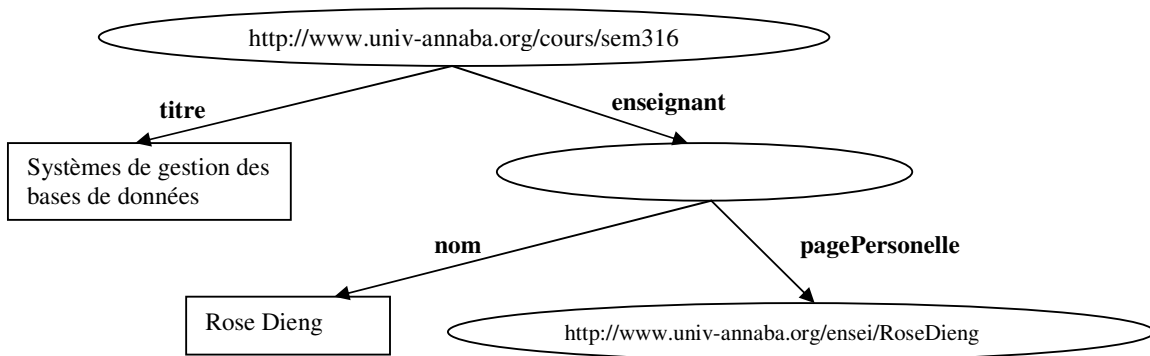


Figure 5 : Exemple d'un graphe RDF

Ce graphe peut être sérialisé en RDF/XML comme suit :

```
<rdf:Description rdf:about='http://www.univ-annaba.org/cours/sem316'>
  <titre>Systèmes de gestion des bases de données</titre>
  <enseignant>
    <nom> Rose Dieng </nom>
    <pagePersonelle>
      <rdf:Description rdf:about='http://www.univ-annaba.org/ensei/RoseDieng'>
        </pagePersonelle>
      </enseignant>
    </enseignant>
  </rdf:Description>
```

3.3. Le langage RDFS (RDF Schema)

RDFS ajoute à RDF la possibilité de définir des hiérarchies de classes et de propriétés dont l'applicabilité et le domaine de valeurs peuvent être contraints à l'aide des propriétés `rdfs:domain` et `rdfs:range`. A chaque domaine applicatif peut être ainsi associé un schéma identifié par un préfixe particulier et correspondant à une URI. Les ressources instances sont ensuite décrites en utilisant le vocabulaire donné par les classes définies dans ce schéma. Les

applications peuvent alors leur donner une interprétation opérationnelle. Néanmoins, RDFS n'intègre pas en tant que tel de capacités de raisonnement.

Les principaux éléments du langage RDFS sont les suivants:

- la classe *rdfs:Class* permet de déclarer une ressource comme étant une classe pour d'autres ressources. Par exemple nous pouvons définir en RDFS la classe *Cours* qui décrit les cours d'une école :

```
<rdfs:Class rdf:ID='Cours'/>
```

L'expression formelle 'Sem316 rdf:type ex:Cours' en RDF traduit l'énoncé en langage naturel : 'Sem316 est un cours'

- la propriété *rdfs:subClassOf* permet de définir des hiérarchies de classes. Par exemple, 'un cours de Master est un cours':

```
<rdfs:Class rdf:ID='CoursMaster'>
  <rdfs:subClassOf rdf:resource='#Cours'/>
</rdfs:Class>
```

- RDFS précise la notion de propriété définie par RDF en permettant de donner un type ou une classe au sujet et à l'objet des triplets. Pour cela, RDFS définit :

- la propriété *rdfs:domain* qui permet de définir la classe des sujets liée à une propriété.
- la propriété *rdfs:range* qui permet de définir la classe ou le type de données des valeurs d'une propriété.

Nous pouvons, par exemple, exprimer en RDFS le fait que l'enseignant d'un cours est une personne et que le nom d'une personne est un littéral (une chaîne de caractères) :

```
<rdf:Property rdf:ID='enseignant' >
  <rdfs:domain rdf:resource='#Cours'/>
  <rdfs:range rdf:resource='#Personne'/>
</rdf:Property>

<rdf:Property rdf:ID='nom'>
  <rdfs:domain rdf:resource='#Personne'/>
  <rdfs:range rdf:resource='&rdfs;Literal'/>
</rdf:Property>
```

Nous pouvons alors exprimer en RDF le fait que la propriété 'enseignant' relie 'sem316' (de type *Cours*) à 'Rose Dieng' (de type *Personne*):


```
<Cours rdf:about=' http://www.univ-annaba.org/cours/sem316'>
  <enseignant>
    <Personne rdf:about=' http://www.univ-annaba.org/Ensei/RoseDieng'>
      <nom>Rose Dieng</nom>
    </Personne>
  </enseignant>
</Cours>
```

Pour résumer, XML peut être vu comme la couche de transport syntaxique, RDF comme un langage relationnel de base et RDFS offre des primitives de représentation de structures ou des primitives ontologiques.

3.4. Le langage OWL

OWL (Web Ontology Language) est un langage basé sur RDF. Il enrichit le modèle RDF Schema en définissant un vocabulaire riche pour la description d'ontologies complexes. OWL est basé sur une sémantique formelle définie par une syntaxe rigoureuse. Il existe trois versions du langage : OWL Lite, OWL DL, et OWL Full.

OWL et RDFS sont tous deux des vocabulaires RDF.

RDFS définit le plus petit nombre de notions et de propriétés nécessaires à la définition d'un vocabulaire simple, essentiellement :

- les notions de classe, de ressource et de littéral
- les propriétés de sous-classe, de sous-propriété, de champ de valeur, de domaine d'application

OWL est un langage beaucoup plus riche qui, aux notions définies par RDF Schema, ajoute les propriétés de classe équivalente, de propriété équivalente, d'identité de deux ressources, de différences entre deux ressources, de contraire, de symétrie, de transitivité, de cardinalité, etc., permettant de définir des relations complexes entre des ressources.

4. Web sémantique & e-Learning

Les propriétés d'une plateforme Web sémantique, à savoir, une description partagée et commune et des métadonnées traitables par les programmes, possibles grâce à la notion d'ontologie, apparaissent suffisamment puissantes pour satisfaire les exigences d'un système pédagogique adaptatif: efficacité, réutilisabilité et pertinence de l'apprentissage. En effet, les ressources pédagogiques sémantiquement annotées peuvent, pour de nouvelles demandes de formation, être combinées en un nouveau cours adapté au profil de l'apprenant et à son

objectif. Le processus pédagogique est basé sur les requêtes Web sémantique et la navigation à travers des ressources pédagogiques activée par un background ontologique. L'ontologie permet le lien entre les besoins des apprenants et les annotations des ressources pédagogiques. Les travaux de (Stojanovic et al., 2001) appuient le fait que le Web sémantique constitue une plateforme pour la recherche et l'extraction de ressources pédagogiques du Web. Les auteurs considèrent que l'annotation sémantique des ressources pédagogiques répond au problème posé par les annotations utilisant des vocabulaires non contrôlés tel que le standard de métadonnées LOM (Learning Object Metadata) ou le Dublin Core. Ils stipulent que les ressources pédagogiques du Web pourraient être traitables par les programmes si celles-ci sont annotées par une ontologie. Ainsi, grâce à une représentation ontologique comprise par les programmes, la réutilisation et la personnalisation des systèmes pédagogiques représentent les points forts de l'utilisation du Web sémantique.

Les travaux de (Vargas-Vera & Lytras, 2008) montrent que la recherche de ressources pédagogiques guidée par une ontologie représente une alternative intéressante à la recherche classique par mots clés. Les résultats de la recherche sont plus adaptés aux apprenants grâce aux capacités de raisonnement intrinsèques aux ontologies. Ces travaux proposent un portail sémantique qui permet aux apprenants d'accéder de manière transparente à des services et à des ressources pédagogiques distribués sur le Web. Ces services peuvent effectuer des raisonnements en arrière plan pour rechercher des ressources pédagogiques pertinentes pour le compte de l'apprenant.

Un framework pour la personnalisation de l'apprentissage en utilisant le Web sémantique est proposé dans (Henze, Dolog, & Nejdl, 2004). Les auteurs montrent comment l'utilisation des descriptions RDF des ressources pédagogiques distribuées permet la génération automatique de cours hypermédias. Ils utilisent le langage de règles TRIPLE pour l'échange des annotations RDF.

Une synthèse sur le nouveau défi des systèmes pédagogiques, à savoir, le Web sémantique éducationnel est présentée dans (Aroyo, & Dicheva, 2004). Les auteurs présentent un état de l'art du e-Learning sémantique et proposent un framework pour assurer l'interopérabilité, la réutilisation et le partage des informations éducationnelles à travers des descriptions sémantiques et la standardisation de la communication à travers des systèmes modulaires et à bases de services.

En résumé, les recherches citées ici et beaucoup d'autres tendent à tirer profit du Web sémantique dans le domaine de l'éducation et en particulier pour la personnalisation des

systèmes pédagogiques, la génération de cours adaptés au profil de l'apprenant et la réutilisation de ressources pédagogique.

5. Métadonnées et e-Learning

Comparativement à l'apprentissage traditionnel, dans lequel l'enseignant joue le rôle d'intermédiaire entre l'apprenant et le matériel d'apprentissage, le scénario d'apprentissage dans le e-Learning est complètement différent : les enseignants ne contrôlent plus la livraison du matériel aux apprenants et ces derniers ont la possibilité de combiner le matériel d'apprentissage en des cours pour leur propre besoin. De ce fait, le contenu d'apprentissage doit exister pour son propre compte. Cependant, indépendamment du temps et du coût pris pour la création du matériel, le contenu demeure inutilisable tant qu'il ne peut être recherché et indexé facilement.

La communauté du e-Learning a convenu d'utiliser des standards de métadonnées. Les métadonnées fournissent un ensemble commun de balises qui peuvent être appliquées à n'importe quelle ressource, indépendamment de qui l'a créée et quels sont les outils qui vont l'utiliser et même où elle sera stockée. Les métadonnées sont des descriptions de ressources ; le balisage des métadonnées permet aux organisations de décrire, indexer et rechercher leurs ressources.

Dans la communauté e-Learning, des standards tels que IEEE-LOM (Learning Object Metadata) (LOM, 2002) et IMS Metadata⁹ sont apparus pour décrire les ressources pédagogiques. Ces modèles de méta-données décrivent comment le matériel d'apprentissage peut être décrit de façon interopérable. Tous les éléments des méta-données nécessaires à la description d'une ressource peuvent être classifiés en plusieurs catégories, chacune offrant un point de vue différent de la ressource. Ce mouvement de standardisation vise à réaliser l'interopérabilité selon différentes dimensions (Duval, 2001) :

- **Fonctions** : Le but est de pouvoir utiliser les mêmes métadonnées, ou des ensembles de métadonnées interopérables, pour la catégorisation, la découverte, la livraison (tout en respectant les droits de l'auteur, de l'éditeur, etc.), l'installation (dans un contexte de documents numériques), etc.
- **Niveaux** : Bien qu'une large majorité des recherches sur le Web se limitent à l'utilisation de mots clefs, il est souhaitable que les mêmes métadonnées puissent être exploitées pour des recherches plus spécifiques (telles que « tous les documents qui

⁹ <http://www.imsglobal.org/metadata/index.html>

traitent de la normalisation en forme Boyce-Codd des bases de données relationnelles, en néerlandais ou en français, pour un niveau universitaire, qui prennent entre 20 et 35 minutes du temps de l'étudiant, ne coûtant rien et étant disponibles sur une plateforme Linux »).

- **Technologies** : L'interopérabilité la plus évidente est celle des technologies. Dans le contexte des standards, on parle de « binding » d'un standard à une technologie. Pour les métadonnées, par exemple, on peut définir des représentations en RDF, XML (DTD ou Schéma), base de données (relationnelle, orientée objet ou autre), etc. Le but de la standardisation est de limiter le nombre de ces « bindings » à un seul, ou bien d'en définir un nombre réduit, chacun pour un contexte technique spécifique, et d'établir les relations entre ces différentes représentations, de sorte que l'on puisse transformer automatiquement les métadonnées d'un contexte (par exemple RDF) à un autre (par exemple base de données).
- **Barrières sémantiques et linguistiques** : Sans doute l'interopérabilité la plus difficile à réaliser est l'interopérabilité sémantique, y compris (mais pas du tout limitée à) l'interopérabilité linguistique. Il est bien évident que, même s'il y avait une interopérabilité parfaite sur le plan des fonctions, niveaux et technologies, le but de la compréhension mutuelle entre producteur de contenu, indexeur, chercheur et autres utilisateurs est complètement évasive si les mots et termes ne sont pas traduits, ou si les vocabulaires sont appliqués d'une manière inconsistante. Pour réaliser une telle interopérabilité, il faut donc que l'on traduise en plusieurs langues, que l'on rende explicite l'application et le sens des vocabulaires, que l'on essaie d'établir des liens entre des classifications et taxonomies relatives, etc.

6. Exemple de Standards de métadonnées : Learning Object Metadata

(LOM)

Le LOM pour Learning Object Metadata est un standard de l'IEEE (Institute of Electrical and Electronics Engineers) qui porte la référence IEEE-1484.12.1-2002. D'un point de vue pratique, le LOM fournit un moyen efficace de repérer et d'indexer des objets d'apprentissage de différents types. Ce standard est seulement la première partie d'un standard multi parties. Cette première partie contient un modèle abstrait (*Abstract Model*) composé de descripteurs ou d'éléments qui sont utilisés pour décrire les objets d'apprentissage et ne concerne pas la réalisation technique de ces éléments. Les éléments LOM seront gérés par différents formats

incluant les tables SQL, les fichiers textes, les méta-balises HTML, etc. De telles réalisations techniques du modèle abstrait dans un format spécifique représentent des 'binding'.

La référence IEEE-1484.12.1 s'accompagne de 3 autres standards IEEE-1484.12.2, IEEE-1484.12.3 et IEEE-1484.12.4 qui traitent respectivement le modèle et des bindings XML et RDF.

LOM a une structure hiérarchique de 78 éléments qui forment un arbre sur 3 niveaux. Il propose 59 champs d'information indépendants. Chaque champ d'information peut apparaître en général 0 ou 1 fois; certains champs particuliers peuvent apparaître un nombre limité de fois 10, 30, 40 ou même 100 fois.

Cependant, il est illusoire de croire pouvoir imposer à tous les auteurs et à tous les documentalistes qui produisent des ressources pédagogiques l'utilisation exhaustive et restrictive du standard LOM.

6.1 Descripteurs du modèles LOM

LOM est un schéma de métadonnées qui comprend 78 éléments hiérarchisés, regroupés en neuf catégories. LOM intègre les 15 champs du Dublin Core¹⁰ qui est un modèle de métadonnées générique défini pour être appliqué à tout type de ressource numérique. Les neuf catégories de LOM sont (extrait de <http://wiki.univ-paris5.fr/wiki/LOM>) :

1. Général

Cette catégorie regroupe les caractéristiques générales indépendantes du contexte qui décrivent la ressource dans son ensemble : identifiant de la ressource (identificateur global unique), son titre, sa description, la liste des langues utilisées, une liste de mots clés, l'étendue de la ressources (géographie, culture), le type de structure (collection, linéaire, hiérarchique, etc.), son niveau de granularité (de 1 à 4, 1 désignant un cours entier).

2. Cycle de vie

Caractéristiques relatives au cycle de vie : historique de la ressource (version), état de la ressource (du point de vue de son achèvement ou de sa disponibilité) et d'autre part les personnes qui ont affecté cette ressource durant son évolution (instance unique). Toutes les sous catégories sont fondamentales pour pouvoir faire fonctionner un système d'information structurée avec métadonnées. La plupart de ces catégories sont bien connues des bibliothécaires, mais doivent être modifiées en conséquence.

¹⁰ <http://dublincore.org/>

3. **Méta-métaDonnées**

Caractéristiques de la description elle-même (spécifiques à l'enregistrement des métadonnées) : la date de contribution, les différents contributeurs à l'établissement des métadonnées, etc. Il est à souligner que la non confusion entre création des métadonnées et création de ressources doit être soulignée comme primordiale. Cette catégorie est également fondamentale pour pouvoir faire fonctionner un système d'information structurée avec métadonnées. Les sous-ensembles que contient cette catégorie étaient pour la plupart inconnus des bibliothécaires qui y ont été confrontés lorsqu'ils ont construit des bibliothèques numériques. Font partie de cette catégorie :

- une entrée unique correspondant au système d'identification (comme ARIADNE, URL, etc.),
- la langue de la description,
- les contributions, y compris les entités (personnes, organisations) ayant contribué et la date de leur contribution (un type spécifique de contribution étant la validation) ;

4. **Technique**

Cette catégorie définit les exigences techniques en terme de navigateur, de système d'exploitation, ou des caractéristiques comme le type des données, le format (type MIME), la taille de la ressource pédagogique (en octets), sa localisation physique en vue de son téléchargement (URL Uniform Resource Locator ou URI Uniform Resource Identifier), sa durée, des informations complémentaires sur l'installation, etc.

5. **Pédagogie**

Cette catégorie se décline en treize sous-catégories pédagogiques. C'est souvent par ces sous-catégories que l'on améliore l'exploitation du contenu pédagogique. Il est par contre fondamental de bien comprendre que les deux items *densité sémantique* et *description* ne sont pas destinés à décrire la ressource pour la retrouver ultérieurement ou la classer.

- *Type d'interactivité* : le type d'interaction entre la ressource et l'utilisateur typique (Active, Expositive, Undefined).
- *Types de ressources pédagogiques* : le type pédagogique (exercice, simulation...) peut être présent plusieurs fois.
- *Niveaux d'interactivité* : degré d'interactivité.
- *Densité sémantique* : (de très basse à très élevée).

- *Rôle présumé de l'utilisateur final* : public ciblé, prioritairement destinataire de la ressource (professeur, auteur, apprenant ou administrateur).
- *Niveau* : niveau du public ciblé.
- *Proposition d'utilisation* : (contexte) description relative à l'utilisation pédagogique de la ressource.
- *Tranche d'âge* : âge de l'apprenant.
- *Difficulté* : difficulté de la ressource.
- *Durée d'apprentissage moyen* : durée approximative ou typique d'apprentissage pour le ou les niveaux indiqués.
- *Description* : commentaires sur l'utilisation de la ressource.
- *Language* : langue de l'utilisateur.

6. **Droits**

Caractéristiques exprimant les conditions d'utilisation : les droits (copyright) liés à la ressource, les conditions légales d'utilisation, éventuellement son coût.

7. **Relation**

Caractéristiques exprimant les liens avec d'autres ressources pédagogiques en précisant la nature de la relation (« ...est requis par... », « ...est une partie de... »).

8. **Annotation**

Annotations ou commentaires sur l'utilisation pédagogique de la ressource.

9. **Classification**

Cette catégorie, essentielle pour proposer un schéma ouvert, permet d'indiquer où se place la ressource dans un système de classification donné. Dans ARIADNE par exemple, on utilise cette catégorie pour placer la ressource dans une classification sémantique dynamique, en indiquant la discipline et les concepts présentés dans la ressource.

Il est important de noter que de nombreux éléments peuvent recevoir plusieurs valeurs (tout comme certaines catégories). Par exemple, en utilisant plusieurs classifications, on pourra préciser où se place la ressource dans différentes taxinomies sémantiques, telles que *Universal Decimal Classification*, *Dewey Decimal Classification*, etc.

Par ailleurs, un des buts explicites du LOM est de faciliter le multilinguisme, aussi bien au sein des ressources que dans leur description. A chaque fois que l'on doit saisir des textes libres (plutôt que de sélectionner des valeurs d'un vocabulaire prédéfini), la notion de '*LangString*' proposée par le schéma permet de définir un texte en différentes langues. Une

indication précise la langue utilisée ainsi que son éventuelle variante géographique, comme par exemple *fr-fr* et *fr-ca* pour « français pratiqué en France » et « français pratiqué au Canada » (Duval, 2002). Ces descripteurs ne sont pas obligatoires mais certains peuvent être répétés.

6.2. Limitations du modèle LOM

LOM est considéré comme le modèle de métadonnées de référence pour décrire des ressources pédagogiques. Cependant, de nombreuses limitations sont constatées :

- **Lourdeur du modèle :**

Arbitrage à réaliser entre le volume et la pertinence des champs du modèle.

- **Complexité technique du modèle :**

Difficulté de remplir les 78 champs du modèle.

- **Insuffisance en termes d'aspects pédagogiques :**

Les 9 catégories de LOM renferment implicitement des modèles pédagogiques sous-jacents, approuvés de facto. Dans l'état présent, ce modèle ressort davantage d'un catalogue de ressources que d'une aide à la mise en place de dispositifs efficaces de transfert ou de construction de savoirs.

- Manque de définition de ce qu'est une ressource pédagogique : Le LOM s'applique indistinctement à des objets qui ont des natures et des fonctionnalités très différentes (ressources, activités et unité d'apprentissage) : cela empêche leur description en tant que tels
- Excès de généralité : Le LOM offre des champs de données qui manquent de précision (Ex : la catégorie "classification" permet à chaque institution de classer l'objet suivant son propre catalogue).
- Rôle de descripteur n'est pas assez poussé pour prendre en compte des ressources plus interactives comme celles mises en œuvre dans l'évaluation des connaissances.
- Manque de pertinence de certains champs en terme d'indexation : Certains descripteurs (tels que la "densité sémantique") apparaissent trop subjectifs, donc difficiles à informer
- Focalisation exclusive sur les contenus sans prise en considération de la démarche pédagogique à leur associer.
- Incompatibilité avec des représentations structurées de l'information : Le LOM ne permet pas, dans sa forme actuelle, de représentations enrichies de connaissances (liens, ontologies...)

- Aspects déterministe de certaines listes de valeurs
Par exemple, la valeur du champ "Niveau d'agrégation" induit celle du champ "structure":
- Ambiguïté dans l'utilisation de la catégorie "Relation"

Conclusion : Dans ce chapitre nous avons présenté le Web sémantique en explicitant ses objectifs et ses langages. En effet, Les spécifications RDF, RDFS et OWL combinées forment une partie de la plateforme Web Sémantique basée sur les concepts suivants :

- un schéma de nommage global (les URI)
- une syntaxe standard pour décrire des connaissances (RDF)
- un moyen standard de décrire les propriétés de ces connaissances (RDF Schema)
- un moyen standard de décrire les relations entre ces connaissances (OWL)
- un moyen de créer un réseau de confiance sécurisé à propos de ces connaissances

Nous avons également montré l'intérêt que pouvait avoir le Web sémantique pour le e-Learning et particulièrement pour les systèmes pédagogiques adaptatifs qui ont comme mots clés : efficacité, réutilisabilité et apprentissage pertinent. Nous avons présenté un aperçu des recherches qui essaient d'utiliser les ontologies et les standards de métadonnées pour améliorer la réutilisation des ressources pédagogiques au sein de systèmes pédagogiques.

Nous avons également présenté le standard LOM, d'une part parce qu'il représente un modèle de métadonnées incontournable dans le domaine de l'éducation et d'autre part parce que nous nous intéressons, dans le cadre de notre travail, uniquement aux ressources pédagogiques du Web annotées par le standard LOM.

Partie II : Un environnement pédagogique adaptatif: le système OrPAF

Dans notre travail de thèse, nous nous intéressons à la construction d'un environnement pédagogique centré sur l'apprenant, qui permet de monter efficacement des formations en réutilisant des ressources pédagogiques existantes du Web. Ce type d'environnements intègre des agents humains (des enseignants, des experts du domaine, des apprenants, etc.) et des agents logiciels qui fournissent aux apprenants des situations d'interaction leur permettant d'accéder à des ressources pédagogiques dans le but d'atteindre leurs objectifs de formation.

La partie logicielle de notre environnement est développée pour assurer trois fonctions :

1. Représentation, gestion et exploitation des connaissances (dans notre cas, basées sur les modèles et les technologies du Web sémantique) (Yessad et al., 2008a ; Yessad et al., 2008b ; Yessad et al., 2008c);
2. Recherche de ressources pédagogiques pertinentes du Web (dans notre cas, recherchées à partir du corpus ARIADNE) (Yessad et al., 2005a) ;
3. Génération et présentation de cours adaptatifs (dans notre cas, à travers la technologie des hypermédias adaptatifs) (Yessad et al., 2005b ; Yessad et al., 2006).

1. Cadre d'analyse et méthodologie de développement

Nous présentons tout au long de cette deuxième partie les choix techniques et didactiques qui sous-tendent notre approche ainsi que la démarche méthodologique utilisée pour le développement de notre environnement pédagogique adaptatif.

Au-delà des contraintes classiques que doit vérifier un système informatique, les environnements pédagogiques embarquent des fonctionnalités supplémentaires liées à l'objectif de susciter ou d'accompagner un apprentissage humain. Dans ce domaine, plusieurs méthodologies ont été proposées (Reigeluth, 1999) pour assister les équipes de développement. Dans notre travail, nous réutilisons une méthodologie de développement proposée dans (Lester et al., 2005) et qui est une extension du processus de développement « en cascade » (Boehm, 1981). Ce modèle de développement de systèmes pédagogiques comporte six phases, à savoir, l'étude de faisabilité, l'analyse, la conception, la construction, la réalisation et l'évaluation du système (voir figure 6).

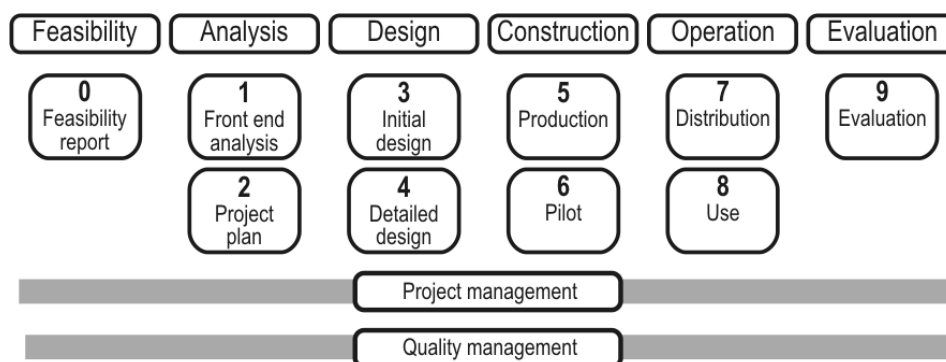


Figure 6: Phases de développement d'un environnement pédagogique (Lester et al., 2005)

2. Problématique de recherche

L'intérêt principal de notre recherche est de proposer un environnement pédagogique qui permet de monter efficacement des formations destinées à un public averti, par exemple des employés impliqués dans un processus de formation continue, qui sont régulièrement confrontés à de nouveaux besoins de formation et qui ne disposent d'aucune assistance humaine. Il s'agit également d'intégrer une démarche didactique au développement technique d'un tel environnement.

Le point de départ de notre travail a été d'identifier les situations d'apprentissage pour lesquelles notre approche serait utile et apporterait une valeur ajoutée par rapport aux approches et systèmes déjà existants.

Besoin 1 : Adaptation à de nouveaux besoins de formation

Premièrement, nous avons voulu réaliser un environnement qui s'adapte avec un minimum d'efforts technique et didactique (de la part des enseignants et des experts du domaine) à de nouveaux besoins de formation. Ce besoin d'adaptation est réel car nous constatons qu'aujourd'hui de plus en plus d'entreprises prévoient régulièrement dans leur calendrier des sessions de formation pour leurs employés. Les entreprises organisent souvent ces formations en présentiel sur site ou délocalisées (par exemple sur des campus universitaires) ou ont recours à des plateformes d'apprentissage avec tuteurs (Learning Management Systems). Ces deux types de solutions (formation en présentiel ou plateforme d'apprentissage) coûtent énormément aux entreprises et donc ne conviennent pas aux contraintes économiques et aux besoins d'efficacité de celles-ci. Notre réponse à ce problème est de proposer un environnement pédagogique adaptatif qui s'adapte efficacement aux nouveaux besoins de formation des employés ainsi qu'à leurs profils. Dans ce contexte, un effort de standardisation des différents composants de l'environnement pédagogique nous a paru indispensable dans le but de faciliter son adaptation et réutilisation pour de nouvelles formations (Voir chapitre 4).

Besoin 2 : Réutilisation des ressources pédagogiques du Web

Ensuite, nous avons constaté que le Web regorge de ressources pédagogiques qui pourraient être réutilisées pour de nouveaux besoins de formation. Cette réutilisation est généralement manuelle. Or, ce travail manuel est à la fois coûteux et complexe. Fidèles à notre logique d'efficacité, nous voulons apporter une assistance automatique aux experts et aux enseignants pour rendre la recherche et la réutilisation des ressources pédagogiques du Web plus pertinente et efficace. Pour ce faire, nous avons restreint notre recherche à une partie du Web, en l'occurrence, la partie annotée qui contient des ressources pédagogiques décrites avec le standard LOM (par exemple le corpus ARIADNE). En effet, la réutilisation des ressources pédagogiques disponibles sur le Web (cours, exercices, simulations, etc.) pourrait améliorer l'efficacité du système en économisant le temps et l'effort de développement de nouvelles ressources pédagogiques. Cette réutilisation est conditionnée par la capacité d'adapter ces ressources aux différents contextes de formation (voir chapitre 5)

Besoins 3 : Présentation personnalisée du contenu pédagogique

Enfin, nous voulons offrir aux apprenants des parcours d'apprentissage adaptatifs où une approche constructiviste et un mécanisme d'apprentissage par la découverte soient possibles. Nous voulons mettre en place un mécanisme d'interactions apprenant-système qui met l'apprenant au centre du système, en s'adaptant à ses caractéristiques et qui dans le même temps le responsabilise. Notre recherche s'est orientée naturellement vers la technologie des

hypermédias éducationnels adaptatifs. En effet, les hypermédias éducationnels adaptatifs, présentés au chapitre 2, nous permettent de construire des cours personnalisés et favorisent une approche constructiviste de l'apprentissage (voir chapitre 5).

3. Scénarios d'utilisation de l'environnement pédagogique adaptatif

Dans un environnement pédagogique, différents acteurs peuvent intervenir : des experts pour ajouter, supprimer ou modifier les connaissances sur le domaine et la pédagogie, des enseignants pour rechercher et annoter les ressources pédagogiques, et des apprenants impliqués dans un processus de formation avec des objectifs et des contraintes précises.

Scénario 1 : Description des connaissances par les experts

D'une part, les experts du domaine ont besoin de décrire de manière informelle les connaissances de leur domaine d'apprentissage (les mathématiques, la biologie, etc.). Ces connaissances sont des notions du domaine (par exemple la notion de polynôme en Algèbre), des relations qui relient ces notions et des règles qui permettent de raisonner sur les notions et les relations.

D'autre part, les experts en pédagogie ont besoin de décrire les connaissances relatives aux différentes stratégies pédagogiques à adopter en fonction des situations d'apprentissage et des profils apprenants (par exemple illustrer une notion du domaine par l'analogie pour un apprenant débutant).

Les connaissances du domaine et les connaissances de la pédagogie sont formalisées par des ingénieurs cogniticiens. Des techniques et des langages de représentation et de gestion de ces connaissances sont utilisés. Nous détaillerons cet aspect dans le chapitre 4.

Scénario 2 : Recherche de ressources pédagogiques par les enseignants

A chaque fois qu'un besoin de formation apparaît, les enseignants se posent la question suivante : Quelles sont les notions que l'apprenant doit acquérir dans le cadre de cette nouvelle formation ? Existe-il déjà sur le Web des ressources pédagogiques qui pourraient être réutilisées dans le cadre de cette formation ? Comment et à partir de quels corpus rechercher ces ressources pédagogiques ? Quelle est la pertinence de ces ressources pour l'apprenant ? Les solutions conceptuelles que nous apportons à ces questions sont exposées dans le chapitre 5.

Scénario 3 : Apprentissage adaptatif de l'apprenant

Un environnement pédagogique adaptatif fournit à l'apprenant les moyens d'atteindre efficacement son objectif en évitant de lui présenter des notions ou des ressources pédagogiques qui ne servent pas son objectif ou ne s'adaptent pas à son profil. L'apprenant a

besoin d'être guidé dans son parcours d'apprentissage sans pour autant perdre le contrôle sur celui-ci. Il s'agit de trouver un compromis entre le guidage absolu de l'apprenant dans les tuteurs intelligents et la liberté absolue de l'apprenant dans hypermédias éducationnels. Ce problème est difficile et relève du domaine de la psychologie pédagogique mais des réponses technologiques existent (Voir chapitre 5).

4. Environnement pédagogique adaptatif: le système OrPAF

Le système OrPAF (Organisateur de Parcours Adaptatifs de Formation) est la solution qui a été construite tout au long de notre travail de thèse pour répondre à notre problématique de recherche. Il s'agit d'un environnement pédagogique adaptatif centré sur l'apprenant où la gestion d'un modèle apprenant permet à la fois de construire des parcours d'apprentissage adaptatifs (les parcours représentent la structure du cours) et de présenter à l'apprenant des ressources pédagogiques pertinentes (les ressources pédagogiques représentent le contenu du cours). Cette adaptation de la structure et du contenu du cours constitue une personnalisation du système au profil de l'apprenant. Le cours est présenté sous la forme d'un hypermédia que l'apprenant parcourt jusqu'à ce qu'il atteigne son objectif. La personnalisation de cet hypermédia préserve la liberté de parcours du cours par l'apprenant tout en réduisant les problèmes de désorientation et de surcharge cognitive intrinsèques à la technologie hypermédia. Au-delà de la personnalisation des cours destinés aux apprenants, nous nous sommes intéressés au problème de l'adaptation efficace de l'environnement pédagogique à différentes demandes de formation. Ceci nous a conduits à nous intéresser et à utiliser les standards des domaines de l'éducation et de la gestion des connaissances. En plus, le besoin d'interopérabilité des composants du système OrPAF avec d'autres systèmes du Web nous a conduits à nous tourner vers les technologies du Web sémantique et en particulier les ontologies. Ces dernières nous permettent de représenter les différentes connaissances (connaissances du domaine, connaissances sur la pédagogie, connaissances sur l'apprenant) et d'annoter sémantiquement les ressources pédagogiques recherchées du Web pour les réutiliser dans le processus d'apprentissage.

4. Représentation des connaissances dans OrPAF

Nous avons expliqué dans le chapitre 3, l'intérêt d'utiliser les ontologies et les technologies du Web sémantique pour (i) répondre à la faiblesse sémantique de la représentation des connaissances dans le domaine de l'apprentissage humain et (ii) associer aux ressources pédagogiques des descriptions formelles permettant de mettre en œuvre des raisonnements formels (recherche de ressources, compositions de ressources, etc.). Dans le cadre de notre travail de thèse, l'opérationnalisation des ontologies nous permet de réaliser des besoins fonctionnels et d'autres non fonctionnels.

Dans le cadre des besoins fonctionnels, l'utilisation des technologies Web sémantique nous permet de rechercher des ressources pédagogiques adaptées à l'apprenant, de composer de nouvelles ressources à partir de ressources existantes permettant d'avoir des parcours d'apprentissage personnalisés et d'adapter l'interaction entre le système et l'apprenant en fonction des objectifs de celui-ci.

Dans la cadre des besoins non fonctionnels, les technologies et les modèles du Web sémantique améliorent l'interopérabilité de l'environnement pédagogique avec d'autres systèmes pédagogiques du Web, en particulier, la réutilisation de ressources pédagogiques et de connaissances (par exemple les connaissances d'un domaine) créées et gérées dans ces

systèmes pédagogiques. D'ailleurs, une ontologie représente une définition explicite d'une conceptualisation partagée d'un domaine de connaissances et donc permet de limiter les projections possibles entre les symboles utilisés et leur signification, ce qui améliore la compréhension automatique des modèles et donc leur réutilisation.

Dans ce chapitre, nous allons présenter comment nous avons utilisé les ontologies pour décrire les connaissances de l'environnement pédagogique adaptatif OrPAF.

1. Modélisation des connaissances dans OrPAF (Yessad et al., 2008b)

Dans le contexte du système OrPAF, l'utilisation des ontologies nous permet de :

- Lever les ambiguïtés dans les échanges entre les personnes, entre les personnes et les composants du système, et entre les composants du système ;
- Améliorer la réutilisation des connaissances liées à l'apprentissage humain (connaissances sur le domaine, connaissances sur l'apprenant, etc.) ;
- Décrire de manière formelle et partagée les notions à appréhender lors d'une formation (le modèle du domaine) ;
- Décrire de manière formelle et partagée les stratégies pédagogiques à appliquer en fonction des situations d'apprentissage et des apprenants (le modèle pédagogique) ;
- Décrire de manière formelle et partagée les caractéristiques des apprenants (le modèle apprenant) ;
- Ces connaissances (du domaine, de la pédagogie et de l'apprenant) servent de vocabulaire formel et partagé pour annoter les ressources pédagogiques du Web et permettre l'accès aux ressources les plus pertinentes.

Pour répondre au besoin d'adaptation de l'environnement pédagogique à un nouveau besoin de formation, c'est-à-dire, à un nouveau domaine (ou discipline), à une nouvelle stratégie pédagogique (par exemple la stratégie inductive pour présenter les activités du cours) ou à un nouveau profil apprenant (par exemple un apprenant avec des difficultés particulières d'accessibilité au système), nous avons fait le choix de décrire toutes les connaissances à gérer dans le système à deux niveaux ontologiques : un niveau méta-modèle où nous décrivons une ontologie dédiée à l'apprentissage en ligne (une ontologie du e-Learning) et un niveau modèle où sont décrites des ontologies dédiées à une formation particulière (un domaine particulier et une stratégie pédagogique particulière), des modèles apprenant (des profils apprenants particuliers) et des annotations pour les ressources pédagogiques du Web.

Le méta-modèle sert à formaliser les connaissances liées au domaine de l'apprentissage en ligne (l'ontologie du e-Learning) et constitue le vocabulaire commun utilisé pour construire

les ontologies et les modèles du niveau inférieur. En d'autres termes, ces derniers sont construits par instanciation des différents concepts définis dans l'ontologie du e-Learning. Etant donné, que certains concepts sont redondants et sont représentés dans plusieurs modèles (modèle du domaine, modèle pédagogique et modèle apprenant), cette approche de modélisation à deux niveaux nous permet de décrire de manière partagée et cohérente les concepts des différents modèles.

Ainsi, la conception d'un méta-modèle permet de réduire l'hétérogénéité :

(1) entre les langages qui servent à décrire les modèles. Cette hétérogénéité se manifeste généralement en quatre points : la syntaxe, la représentation logique, la sémantique des primitifs et l'expressivité du langage.

(2) entre les modèles. C'est-à-dire qu'une différence dans la façon d'interpréter un domaine implique une différence entre les concepts et leurs relations qui se manifeste à la fois au niveau de la spécification et au niveau de la conceptualisation du domaine, comme par exemple avoir plusieurs termes pour désigner le même concept. En effet, nous pouvons avoir dans le système une connaissance qui spécifie que l'« apprenant A doit apprendre la notion désignée par le terme t_1 » et une autre connaissance qui spécifie que « la ressource pédagogique R est un exercice sur la notion désignée par le terme t_2 ». Si dans l'ontologie nous disposons de la connaissance qui spécifie que les termes t_1 et t_2 sont relatifs à la même notion alors R devient une ressource pertinente pour l'apprenant A.

2. Acteurs du système OrPAF

Dans le système OrPAF, différents types d'acteurs ont différents types d'interactions avec la base de connaissances (voir figure 7).

Un acteur Producteur : ce rôle est joué par l'expert qui produit la connaissance. Il spécifie de manière informelle ou semi-formelle (structurée) tous les concepts qui appartiennent au domaine de la formation, à la pédagogie adoptée et ceux qui caractérisent les profils apprenants. Dans le système OrPAF, différentes personnes jouent le rôle d'expert : les experts du domaine de la formation (par exemple un mathématicien, un biologiste, etc.), les experts en pédagogie, les experts en psychologie didactique, etc. Ces personnes détiennent des connaissances statiques et opérationnelles dans leurs domaines de compétences.

Un acteur Consommateur : ce rôle est joué par les apprenants. Les apprenants empruntent des parcours d'apprentissage formalisés dans les différentes ontologies (l'ontologie du e-Learning et les ontologies de formation). Ils consultent également des ressources pédagogiques adaptées à leur profil. La pertinence de ces ressources pédagogiques pour un

contexte de formation donné est identifiée grâce leurs annotations sémantiques basées sur les différentes ontologies.

Un acteur modélisateur : ce rôle est joué par l'ingénieur de la connaissance qui utilise un langage de représentation de connaissances (dans notre cas, un langage ontologique) pour formaliser les connaissances informelles ou semi-formelles spécifiées par les experts.

Dans le système OrPAF, les enseignants ont un double rôle : ils recherchent sur le Web des ressources pédagogiques et les annotent (dans ce cas, ils sont producteurs de connaissances). Cependant, le processus d'annotation est guidé par le système car les enseignants sont contraints d'utiliser exclusivement le vocabulaire spécifié dans les ontologies (dans ce cas, ils sont consommateurs de connaissances). Par exemple, si un enseignant annote une ressource pédagogique qui définit la notion 'Polynôme' du domaine de l'algèbre, il doit utiliser la notion *Polynomial* ou une notion « proche » de *Polynomial* dans l'ontologie qui décrit le domaine de l'algèbre. Cette contrainte assure l'utilisation d'un vocabulaire commun entre les différents acteurs et composants du système et réduit ainsi le risque d'hétérogénéité sémantique.

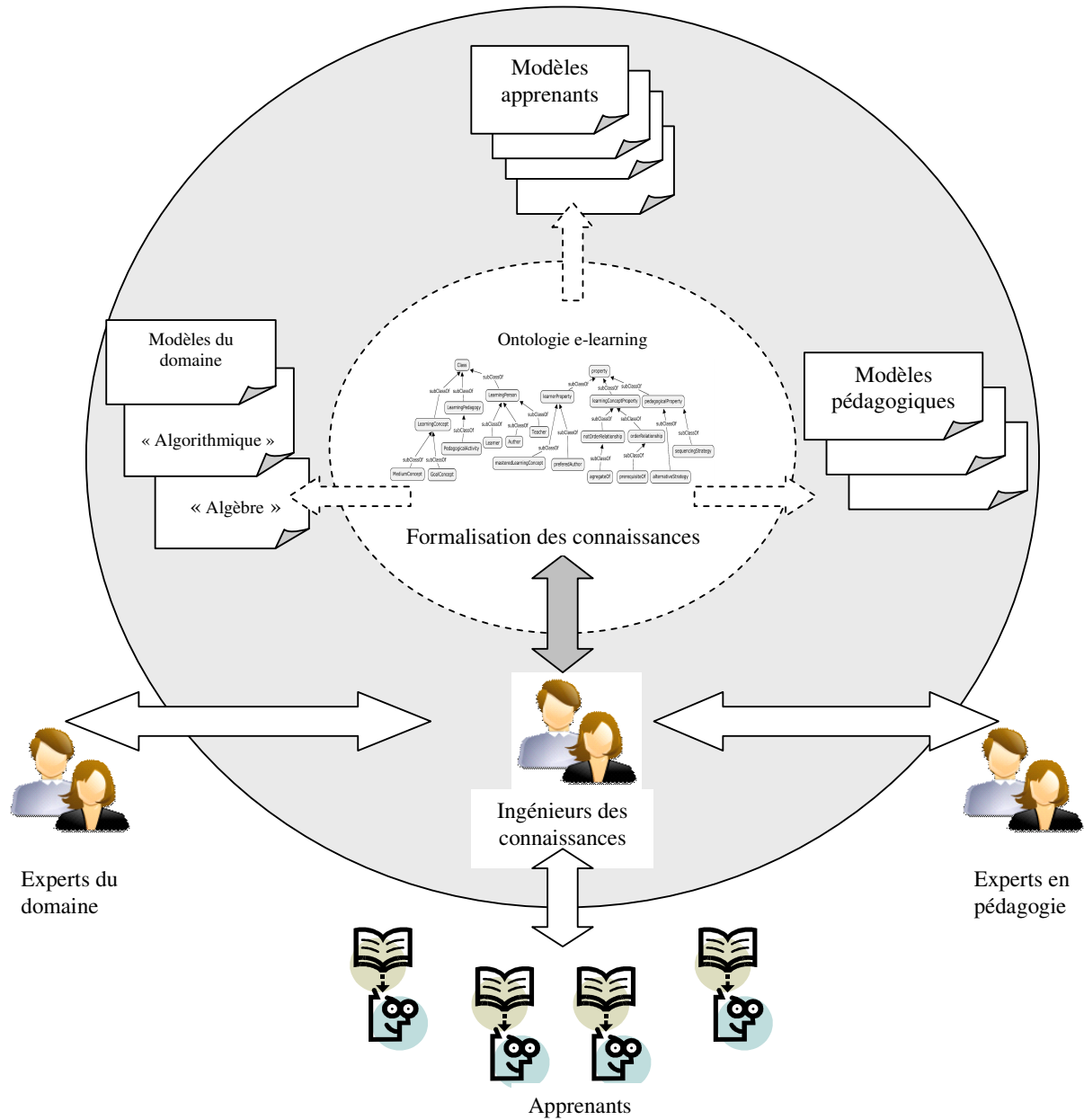


Figure 7: Les acteurs et l'organisation des connaissances dans OrPAF

3. Ontologies du système OrPAF

Une ontologie comprend deux catégories d'entités : les concepts qui sont organisés hiérarchiquement et les relations sémantiques qui relient les concepts. Pour les besoins de notre recherche, nous avons défini deux types d'ontologies : une ontologie du e-Learning et des ontologies de formation.

3.1. Ontologie du e-Learning

3.1.1. Définition informelle

Cette ontologie représente un modèle générique qui formalise des connaissances du domaine de l'apprentissage en ligne (e-Learning) indépendamment de la formation visée (la discipline à apprendre). Elle décrit deux types de connaissances: des concepts et des propriétés organisés hiérarchiquement. Pour chaque propriété, on définit un concept de départ (le domaine de la propriété) et un concept d'arrivé (la valeur ou le co-domaine de la propriété). Ces connaissances sont instanciées pour décrire les notions d'un domaine (par exemple les notions de l'algèbre), des stratégies pédagogiques (par exemple la stratégie inductive), des caractéristiques de l'apprenant (par exemple son niveau de connaissances) et des annotations conceptuelles pour les ressources pédagogiques (par exemple le type pédagogique de la ressource, son auteur, etc.).

La figure 8 représente un extrait de l'ontologie du e-Learning qui a été construite dans le cadre du projet OrPAF. Elle est composée de différents concepts et propriétés dédiés à la formalisation de différentes connaissances : connaissances sur le domaine, connaissances sur la pédagogie, connaissances sur l'apprenant et connaissances sur les ressources pédagogiques.

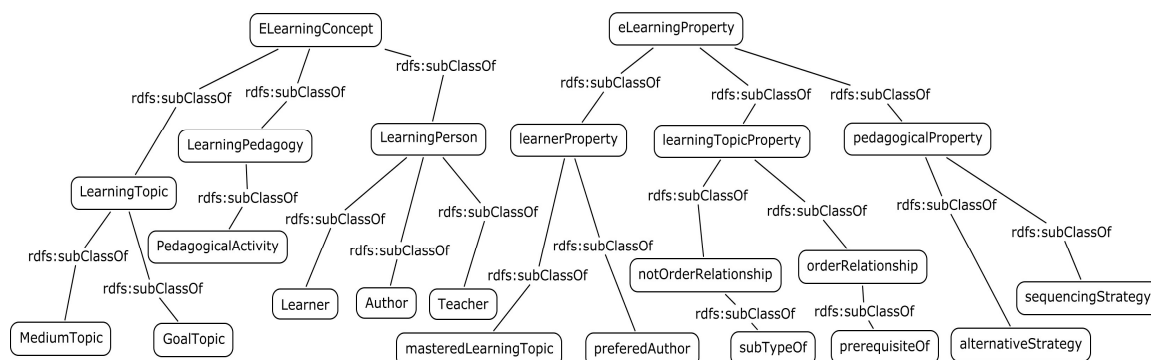


Figure 8: Un extrait de l'ontologie du e-Learning

Nous notons qu'une propriété qui relie deux concepts est orientée du concept « domaine » vers le concept « co-domaine ». Par exemple, la propriété *preferredAuthor* relie le concept *Learner* (domaine de la propriété) au concept *Author* (le co-domaine de la propriété).

3.1.2. Définition formelle

Formellement, nous définissons l'ontologie du e-Learning LO (pour Learning Ontology) comme suit :

LO = (C, P, H^c, H^p, Signature, Rules), où:

- C et P sont deux ensembles disjoints. Les éléments de C sont appelés concepts et les éléments de P sont appelés propriétés. Chaque élément de ces deux ensembles est identifié par une URI (Unique Resource Identifier).
- H^c est la hiérarchie qui relie les différents concepts. Elle représente un graphe acyclique et orienté avec $H^c \subseteq C \times C$. L'écriture $H^c(c_1, c_2)$ signifie que c_1 est un sous-concept de c_2 .
- H^p est la hiérarchie qui relie les différentes propriétés. Elle représente un graphe acyclique et orienté avec $H^p \subseteq P \times P$. $H^p(p_1, p_2)$ signifie que p_1 est une sous-propriété de p_2 .
- Signature: $P \rightarrow C \times C$ est une fonction qui définit les concepts de départ et d'arrivé d'une propriété p. La fonction domain: $P \rightarrow C$ avec $\text{domain}(p) = \prod 1(\text{Signature}(p))$ retourne le domaine de p et la fonction range(P) = $\prod 2(\text{Signature}(P))$ retourne son co-domaine.
- Rules est un ensemble de règles sur C et P.

Les concepts et les propriétés sont considérés comme des objets généraux représentant des connaissances générales dans le domaine du e-Learning. Cette ontologie du e-Learning est instanciée pour construire des ontologies de formation spécialisées : le modèle du domaine, le modèle pédagogique et le modèle apprenant. Elle est également utilisée pour annoter les ressources pédagogiques en construisant des annotations conceptuelles (voir chapitre 5).

Contrairement à l'ontologie du e-Learning, les ontologies de formation, c'est-à-dire, le modèle du domaine, le modèle pédagogique et le modèle apprenant décrivent des objets concrets et spécifiques. Par exemple, nous retrouvons dans l'ontologie du e-Learning le concept *MediumTopic* (une notion intermédiaire) et *GoalTopic* (une notion cible) qui représentent des types de notions alors que dans le modèle du domaine nous décrivons des notions concrètes qui appartiennent à un domaine de formation précis, par exemple *Operator* et *Statement* pour le domaine de l'algorithmique. En outre, le fait que *Operator* et *Statement* soient de type *MediumTopic* exprime explicitement la relation qui existe entre l'ontologie du e-Learning et le modèle du domaine : le modèle du domaine est une instance de l'ontologie du e-Learning.

Les concepts de l'ontologie du e-Learning sont organisés hiérarchiquement avec un concept racine et abstrait appelé *ELearningConcept*. Le concept *ELearningConcept* possède trois sous-concepts: *LearningTopic*, *LearningPerson*, *LearningPedagogy* qui sont utilisés respectivement ainsi que leurs sous-concepts pour construire le modèle du domaine, le modèle pédagogique et le modèle apprenant. De la même manière, les propriétés sont organisées hiérarchiquement avec une propriété racine abstraite appelée *eLearningProperty*. Cette

dernière est spécialisée en trois sous-propriétés appelées : *learningTopicProperty*, *learnerProperty* et *pedagogicalProperty*.

3.1.3. Langage de description de l'ontologie du e-Learning

Dans le système OrPAF, l'ingénieur des connaissances prend en charge la représentation des connaissances exprimées par l'expert en utilisant des formalismes de représentation.

Les standards actuels de représentation ontologique permettent :

- La formalité (logique du premier ordre pour RDF et logique de description pour OWL) ;
- La multi-instanciation : instances de plusieurs classes ;
- La réutilisation de plusieurs schémas grâce aux namespaces ;
- Des règles d'inférence standards pour raisonner sur les connaissances (OWL).

Dans notre travail, le formalisme choisi pour la modélisation des connaissances doit permettre:

- l'exploitation non ambiguë des différentes connaissances (les différentes ontologies) ;
- la possibilité d'adapter les différents modèles (domaine, pédagogique et apprenant) ou de les substituer par d'autres modèles correspondants à une nouvelle demande de formation;
- la validité des trois modèles de connaissances par rapport au méta-modèle (l'ontologie du e-Learning)
- l'annotation des ressources pédagogiques par association des différents modèles (plusieurs schémas). Cela signifie l'utilisation de formalismes de représentation modulaires.

La description de l'ontologie du e-Learning *LO* est faite dans le langage OWL. Etant donné, d'une part, les caractéristiques de ce langage de description d'ontologies (voir chapitre 3) et d'autre part les besoins cités ici, OWL est parfaitement adapté pour décrire les concepts de *C*, les propriétés de *P* et leur signature *Signatures*, leurs liens de subsomption H^C et H^P et les règles *Rules* qui s'appliquent sur les concepts et les propriétés.

3.1.4. Exemples de sérialisation de l'ontologie du e-Learning dans OWL

La classe *ELearningConcept* représente le concept racine de tous les concepts de *C*.

```
<owl:Class rdf:ID="ELearningConcept"/>
```

Les concepts *LearningTopic*, *GoalTopic* et *MediumTopic* sont utilisés pour instancier des notions du domaine. *GoalTopic* et *MediumTopic* sont des sous-concepts de *LearningTopic* et permettent pour le premier d'instancier des notions cibles (des notions que l'apprenant cherche à atteindre) et pour le deuxième d'instancier les autres notions du domaine.

```
<owl:Class rdf:ID="LearningTopic">
    <rdfs:subClassOf rdf:resource="#ELearningConcept"/>
</owl:Class>
<owl:Class rdf:ID="GoalTopic">
    <rdfs:subClassOf rdf:resource="#LearningTopic"/>
</owl:Class>
<owl:Class rdf:ID="MediumTopic">
    <rdfs:subClassOf rdf:resource="#LearningTopic"/>
</owl:Class>
```

Les concepts *Learner* et *Author* sont des sous-concepts du concept *LearningPerson* qui représente de manière abstraite toutes les personnes impliquées dans un environnement d'apprentissage humain (apprenant, auteur, enseignant, expert, etc.).

```
<owl:Class rdf:ID="LearningPerson">
    <rdfs:subClassOf rdf:resource="#ELearningClass"/>
</owl:Class>
<owl:Class rdf:ID="Learner">
    <rdfs:subClassOf rdf:resource="#LearningPerson"/>
</owl:Class>
<owl:Class rdf:ID="Author">
    <rdfs:subClassOf rdf:resource="#LearningPerson"/>
</owl:Class>
```

Le concept *LearningPedagogy* décrit de manière abstraite les notions liées au volet pédagogique de l'apprentissage. Il est spécialisé par le concept *PedagogicalActivity* qui sert à instancier les différentes activités pédagogiques (définition, exercice, simulation, etc.).

```
<owl:Class rdf:ID="LearningPedagogy">
    <rdfs:subClassOf rdf:resource="#ELearningClass"/>
</owl:Class>
<owl:Class rdf:ID="PedagogicalActivity">
    <rdfs:subClassOf rdf:resource="#LearningPedagogy"/>
</owl:Class>
```

Les propriétés *learningTopicProperty*, *LearnerProperty* et *pedagogicalProperty* sont des sous-propriétés de la propriété racine *eLearningProperty*. Ces trois propriétés permettent de décrire des propriétés relatives respectivement au domaine, à l'apprenant et aux activités pédagogiques.

```

<owl:ObjectProperty rdf:ID="learningTopicProperty">
  <rdfs:subPropertyOf rdf:resource="#eLearningProperty"/>
  <rdfs:domain rdf:resource="#LearningTopic"/>
  <rdfs:range rdf:resource="#LearningTopic"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="learnerProperty">
  <rdfs:subPropertyOf rdf:resource="#eLearningProperty"/>
  <rdfs:domain rdf:resource="#Learner"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="pedagogicalProperty">
  <rdfs:subPropertyOf rdf:resource="#eLearningProperty"/>
  <rdfs:domain rdf:resource="#PedagogicalActivity"/>
  <rdfs:range rdf:resource="#PedagogicalActivity"/>
</owl:ObjectProperty>

```

La propriété *prerequisiteOf* permet de relier une notion (de type *LearningTopic*) à son post-requis (de type *LearningTopic*). La propriété *prerequisiteOf* est une sous propriété de la propriété *orderRelationship*. Cette dernière définit des relations d'ordre entre les notions du domaine, c'est-à-dire des relations où les notions se succèdent séquentiellement selon une certaine logique (ordre temporel, ordre d'importance, etc.).

```

<owl:ObjectProperty rdf:ID="prerequisiteOf">
  <rdfs:subPropertyOf rdf:resource="#orderRelationship"/>
  <rdfs:domain rdf:resource="#LearningTopic"/>
  <rdfs:range rdf:resource="#LearningTopic"/>
</owl:ObjectProperty>

```

La propriété *preferredAuthor* relie le concept *Learner* au concept *Author* et exprime le fait qu'un apprenant préfère un auteur en particulier.

```

<owl:ObjectProperty rdf:ID="preferredAuthor">
  <rdfs:subPropertyOf rdf:resource="#learnerProperty"/>
  <rdfs:domain rdf:resource="#Learner"/>
  <rdfs:range rdf:resource="#Author"/>
</owl:ObjectProperty>

```

La propriété *sequencingStrategy* permet de relier deux activités pédagogiques séquentielles dans le cadre d'une stratégie pédagogique donnée.


```

<owl:ObjectProperty rdf:ID="sequencingStrategy">
    <rdfs:subPropertyOf rdf:resource="#pedagogicalProperty"/>
    <rdfs:domain rdf:resource="#LearningActivity"/>
    <rdfs:range rdf:resource="# LearningActivity"/>
</owl:ObjectProperty>

```

Une fois les différents concepts et propriétés définis dans l'ontologie OWL, ils sont instanciés pour construire les différents modèles du e-Learning: le modèle du domaine, le modèle apprenant et le modèle pédagogique (voir la section suivante).

3.2. Modèles du e-Learning

L'ontologie du e-Learning décrite ci-dessus est instanciée pour construire les ontologies de formation et les annotations conceptuelles. Les ontologies de formation sont le modèle du domaine, le modèle pédagogique et le modèle apprenant, les annotations conceptuelles sont les descriptions des ressources pédagogiques (voir chapitre 5 pour les annotations conceptuelles). Conceptuellement, les modèles sont spécifiés de manière identique comme suit :

LM = (LO, I, InstC, InstP), où:

- LO=(C, P, H^c, H^p, Signature, Rules) représente l'ontologie du e-Learning.
- I est un ensemble d'instances (C, P et I sont des ensembles disjoints).
- InstC: C → 2^I est une fonction qui permet d'associer des instances aux concepts de LO.
- InstP: P → 2^{I × I} est une fonction qui permet d'associer des instances aux propriétés de LO.

Il ressort de cette définition formelle que le lien conceptuel qui relie les différents modèles à l'ontologie du e-Learning LO est le lien d'instanciation, c'est-à-dire, les concepts et les propriétés de LO sont instanciés pour construire les différents modèles du e-Learning.

3.2.1. Modèle du domaine

Ce modèle est une ontologie de formation. Il permet de décrire les notions (en anglais topics) qui constituent le savoir relatif à un domaine ou à une discipline. Un domaine particulier (par exemple le domaine de l'algorithmique) est décrit par un ensemble de notions et un ensemble de relations qui relient ces notions. La figure 9 représente une partie du domaine de l'algorithmique décrit par un ensemble de notions reliées par deux types de relation : la relation *prerequisiteOf* (pré-requis de) et la relation *subTypeOf* (sous-type de). Ces deux relations sont définies comme des propriétés dans l'ontologie du e-Learning (voir figure 8) avec comme domaine et co-domaine le concept *LearningTopic*. Les relations *prerequisiteOf*

et *subTypeOf* ne sont pas réflexives (cette règle de non-réflexivité est définie dans Rules de LO).

Dans notre travail de thèse, nous avons construit deux modèles de domaine :

- Une ontologie pour représenter les notions du domaine de l'« algorithmique et des langages de programmation »
- Une autre ontologie pour représenter les notions du domaine de l'« algèbre ».

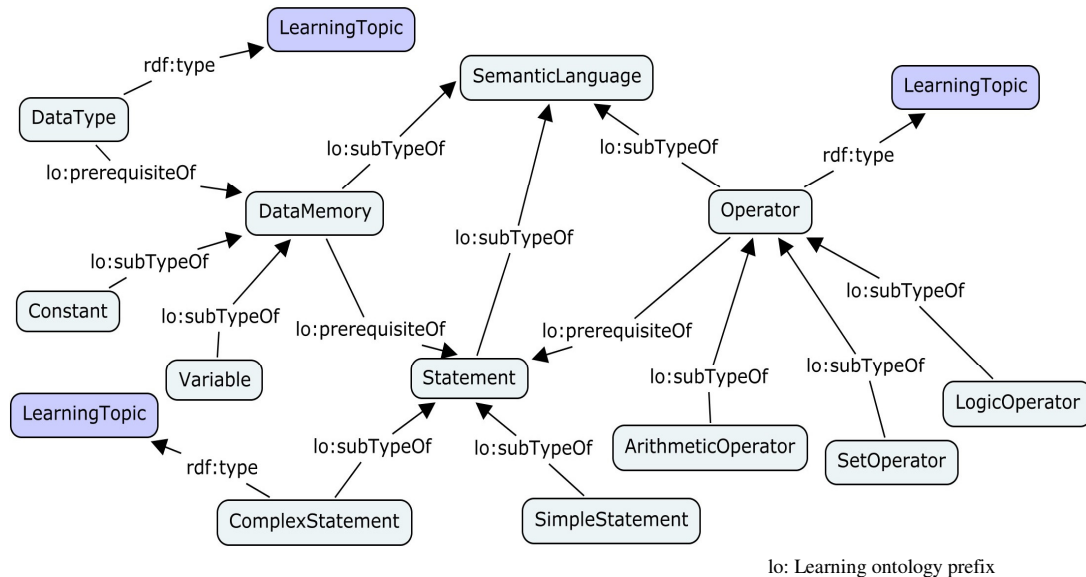


Figure 9 : Un extrait du modèle du domaine de l'algorithmique

3.2.2. Modèle pédagogique

A l'instar des systèmes tuteurs intelligents, nous utilisons un modèle pédagogique pour décrire comment se fait l'enchaînement des différentes activités pédagogiques (exercice, explication, définition, etc.). Ce modèle permet de décider quel type de ressources pédagogiques présenter à l'apprenant en tenant compte du profil de celui-ci. Par exemple, certains apprenants préfèrent, pour appréhender une notion particulière, partir d'un exemple concret sur la notion et ensuite aller vers une conceptualisation de la notion (stratégie inductive).

Ce modèle représente une ontologie de formation qui décrit différentes stratégies pédagogiques pour appréhender une notion du domaine. Concrètement, nous construisons cette ontologie en définissant les différentes activités pédagogiques (exercice, définition, exemple, etc.) et leur enchaînement dans le temps. Certaines activités pédagogiques sont séquentielles et d'autres peuvent se dérouler en parallèle (elles peuvent être présentées à

l'apprenant en parallèle). Dans la figure 10, nous représentons différentes activités pédagogiques reliées par deux types de relations *alternativeStrategy* et *sequencingStrategy*, tous deux définis en tant que propriétés dans LO (sous-propriétés de *pedagogicalProperty*). Les activités pédagogiques sont définies comme des instances du concept *PedagogicalActivity* définis dans LO.

Nous rappelons que l'approche pédagogique adoptée dans notre travail est l'approche constructiviste qui s'adapte bien à notre problématique de recherche : offrir un environnement pédagogique centré sur l'apprenant et qui permet à celui-ci de construire ses connaissances sans l'influence d'un tuteur/enseignant mais par la découverte et la responsabilisation. Dans le système OrPAF, le modèle pédagogique n'est pas utilisé pour modéliser plusieurs théories d'apprentissage (autres que le constructivisme) mais plutôt permet de modéliser différentes stratégies d'apprentissage à l'intérieur de la théorie constructiviste.

Le modèle pédagogique de la figure 10 indique, par exemple, que la définition d'une notion peut être indifféremment une définition formelle (*FormalDefinition*) ou une définition informelle (*UnformalDefinition*) et que l'explication d'une notion (*Explanation*) précède l'exercice sur la notion (*Exercise*).

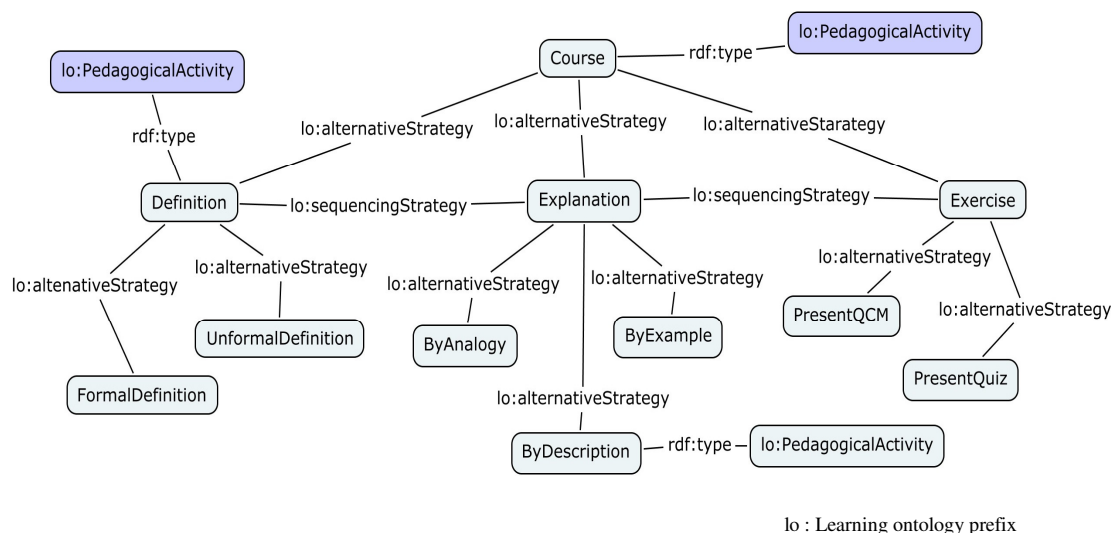


Figure 10 : Un extrait du modèle pédagogique

3.2.3. Modèle apprenant

Le modèle apprenant représente les connaissances que possède le système sur l'apprenant. Ces connaissances ne sont pas fiables et sont remises en cause pendant le processus d'apprentissage. Dans notre travail, nous avons modélisé deux types de connaissances sur

l'apprenant : d'une part sa maîtrise des différentes notions du domaine cible et d'autres part ses préférences (son auteur préféré, son format de ressource préféré, son type cognitif, etc.).

Pour représenter la connaissance qu'a l'apprenant sur les différentes notions du domaine, nous avons opté pour le modèle de recouvrement binaire (Overlay model) (Galeev et al., 2004). Ce modèle est une estimation du niveau de connaissance de l'apprenant pour chaque notion du domaine. Dans notre cas, cette connaissance est binaire : soit l'apprenant connaît la notion soit il ne la connaît pas. Cette connaissance est acquise à travers les différents tests qu'exécute l'apprenant pendant l'apprentissage.

Les préférences de l'apprenant sont fournies par lui au début de l'utilisation du système et sont représentées dans son modèle apprenant. Les préférences de l'apprenant servent principalement à sélectionner les ressources pédagogiques qui s'adaptent son profil.

Dans notre travail, le modèle apprenant est construit en instanciant des concepts appartenant à l'ontologie du e-Learning LO, au modèle du domaine ou au modèle pédagogique. Par exemple, dans la figure 11, l'apprenant *Maria* est une instance du concept *Learner* défini dans LO et elle est reliée par la propriété *masteredLearningTopic* (définie dans LO) à la notion *Operator* (définie dans le modèle du domaine), instance de la classe *LearningTopic* (définie dans LO). La propriété *masteredLearningTopic* décrit le fait que l'apprenant *Maria* maîtrise la notion *Operator*.

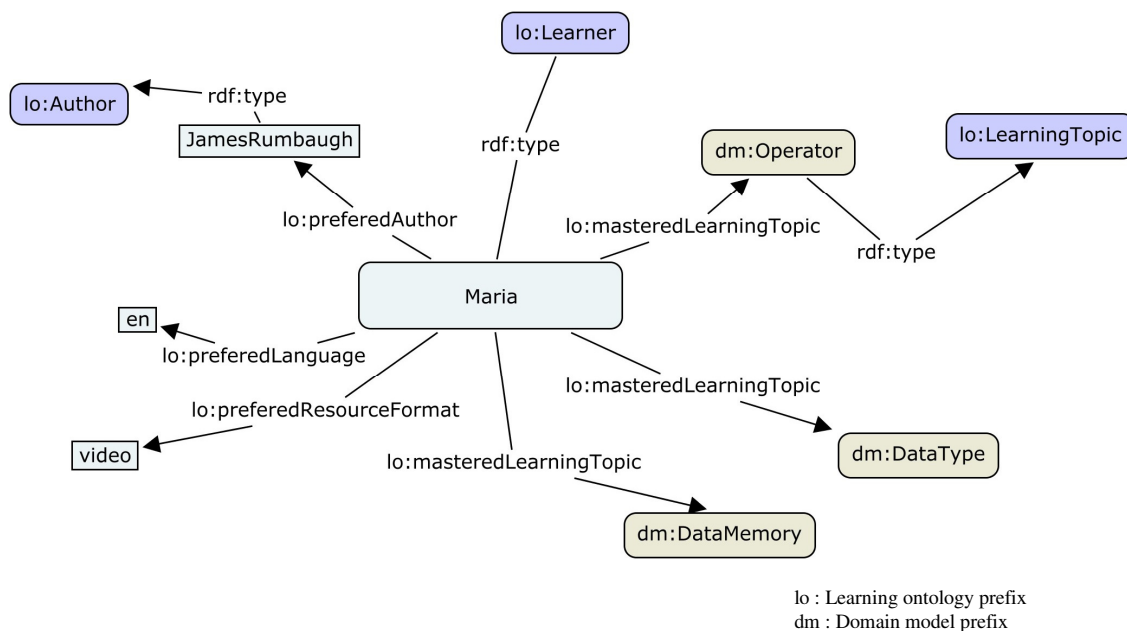


Figure 11 : Un extrait du modèle apprenant

3.2.4. Langage de description des différents modèles

La description des différents modèles du e-Learning (le modèle du domaine, le modèle pédagogique et le modèle apprenant) est faite dans le langage RDF. Le choix de RDF est justifié par le constat suivant : dans le système OrPAF, les différents modèles utilisent les concepts et les propriétés de l'ontologie du e-Learning pour décrire des connaissances concrètes. De manière similaire, les annotations RDF utilisent les concepts et les propriétés d'un modèle OWL pour la description de ressources. En effet, il existe un parallèle évident entre l'ontologie du e-Learning et les modèles du e-Learning d'une part et le modèle OWL et ses annotations RDF d'autre part.

La figure 12 synthétise le lien qui relie l'ontologie du e-Learning aux différents modèles. Il y apparaît clairement que les annotations RDF sont l'instanciation de connaissances décrites au niveau du modèle OWL.

Par exemple le triplet $\langle \textit{Maria}, \textit{masteredLearningTopic}, \textit{Transitivity} \rangle$ est l'instanciation de la connaissance suivante décrite en OWL : $\textit{masteredLearningTopic}(\textit{Learner}, \textit{LearningTopic})$ (en français, un apprenant maîtrise une notion). Cette annotation RDF appartient au modèle apprenant qui sert à modéliser les caractéristiques de l'apprenant *Maria*.

Un autre exemple, le triplet RDF $\langle \textit{EquivalenceRelation}, \textit{hasPrerequisite}, \textit{Transitivity} \rangle$ est l'instance de la connaissance générique « une notion a comme prérequis une autre notion », décrite en OWL comme suit : $\textit{hasPrerequisite}(\textit{LearningTopic}, \textit{LearningTopic})$.

Notons que $\textit{masteredLearningTopic}$ et $\textit{hasPrerequisite}$ sont des propriétés dans *P* de LO.

Modèle OWL

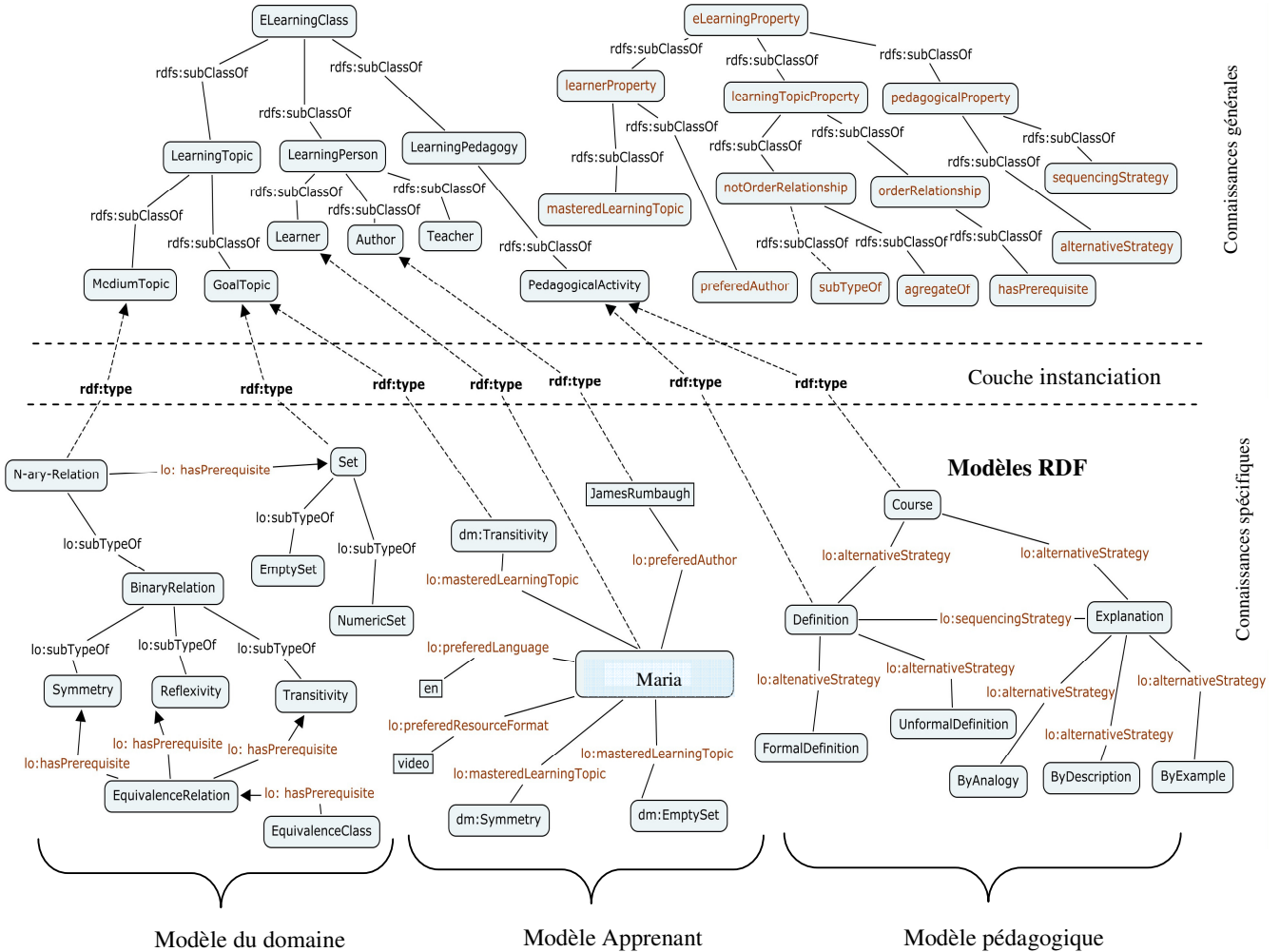


Figure 12: Extrait de l'intégration des connaissances ontologiques

3.2.5. Exemples de sérialisation des modèles du e-Learning par des annotations RDF

Annotations RDF du modèle du domaine:

Dans le domaine de l'algorithmique, les experts définissent la notion *Statement* comme une instance du concept *MediumTopic*, une généralisation de la notion *Statement* et un postrequis de la notion *Operator*:

```

<MediumTopic rdf:ID="ArithmeticStatement">
  <lo:subTypeOf>
    <GoalTopic rdf:ID="Statement">
    </GoalTopic>
  </lo:subTypeOf>
  <lo:hasprerequisite>
    <MediumTopic rdf:ID="Operator"/>
  </lo:hasprerequisite>
</MediumTopic>

```

Annotations RDF du modèle pédagogique:

Dans le modèle pédagogique qui décrit la stratégie inductive, l'expert en pédagogie spécifie qu'il existe deux types de définitions, une définition formelle et une autre informelle et c'est seulement une fois que la définition d'une notion est présentée à l'apprenant que celui-ci peut accéder à une explication de la notion :

```

<PedagogicalActivity rdf:ID="Definition">
  <lo:alternativeStrategy>
    < PedagogicalActivity rdf:ID="UnformalDefinition"/>
  </lo:alternativeStrategy>
  <lo:alternativeStrategy>
    < PedagogicalActivity rdf:ID="FormalDefinition"/>
  </lo:alternativeStrategy>
  <lo:sequencingStrategy>
    <PedagogicalActivity rdf:ID="Explanation"/>
  <lo:sequencingStrategy />
<PedagogicalActivity />

```

Annotations RDF du modèle apprenant:

Dans le modèle de l'apprenant *Maria*, nous représentons, par exemple, le fait que *Maria* maîtrise la notion cible *Statement* et que *Maria* préfère l'auteur *RoseDieng* aux autres auteurs comme suit :

```

<Learner rdf:ID="Maria">
  <lo:masteredLearningTopic>
    <GoalTopic rdf:resource="#Statement"/>
  </lo:masteredLearningTopic>
  <lo:preferredAuthor>
    <Author rdf:ID="RoseDieng"/>
  <lo:preferredAuthor/>
</Learner/>

```

4. Développement des ontologies : Environnement KAON (Maedche et al, 2000)

De nombreuses recherches se sont penchées sur le développement, la construction et l'évolution d'ontologies. La difficulté, les coûts élevés en temps, en effort et en argent du développement d'ontologies freinent énormément le développement des applications basées sur celles-ci.

La plupart des recherches dans le domaine du développement d'ontologies s'orientent vers la construction d'ontologies à partir de texte. En effet, la construction d'ontologies à partir de ressources textuelles pertinentes est une approche viable et de nombreux systèmes basés sur cette approche sont apparus : par exemple, les systèmes ASIUM (Faure et al., 1998), TextToOnto (Maedche et Staab, 2000), Ontolearn (Navigli et al., 2003), OntoLT (Buitelaar et al., 2004) ou Kaso (Yimin et al., 2006). Tous ces systèmes combinent une analyse linguistique des textes et des algorithmes d'apprentissage automatique pour extraire à partir du texte des concepts candidats ou potentiels et des relations entre ces concepts.

Dans notre recherche, nous adoptons différentes approches pour construire les ontologies du système OrPAF: une approche semi-automatique pour la construction du modèle du domaine et une approche manuelle pour la construction des autres ontologies.

Le choix de semi-automatiser la construction de l'ontologie du domaine s'explique par la taille importante des ontologies de domaines qui nécessitent le recours à des outils automatiques pour l'acquisition de notions candidates à partir de documents de référence du domaine en question (par exemple l'algèbre).

Dans notre travail, les ingénieurs de la connaissance et les experts ont utilisé pour la construction du modèle du domaine, le système TextToOnto (Maedche et al, 2000). Le système TextToOnto fait partis de l'environnement de construction d'ontologies KAON (KArllsruhe ONtology) qui associe différents modules, correspondant à autant de phases identifiées de l'ingénierie ontologique : création, stockage, raffinage, exploitation et maintenance. Le module TextToOnto applique des stratégies de fouille de texte (Text mining) à des corpus textuels pour la création semi-automatique d'ontologies. Il nous permet de faire de l'extraction de termes (notions candidates et relations sémantiques entre les notions) à partir de ressources pédagogiques de différents formats (.pdf, .xml, .txt, .html). Ces termes sont ajoutés à l'ontologie du domaine sans ou avec modification (une modification classique est la lemmatisation du terme). En outre, l'environnement KAON nous permet d'économiser du temps et de l'effort dans la construction et la maintenance des différentes ontologies.

Cependant, les concepts et les relations proposées ou extraites doivent nécessairement être validés par les experts (experts du domaine, expert de la pédagogie, etc.) avant d'être ajoutées aux ontologies.

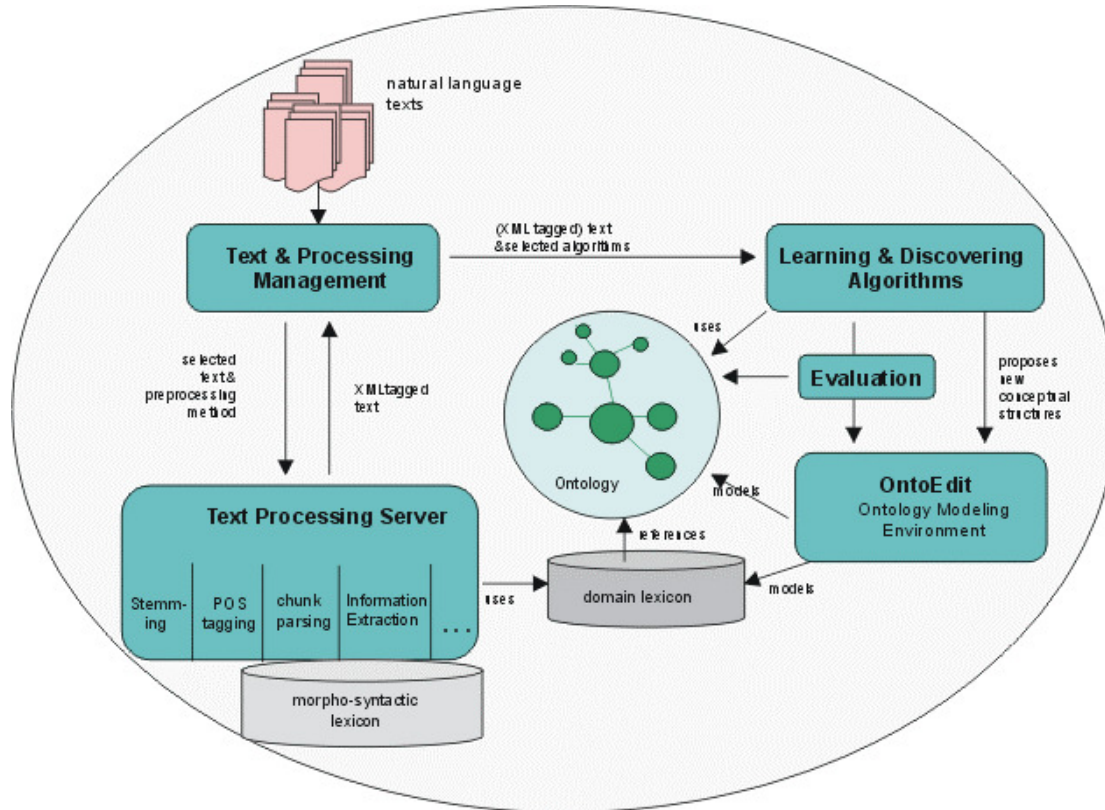


Figure 13: Architecture de l'environnement TextToOnto

Le processus d'ingénierie que nous avons suivi pour la construction du modèle du domaine est le suivant :

- 1- collecter les documents pertinents du domaine cible (dans notre cas, nous avons modélisé deux domaines : l'Algorithmique et l'Algèbre)
- 2- Extraire automatiquement des documents les termes et les relations qui représentent des candidats potentiels pour appartenir à l'ontologie du domaine : (i) un analyseur (parser) linguistique est utilisé par TextToOnto pour séparer/identifier les différentes formes grammaticales du document, (ii) le document est annoté par une analyse linguistique et transformé en un fichier XML (iii) Un module extracteur est ensuite utilisé pour extraire les notions et les relations du document XML.

3- Décider quel terme candidat est retenu comme notion dans l'ontologie du domaine, éventuellement après une transformation morphologique (par exemple le lemme ou la forme canonique du terme candidat).

4- Identifier les relations entre les notions du domaine. Il s'agit dans cette étape de vérifier si la relation extraite du document correspond sémantiquement à une des relations décrites dans l'ontologie du e-Learning (sous-propriété de la propriété *learningTopicProperty*), sinon considérer cette relation comme une nouvelle sous-propriété de la propriété *learningTopicProperty* et l'ajouter dans l'ontologie du e-Learning.

5- Valider l'ontologie finale du domaine.

Les étapes 3, 4 et 5 sont réalisées par des experts du domaine en collaboration avec les ingénieurs de la connaissance.

4.1. Acquisition des connaissances du domaine avec TextToOnto

Le module d'acquisition des connaissances du système TextToOnto est un processeur textuel composé d'un analyseur lexical, d'un analyseur syntaxique et d'un module d'apprentissage automatique d'ontologie (voir figure 14).

L'analyseur lexical prend en charge l'éclatement du texte en tokens (termes atomiques) et l'analyse lexicale de ces tokens (analyse morphologique et mise en forme canonique des tokens).

L'analyseur syntaxique travaille au niveau des phrases du texte pour déterminer les fonctions grammaticales des différentes parties de la phrase (tels que le sujet, le verbe, le complément d'objet direct, etc.).

Les termes potentiels sont extraits de manière statistique (ex. occurrence du terme) ou au moyen d'expressions régulières définies par l'expert. Les expressions régulières peuvent porter sur un étiquetage morphosyntaxique du corpus. L'extraction de relation se fait sur la base de patrons lexico-syntaxiques, d'un calcul de proximité ou des deux (selon le choix de l'expert).

L'apprentissage d'ontologie repose sur des modules qui intègrent différents algorithmes pour extraire les notions les plus pertinentes, la taxonomie des notions et leurs relations à partir du document annoté XML. Ce module permet également de déduire une ontologie propre à un domaine à partir d'une structure terminologique générique, comme WordNet.

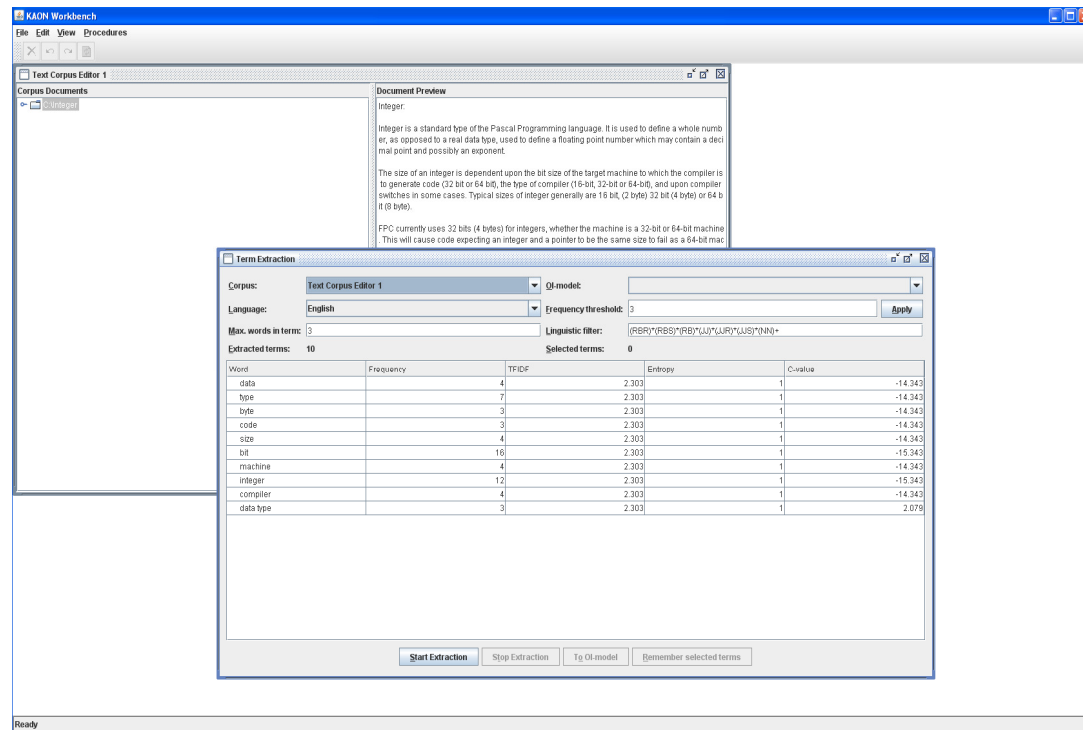


Figure 14 : Extraction de termes candidats dans KAON à partir d'un document textuel

4.2. Edition des ontologies avec OntoEdit

KAON est associé à l'éditeur d'ontologie OntoEDIT, dans lequel il est possible de poursuivre la construction d'une ontologie amorcée avec le module TextToOnto et son exportation dans des fichiers OWL/RDF après sa validation par les experts (voir figure 15).

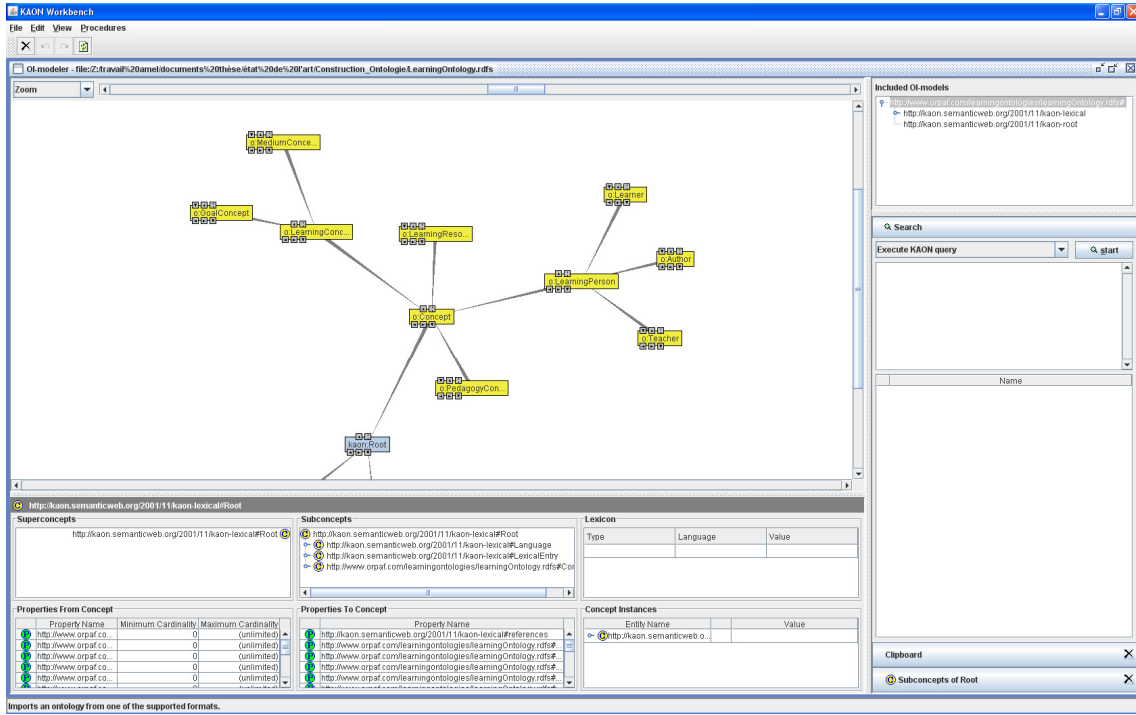


Figure 15 : Construction de l'ontologie du e-Learning avec OntoEdit

Conclusion : Dans ce chapitre, nous nous sommes intéressés à l'aspect représentation des connaissances dans l'environnement pédagogique OrPAF et nous avons montré l'utilisation du formalisme ontologique pour la description de ces connaissances. Notre démarche d'ingénierie combine l'ingénierie éducative réalisée par des experts du domaine et des experts de la pédagogie et l'ingénierie des connaissances réalisée par des ingénieurs de la connaissance. D'une part, les experts du domaine définissent de manière informelle les notions du domaine cible et les experts de la pédagogie définissent les stratégies pédagogiques à appliquer en terme d'activités pédagogiques. D'autre part, les ingénieurs de la connaissance formalisent ces connaissances par une représentation ontologique à deux niveaux : une ontologie du e-learning et les ontologies de formation, à savoir, le modèle du domaine, le modèle apprenant et le modèle pédagogique.

Nous avons choisi le couple (OWL, RDF) pour la représentation ontologique du couple (ontologie du e-learning, modèle du e-learning). Ce choix est en conformité avec notre choix conceptuel qui décrit les modèles du e-learning comme une instantiation de l'ontologie du e-learning pour une formation particulière.

Nous avons adopté l'outil TextToOnto pour la construction des ontologies et particulièrement pour la construction du modèle du domaine à partir de documents textuels. En effet, cet outil nous permet d'extraire des notions du domaine et leurs relations à partir de documents de référence dans le domaine (des documents considérés comme pertinents par les experts du domaine). Cependant, les connaissances extraites sont d'abord traitées par les experts avant d'être ajoutées au modèle du domaine : la construction du modèle du domaine est semi-automatique.

5. Génération de cours hypermédias adaptatifs dans OrPAF

Actuellement, les efforts de développement des environnements pédagogiques en ligne tendent vers la mise en œuvre de supports qui fournissent des fonctionnalités telles que la personnalisation, l'adaptation et la génération automatique de cours.

Dans le chapitre précédent, nous avons montré que face au besoin de personnalisation et d'adaptation des environnements pédagogiques, les ontologies peuvent constituer une réponse viable et satisfaisante car elles permettent de décrire de manière explicite, partagée et formelle la conceptualisation d'un domaine. L'aspect consensuel et formel de la description permet la réutilisation, l'adaptation et la personnalisation des objets pédagogiques (services, modèles, ressources, etc.) pour différents contextes de formation. L'opérationnalisation des ontologies permet la génération automatique des cours adaptés aux apprenants et aux différents contextes de formation. Dans ce chapitre, nous présentons notre approche qui permet de générer automatiquement des cours hypermédias adaptatifs en exploitant les ontologies construites dans le système OrPAF. La tâche de génération des cours hypermédias adaptatifs est réalisée

par le composant `AdaptationCourse` dans lequel nous avons effectué un certain nombre de choix conceptuels (Yessad et al., 2008a ; Yessad et al., 2008b ; Yessad et al., 2008c) :

D'une part, nous réutilisons deux techniques empruntées des systèmes hypermédias adaptatifs (voir chapitre 3), à savoir, les cartes conceptuelles (concept map) et l'annotation des hyperliens. Ces deux techniques permettent de mettre en œuvre l'adaptativité de la navigation dans le cours hypermédia et ont pour objectif de réduire la désorientation et la surcharge cognitive des apprenants (Brusilovsky, 1996).

D'autre part, pour générer du contenu adaptatif, nous proposons une approche qui permet de rechercher, d'annoter et de réutiliser des ressources pédagogiques à partir de corpus du Web.

Le cours généré dans le système OrPAF est un hypermédia adaptatif où :

- les hyperliens sont organisés sous la forme d'une carte conceptuelle qui représente la structure navigationnelle de l'hypermédia. L'exploration de la carte conceptuelle permet à l'apprenant d'acquérir des connaissances sur le domaine.
- les pages sont des ressources pédagogiques recherchées à partir du Web (page HTML ou autres). Ces ressources pédagogiques subissent en amont un travail d'annotation semi-automatique avant d'être réutilisées automatiquement pendant le processus d'apprentissage.

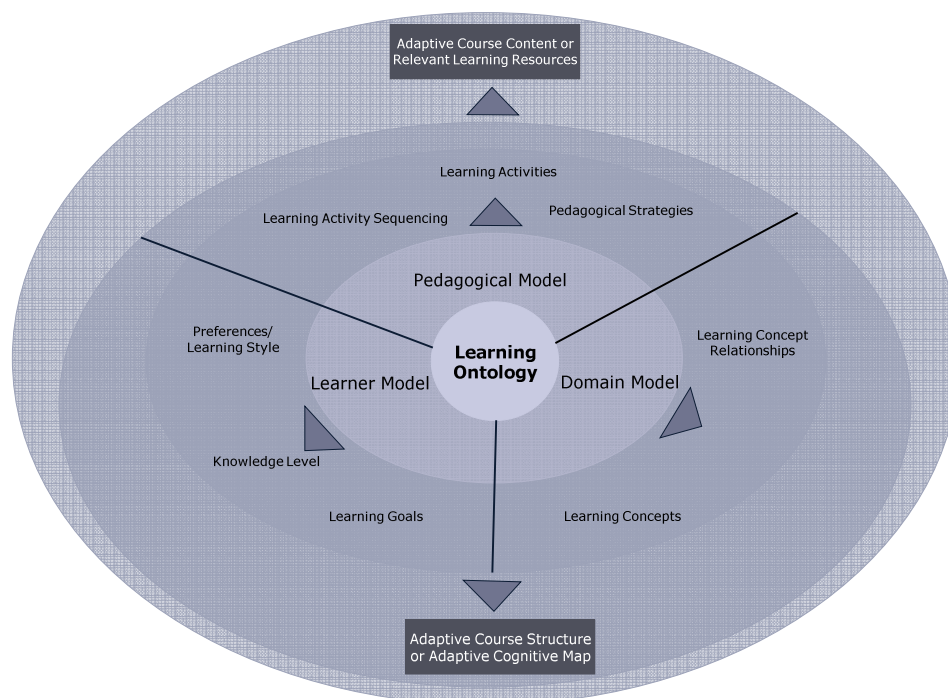


Figure 16 : Moteur d'Adaptation du cours dans OrPAF

1. Structure adaptative du cours hypermédia

1.1. Définition de la carte conceptuelle

Une carte conceptuelle (CM pour Conceptual Map) est un graphe qui organise la connaissance et qui présente les notions d'un domaine ainsi que les relations qui les relient. Une carte conceptuelle peut servir comme support de présentation de cours (Horton et al., 1993). Dans notre approche, une carte conceptuelle représente la structure navigationnelle du cours hypermédia généré. Les nœuds de la carte conceptuelle sont des notions définies dans le modèle du domaine et les liens entre les nœuds de la carte sont des relations qui relient les notions dans le modèle du domaine. Dans une carte conceptuelle, au-delà du graphe de notions (la partie visible de l'iceberg), nous retrouvons d'autres informations : des ressources pédagogiques pertinentes sont associées aux notions de la carte. L'apprenant en parcourant la carte conceptuelle peut visualiser des contenus pédagogiques. Etant donné que la carte conceptuelle a une structure graphique, l'apprenant dispose de plusieurs chemins et parcours d'apprentissage et il est libre d'explorer n'importe quel parcours de la CM. Cette situation est à l'origine de deux problèmes : la désorientation et la surcharge cognitive de l'apprenant.

Nous répondons à ces deux problèmes par :

- La CM ne contient que les notions qui permettent à l'apprenant d'atteindre son l'objectif.
- Les notions de la CM sont annotées selon le principe suivant : une notion de la carte est accessible à l'apprenant si toutes ses notions pré-requis sont déjà connues (maîtrisées) par celui-ci. Ce principe assure que l'apprenant n'aborde que les notions qu'il peut appréhender en vu de son état de connaissances.

1.2. Construction de la carte conceptuelle

La carte conceptuelle est un graphe de notions appartenant à un domaine. Dans notre approche, nous construisons une carte conceptuelle représentant la structure du cours adaptatif par filtrage du modèle du domaine construit dans le système OrPAF. Le filtre utilisé dépend des caractéristiques suivantes de l'apprenant :

- l'objectif de l'apprenant qui est une notion pédagogique particulière, instance du concept *GoalTopic* (défini dans LO), cette information est stockée dans le modèle de l'apprenant.
- les contraintes temporelles de l'apprenant qui sont une évaluation qualitative de la quantité de temps dont dispose l'apprenant pour atteindre son objectif. Nous avons défini trois valeurs pour les contraintes temporelles: « l'apprenant a le temps », « l'apprenant a moyennement le temps » et « l'apprenant n'a pas le temps ».

- l'état des connaissances de l'apprenant est défini par la propriété *masteredLearningTopic* et représenté dans le modèle de l'apprenant. Cette propriété est utilisée pour représenter les notions du domaine maîtrisées par l'apprenant.

Etant donné que nous avons défini trois contraintes temporelles pour un apprenant, nous modélisons trois filtres qui permettent de construire trois types de cartes conceptuelles : les cartes simples, les cartes hiérarchiques et les cartes relationnelles. Nous sommes partis pour la définition des filtres d'une hypothèse assez forte mais qui est vérifiée dans de nombreux cas : « *la taille de la carte conceptuelle en terme de notions et de relations détermine souvent le temps nécessaire à l'apprenant pour atteindre son objectif. Plus la taille de la carte est grande plus le temps d'apprentissage est important.* ».

Cependant, dans certains cas, le temps que passe l'apprenant à apprendre les notions de la carte conceptuelle n'a pas de lien direct avec la taille de la carte conceptuelle mais avec la complexité des notions. Pour réduire la probabilité d'apparition de ces cas, les experts du domaine veillent lors de la construction du modèle du domaine, à définir des notions de granularités similaires. Ceci nous permet de considérer, dans une certaine limite, que le temps passé à apprendre une notion est plus ou moins le même.

1.3. Filtres du modèle du domaine

Nous définissons trois filtres pour la construction de trois types de cartes conceptuelles (CM) : la carte simple $M_s(G)$, la carte hiérarchique $M_h(G)$ et la carte relationnelle $M_{r,m}(G)$. Ces filtres dépendent des contraintes temporelles de l'apprenant : la carte simple $M_s(G)$ s'adapte aux apprenants avec des contraintes sévères ou critiques, la carte hiérarchique $M_h(G)$ s'adapte aux apprenants avec des contraintes temporelles moyennes et la carte relationnelle $M_{r,m}(G)$ s'adapte aux apprenants avec des contraintes temporelles souples.

En fonction des contraintes temporelles et de l'objectif de l'apprenant, le composant *AdaptationCourse* du système OrPAF applique sur le modèle du domaine le filtre adéquat. Seules les notions et les relations qui passent le filtre sont présentées à l'apprenant. Etant donnée une notion cible G (l'objectif de l'apprenant), les filtres sont définis sur le modèle du domaine DM et ses notions qui appartiennent à l'ensemble $InstC$ comme suit :

- $M_s(G)$ est la CM qui contient le moins de notions et de relations. Elle est composée de la notion cible G et de toutes les notions reliées à elle par une relation d'ordre directe ou par fermeture transitive (la propriété *orderRelationship* et ses sous-propriétés dans l'ontologie du e-Learning). Un cas fréquent de relations d'ordre est la relation « prérequis de » qui relie les notions d'un domaine (*prerequisiteOf* dans l'ontologie du e-Learning) et qui signifie qu'une

notion est le prérequis d'une autre, c'est-à-dire, qu'elle doit être présentée à l'apprenant en premier.

$$M_s(G) = \{G\} \cup \{t \in \text{InstC}(\text{DM}) / \exists p \in \text{orderRelationship } p^t(t, G)\}$$

- $M_h(G)$ est la CM obtenue en appliquant le filtre simple à toutes les notions descendantes et ascendantes de la notion cible G (y compris la notion G). Dans cette CM, nous ajoutons en plus des relations d'ordre, la relation de subsomption qui relie une notion à ses parents ou à ses filles. La relation de subsomption est définie par la propriété *subTypeOf* dans l'ontologie du e-Learning.

$$M_h(G) = M_s(F), \text{ where } F = \{G\} \cup \{c \in \text{InstC}(\text{DM}) / \text{subTypeOf}(G, c) \vee \text{subTypeOf}(c, G)\}$$

- $M_{r,m}(G)$ est la CM obtenue en appliquant le filtre simple à toutes les notions reliées à la notion cible G par un chemin de relations inférieur ou égal à m (y compris la notion G):

$$M_{r,m}(G) = M_s(F), \text{ where } F = \{G\} \cup \{t \in \text{InstC}(\text{DM}) / \forall p \in \text{learningTopicProperty } (p^t(G, t) \vee p^t(t, G)) \wedge \text{length}(G, t) \leq m\}$$

Nous reprenons l'exemple de la figure 9 (voir chapitre 4) qui représente une partie du modèle du domaine de l'algorithmique et nous supposons que la notion cible de l'apprenant est la notion *Statemen*. En fonction du filtre appliqué, nous obtenons les cartes conceptuelles de la figure 17 (a), (b) ou (c).

Nous notons qu'à chaque fois qu'une notion est ajoutée à la CM, les notions reliées à cette notion par une relation d'ordre (par exemple la relation *prerequisiteOf*) sont également ajoutées à la CM. Cette approche est purement pédagogique car elle permet à l'apprenant de parcourir les notions prérequis d'une notion avant d'aborder celle-ci même si elles sont déjà maîtrisées par l'apprenant. Nous pouvons également constater que l'apprentissage dans une CM simple est plus directif que celui d'une CM hiérarchique ou d'une CM relationnelle. L'apprentissage dans cette dernière est beaucoup plus libre puisque le nombre de parcours augmente rapidement avec le facteur m .

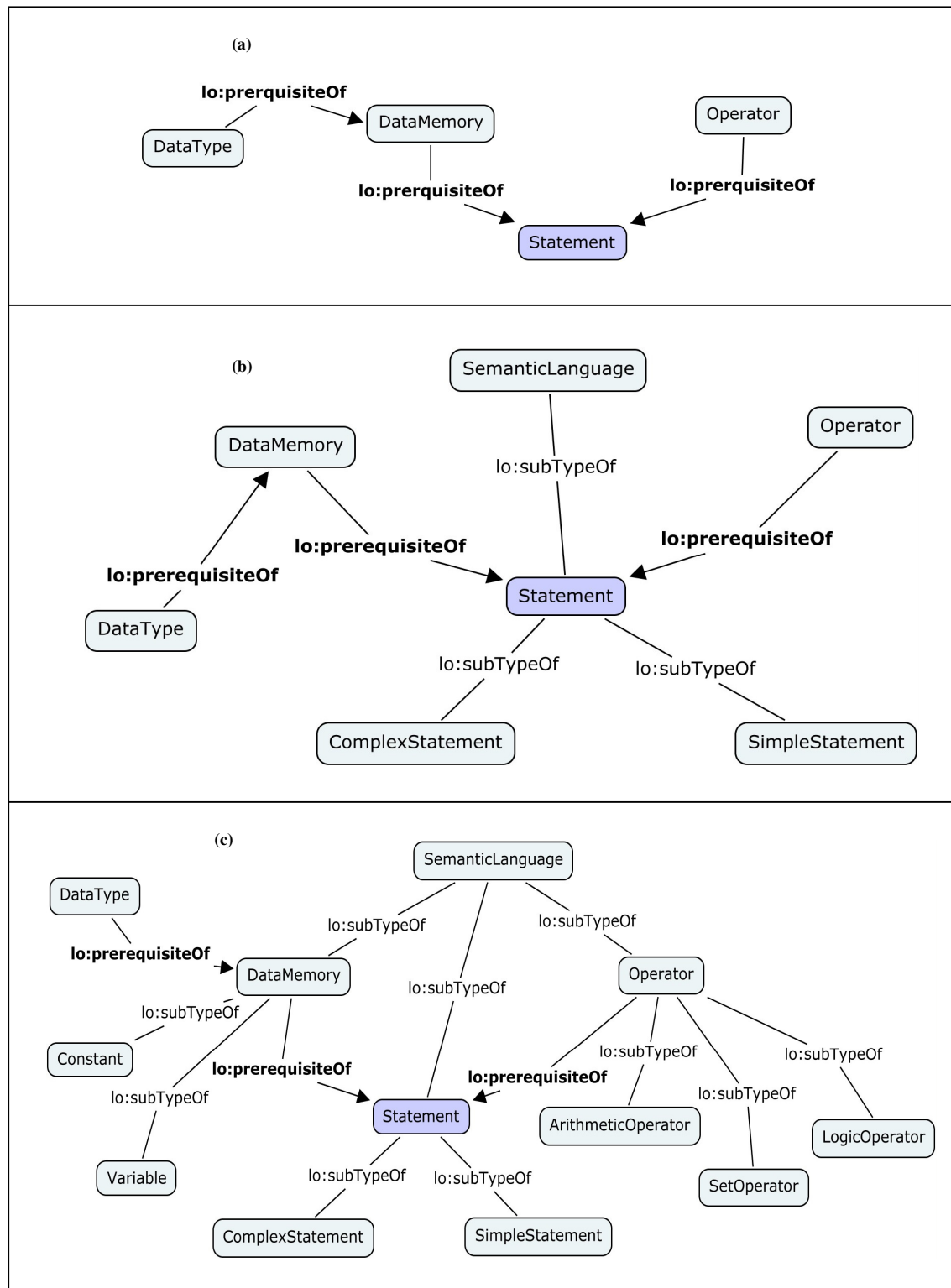


Figure 17 : Filtrés du modèle du domaine : (a) une CM simple, (b) une CM hiérarchique et (c) une CM relationnel (m=2)

1.4. Présentation adaptative de la carte conceptuelle

Avant d'afficher à l'apprenant la CM calculée, nous appliquons à celle-ci des techniques d'adaptation issues des recherches sur les hypermédias adaptatifs (De Bra et al., 1999). Nous utilisons particulièrement la technique d'annotation des hyperliens d'un hypermédia. Dans notre approche, nous annotons les notions de la CM en utilisant le modèle apprenant et nous obtenons une carte conceptuelle adaptative ou ACM (Adaptive Conceptual Map). En cet effet, nous distinguons entre trois états différents pour une notion de l'ACM : *AccessibleAndMastered*, *AccessibleAndNotMastered* et *NotAccessible*.

Une notion est accessible si l'apprenant peut cliquer dessus pour consulter des ressources pédagogiques pertinentes pour son niveau de connaissances et son objectif d'apprentissage. Sur le plan des connaissances, pour qu'une notion soit accessible à l'apprenant, il faut que tous ses prérequis soient maîtrisés par l'apprenant. Une notion accessible peut être maîtrisée (en anglais *mastered*) ou pas par l'apprenant.

Nous pouvons définir les trois états d'un nœud de l'ACM comme suit :

AccessibleAndMastered qui signifie que la notion est accessible à l'apprentissage et qu'elle est maîtrisée par l'apprenant. *Mastered* signifie qu'il existe dans le modèle de l'apprenant un triplet RDF qui relie l'apprenant à cette notion par la propriété *masteredLearningTopic*.

AccessibleAndNotMastered qui signifie que la notion est accessible mais qu'elle n'est pas encore maîtrisée par l'apprenant.

NotAccessible qui signifie que la notion est inaccessible pour l'apprenant, c'est dire, qu'il ne peut pas l'appréhender en vue de son état de connaissances. Cela signifie qu'un ou plusieurs prérequis de la notion ne sont pas encore maîtrisés par l'apprenant.

Nous utilisons pour annoter les nœuds de la CM trois icônes pour représenter chacun des trois états. Dans la figure 18, par exemple, nous avons une ACM où la notion *Operator* est *AccessibleAndMastered*, la notion *Statement* est *AccessibleAndNotMastered* et la notion *DataType* est *NotAccessible*.

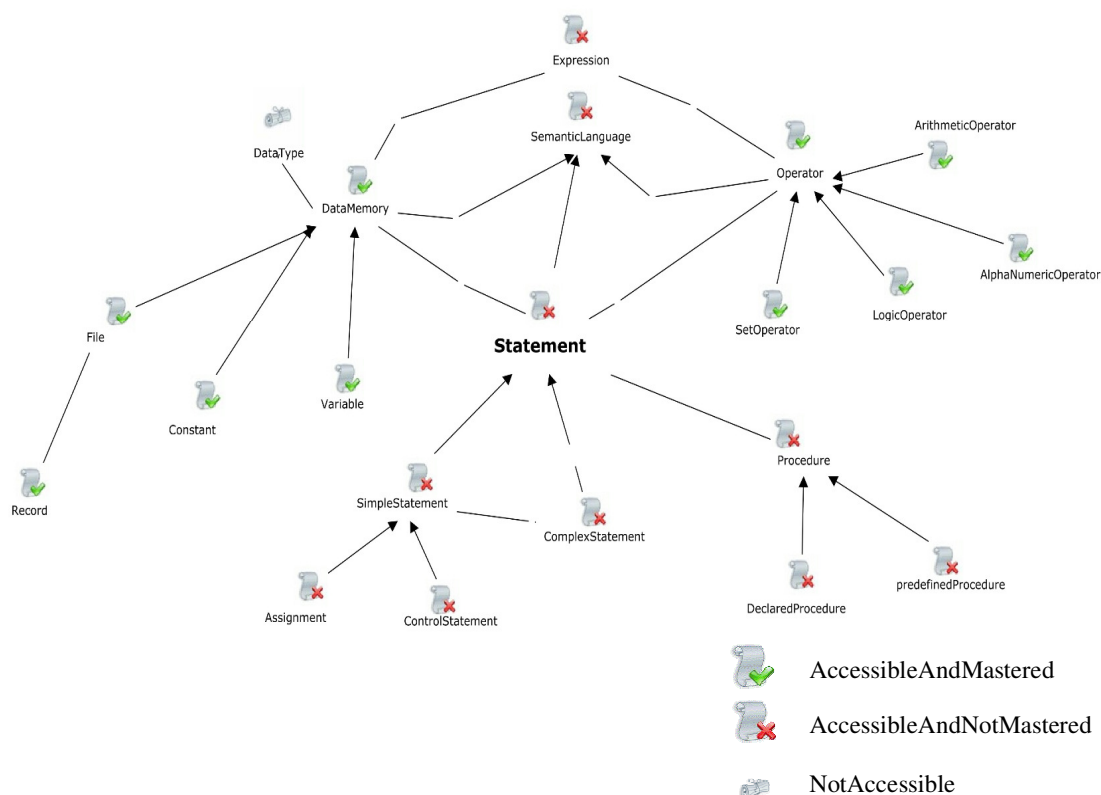


Figure 18 : une ACM du domaine de l'algorithmique

Les règles de changement d'état pour une notion de l'ACM sont modélisées dans le diagramme d'état de la figure 19. Le changement d'état d'une notion de l'ACM est conditionné par le résultat du test sur la notion en question, sur ses prérequis ou sur ses postrequis (les notions dont elle est prérequis). Cependant, la décision de passer un test appartient à l'apprenant car il est complètement responsable de son évolution. C'est l'essence même de l'approche constructiviste.

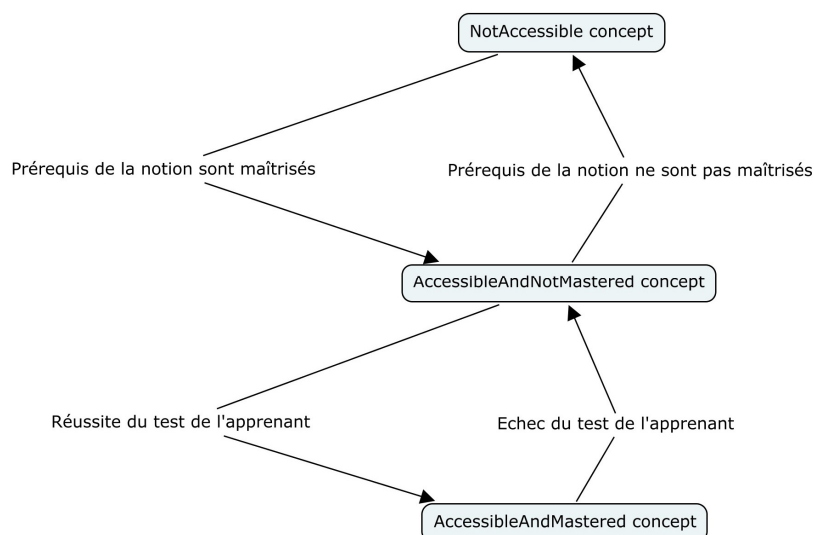


Figure 19 : Diagramme d'états représentant les changements d'état d'une notion de l'ACM

Le principe énoncé au début du chapitre « une notion de la carte conceptuelle n'est accessible à l'apprenant que si toutes ses notions pré-requis sont déjà connues (maîtrisées) par ce même apprenant » est vérifié car une notion devient accessible si et seulement si ses notions prérequis sont déjà maîtrisées (voir figure 19).

En résumé, nous considérons que l'ACM définit le **contexte de la formation** puisque elle représente l'état des connaissances de l'apprenant, sa contrainte temporelle et son objectif d'apprentissage. Dans la suite de ce mémoire, nous utilisons indifféremment les termes ACM et contexte de formation.

2. Contenu adaptatif du cours hypermédia

Dans notre travail de thèse, nous avons fait le choix de ne pas développer nous même les contenus pédagogiques mais de réutiliser des ressources pédagogiques existantes dans le Web. Particulièrement, nous nous sommes intéressés aux ressources pédagogiques annotées par le standard LOM (voir chapitre 3). Ainsi, nous proposons à nos enseignants un assistant qui leur permet de rechercher des ressources pédagogiques de corpus distants. Cette automatisation de la recherche n'est possible que si les métadonnées sont traitables par les programmes et donc les contenus sont automatiquement exploitables (machine-readable content). Les enseignants peuvent également enrichir le corpus local du système OrPAF en recherchant manuellement des ressources pédagogiques qu'ils jugent pertinentes pour le domaine en question.

Dans le système OrPAF, nous avons développé un composant `AdaptationContent` invoqué par le composant `AdaptationCourse` et qui réalise les tâches suivantes:

- la recherche distante de ressources pédagogiques (le composant `QueryingDistantResource`)
- l'annotation des ressources pédagogiques (le composant `AnnotationResource`)
- la recherche locale de ressources pédagogiques (le composant `QueryingDistantResource`)

2.1. Recherche distante des ressources pédagogiques

Il s'agit d'un assistant logiciel qui interroge des corpus distants du Web à la recherche de ressources pédagogiques qui font référence à des notions particulières du domaine, réalisées par des auteurs particuliers, éditées dans une langue particulière, etc. Tous ces critères de recherche sont représentés au niveau des descripteurs du standard LOM. Par exemple, le descripteur LOM « `General.Language` » nous renseigne sur la langue de la ressource pédagogique. Les requêtes distantes sont exécutées sur les métadonnées de la ressource et son contenu reste transparent pour le système OrPAF. Le fait de réduire la recherche à la partie annotée du Web est avantageux pour deux raisons :

- simplifie considérablement l'interaction avec les différents systèmes de gestion du contenu sur le Web. En effet, les métadonnées LOM sont des données structurées à l'inverse des contenus qui peuvent être de types différents (texte, image, vidéo, graphique, etc.) et très peu structurés.
- facilite la détection de la pertinence des ressources pédagogiques recherchées. Ceci n'est possible que si les auteurs des ressources externes annotent véritablement et suffisamment leurs ressources.

Dans notre travail de thèse, nous avons interrogé le corpus ARIADNE qui est lui-même un portail donnant accès à d'autres corpus du Web (Merlot, EducaNext, etc.).

Exemples de requêtes de recherche :

- 1- Rechercher les ressources pédagogiques LR (`Learning Resource`) qui traitent de la notion t (le prédicat `learningResourceTopic(LR, t)`) :

$$\forall LR \text{ (learningResource(LR) } \wedge \text{ learningResourceTopic(LR, t)) } \Rightarrow \text{adaptiveContent(LR)}$$

Cette règle signifie qu'une ressource LR est pertinente (le prédicat `adaptiveContent(LR)`) si elle traite de la notion t (le prédicat `learningResourceTopic(LR, t)`)

- 2- Rechercher les ressources pédagogiques LR qui sont des exercices sur la notion t (le prédicat (`learningResourceType(LR, Exercice)`)) :

$$\forall LR (\text{learningResource}(LR) \wedge \text{learningResourceTopic}(LR, t) \wedge \text{learningResourceType}(LR, \text{Exercise})) \Rightarrow \text{adaptiveContent}(LR)$$

Etant donné que le résultat de la recherche dépend de la correspondance exacte entre la requête et une ou plusieurs valeurs des descripteurs LOM de la ressource, il est possible d'avoir des silences de requêtes. Pour remédier à cette situation, nous avons proposé un mécanisme de relaxation de requêtes si celles-ci ne renvoient pas de résultats. Le mécanisme de relaxation est basé sur le modèle du domaine. Par exemple si l'enseignant lance une requête pour rechercher des ressources sur une notion t et que le résultat de la requête est nul alors nous utilisons les relations qui relient les notions du domaine pour relaxer la requête et éviter le silence.

Exemples de relaxation de requêtes :

Nous recherchons les ressources pédagogiques LR (Learning Resource) qui traitent de la notion t .

1- Relaxation de la requête de recherche en utilisant la propriété *subTypeOf* du modèle du domaine:

$$\forall LR \forall t (\text{learningResource}(LR) \wedge (\text{learningResourceTopic}(LR, t) \vee \exists t' (\text{subTypeOf}(t, t') \vee \text{subTypeOf}(t', t)) \wedge \text{learningResourceTopic}(LR, t'))) \Rightarrow \text{adaptiveContent}(LR)$$

Cette règle de relaxation signifie qu'une ressource est pertinente ($\text{adaptiveContent}(LR)$) si elle traite de la notion courante t ($\text{learningResourceTopic}(LR, t)$) ou bien si elle traite d'une notion descendante ou ascendante t' de la notion t .

2- Relaxation de la requête de recherche à toutes les notions qui sont reliées directement à la notion courante t :

$$\forall LR \forall t (\text{learningResource}(LR) \wedge (\text{learningResourceTopic}(LR, t) \vee \exists t' (\text{learningTopicProperty}(t, t') \vee \text{learningTopicProperty}(t', t)) \wedge \text{learningResourceTopic}(LR, t'))) \Rightarrow \text{adaptiveContent}(LR)$$

Cette règle de relaxation signifie qu'une ressource est pertinente si elle traite de la notion courante t ou bien si elle traite de n'importe quelle notion t' reliée directement à la notion t (le prédicat $\text{learningTopicProperty}(x, y)$).

Le système OrPAF assiste les enseignants dans la recherche de ressources pédagogiques à partir du corpus ARIADNE en implémentant ces différentes règles de relaxation. Une fois le résultat de la requête obtenu, l'enseignant décide d'ajouter ou pas la ressource trouvée au corpus local. Ceci dépend évidemment de la pertinence de la ressource pour le domaine. L'enseignant est également libre de rechercher manuellement des ressources pédagogiques à partir du Web à l'aide des moteurs de recherche classiques.

2.2. Annotation conceptuelle des ressources pédagogiques

Les ressources pédagogiques recherchées des corpus du Web et considérées pertinentes pour le domaine par les enseignants sont annotées et ajoutées au corpus local. Pour une plus grande fiabilité, l'approche d'annotation que nous utilisons est semi-automatique : l'enseignant doit lui-même indiquer les notions traitées dans la ressource pédagogique. Le composant d'annotation d'OrPAF se base également sur les descripteurs LOM de la ressource pédagogique pour compléter les annotations de la ressource avec d'autres caractéristiques (par exemple l'auteur de la ressource, le format de la ressource, le type pédagogique de la ressource, etc.). Des annotations conceptuelles au format RDF sont créées et représentent les métadonnées de la ressource. Ces métadonnées utilisent comme vocabulaire les différentes ontologies du système OrPAF : l'ontologie du e-Learning et les ontologies de formation. Ellesinstancient des concepts de ces ontologies.

Pour annoter les ressources pédagogiques, nous étendons l'ontologie du e-Learning en définissant un sous-concept du concept *ELearningConcept* appelé *LearningResource* qui sert à annoter les ressources pédagogiques. Nous définissons également une sous-propriété à la propriété *eLearningProperty* appelée *learningResourceProperty* qui sert à représenter les différentes propriétés de la ressource (voir figure 20).

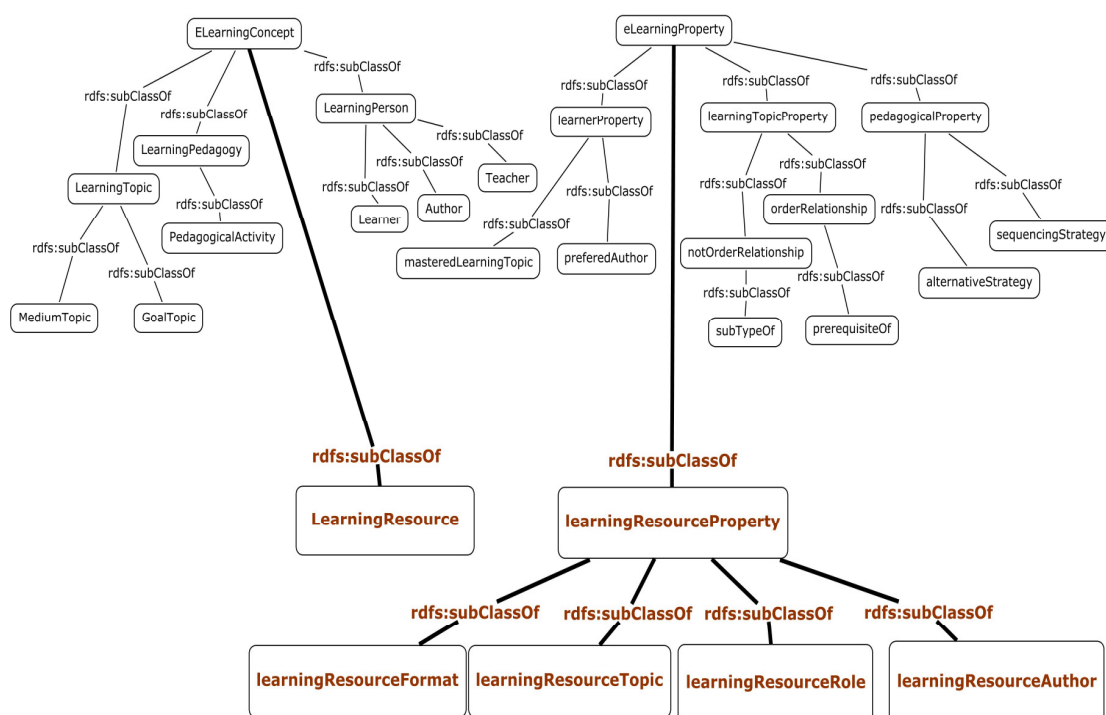


Figure 20: Extension de l'ontologie du e-Learning pour définir un schéma pour l'annotation des ressources pédagogiques

L'exemple de la figure 21 représente une partie des annotations conceptuelles de la ressource identifiée par l'URI « <http://www-sop.inria.fr/.../Course1> » et instance du concept *LearningResource* (défini dans l'ontologie du e-Learning). Cette ressource a comme type pédagogique *FormalDefinition* (défini dans le modèle pédagogique) instance du concept *LearningActivity* (défini dans l'ontologie du e-Learning). Les notions abordées dans cette ressource sont *reflexivity*, *Symmetry* et *Transitivity* (définies dans le modèle du domaine de l'algèbre) instances du concept *LearningTopic* (défini dans l'ontologie du e-Learning).

Une fois que la ressource pédagogique est annotée, elle est stockée avec ses annotations conceptuelles RDF dans le corpus local. Les annotations conceptuelles sont exploitées par le composant *AdaptationContent* pour identifier les ressources les plus pertinentes pour un contexte de formation donné (voir la section suivante).

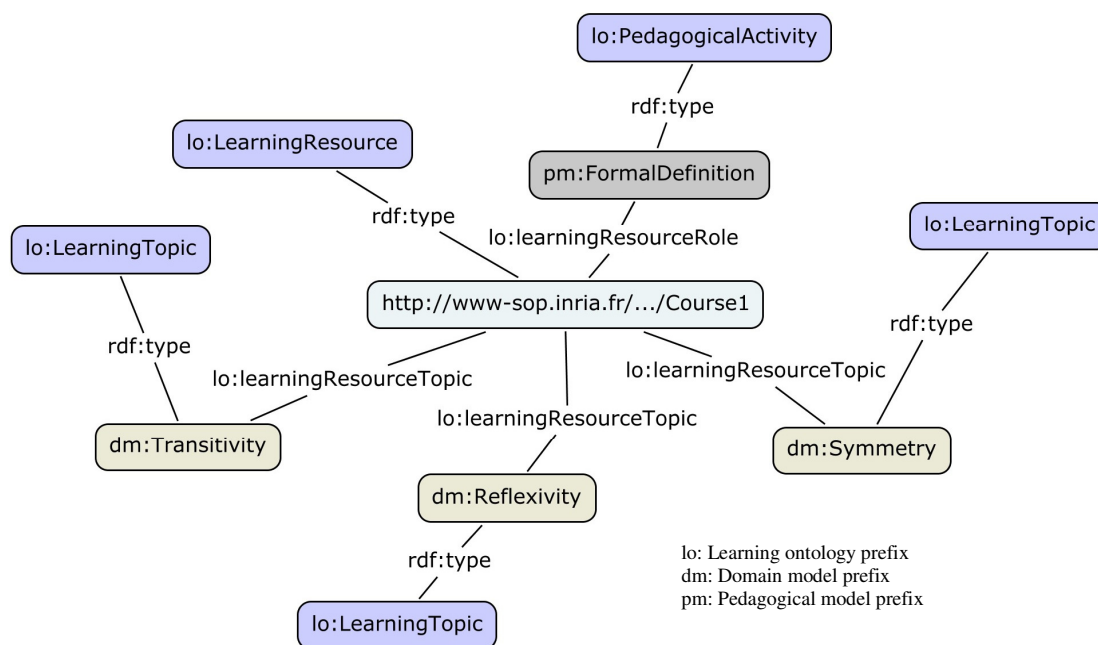


Figure 21 : Annotations conceptuelles de la ressource <http://www-sop.inria.fr/.../Course1>
(Le domaine de l'Algèbre)

2.3. Recherche locale des ressources pédagogiques

Cette tâche consiste à rechercher du corpus local des ressources pédagogiques pertinentes pour un contexte de formation donné. La recherche locale est mise en œuvre par Corese¹¹ (Corby et al., 2006), un moteur de recherche développé au sein de l'équipe Edelweiss. Corese nous permet d'exécuter des requêtes SPARQL sur des données de type RDF/OWL. Ainsi, nous utilisons localement Corese pour interroger les annotations conceptuelles des ressources pédagogiques. Nous pouvons, par exemple, rechercher toutes les ressources pédagogiques qui concernent une notion particulière du modèle du domaine, un auteur particulier, une activité pédagogique, etc. Le moteur Corese exploite la structure graphique des annotations RDF pour trouver des ressources pédagogiques répondant aux requêtes SPARQL émises.

Exemple d'une recherche locale :

Nous annotons une ressource pédagogique identifiée par l'URI « <http://www-sop.inria.fr/.../Cours1> », réalisée par Rose Dieng, de format HTML et qui représente une

¹¹ COnceptual REsource Search Engine

définition formelle des notions *Reflexivity*, *Symmetry* et *Transitivity* comme suit (lo pour learning ontology) :

```

<LearningResource rdf:about="http://www-sop.inria.fr/...Cours1">

  <lo:learningResourceRole>
    <lo:LearningActivity rdf:resource="#FormalDefinition"/>
  </lo:learningResourceRole >

  <lo:learningResourceTopic>
    <lo:LearningTopic rdf:resource="#Reflexivity"/>
  </lo:learningResourceTopic>

  <lo:learningResourceTopic>
    <lo:LearningTopic rdf:resource="#Symmetry"/>
  </lo:learningResourceTopic>

  <lo:learningResourceTopic>
    <lo:LearningTopic rdf:resource="#Transitivity"/>
  </lo:learningResourceTopic>

  <lo:learningResourceAuthor>
    <lo:Author rdf:ID="RoseDieng"/>
  </lo:learningResourceAuthor>

  <lo:learningResourceFormat> text/html <lo:learningResourceFormat />
  .....
</LearningResource>

```

La requête SPARQL suivante est utilisée pour rechercher une ressource pédagogique *r* qui traite de la notion pédagogique *EquivalenceRelation*, de type *Exercice* ou *Illustration* et dont l'auteur est un membre de l'équipe Edelweiss (dm pour domain model et pm pour pedagogical model) :

```

SELECT ?r WHERE
{?r rdf:type lo:LearningResource.
?r lo:learningResourceTopic dm:EquivalenceRelation.
{{?r lo:learningResourceRole pm:Exercice.} UNION
 {?r lo:learningResourceRole pm:Illustration. }}
?r lo:learningResourceAuthor ?a.
?a lo:memberOf "Edelweiss".
}.

```

3. Evaluation de la pertinence des ressources pédagogiques

Notre approche d'évaluation de la pertinence d'une ressource pédagogique pour un contexte de formation défini par une ACM est composée de deux phases d'analyse :

1- l'analyse sémantique : dans cette phase, pour calculer la pertinence d'une ressource pédagogique pour un contexte de formation, nous utilisons d'un côté les notions de l'ACM, leurs états et leurs relations et de l'autre côté les notions de l'annotation conceptuelle de la ressource.

2- une analyse de personnalisation : dans cette analyse, nous utilisons d'autres informations définies dans l'annotation de la ressource pédagogique tel que le type pédagogique de la ressource, le format de la ressource, l'auteur de la ressource, etc. dans le but de décider si la ressource est adaptée ou pas au contexte de formation de l'apprenant. Cette analyse est appliquée sur l'ensemble des ressources considérées comme pertinentes à l'issue de l'analyse sémantique.

3.1. Analyse Sémantique

À l'instar, du mécanisme de relaxation de requêtes utilisé dans la recherche distante pour répondre au problème de silences des requêtes (pas de réponse à la requête), nous proposons dans le cadre de l'analyse sémantique un mécanisme de calcul de la similarité sémantique entre les notions de la ressources et les notions constituant l'ACM de l'apprenant. Ce mécanisme de calcul de la similarité sémantique nous permet de dire si une ressource pédagogique est pertinente pour un contexte de formation et à quelle degré.

Dans cette analyse, nous considérons un seuil de similarité au-dessous duquel la ressource est considérée impertinente. Nous pouvons varier ce seuil de similarité en fonction des situations pour éviter les silences de requête. En effet, si une requête de recherche locale ne renvoie pas de réponse alors nous diminuons la valeur du seuil de similarité. Ceci permet d'avoir toujours des ressources pédagogiques même si elles ne sont pas toujours complètement adaptées au contexte de formation de l'apprenant. Il est à noter que si le corpus local est suffisamment peuplé de ressources pédagogiques et que si ces ressources assurent une bonne couverture des notions du domaine alors nous pouvons avoir à la fois des pertinences de ressources et des seuils de similarité élevés.

3.1.1. Démarche adoptée pour l'évaluation de la pertinence sémantique

Nous avons proposé une mesure qui permet de calculer la pertinence d'une ressource pédagogique pour un contexte de formation donné (Yessad et al., 2008a). Elle permet de dire

quel est le lien de similarité entre l'annotation conceptuelle d'une ressource pédagogique et l'ACM de l'apprenant en se basant sur les liens de similarité entre d'une part les notions de la ressource et d'autre part les notions accessibles de l'ACM. Nous nous intéressons exclusivement aux notions accessibles de l'ACM puisque les ressources annotées avec des notions inaccessibles dans l'ACM ne sont pas sensées intéressées l'apprenant. La valeur de similarité dépend de la notion courante qui intéresse l'apprenant, c'est-à-dire, la notion sur laquelle clique l'apprenant.

La démarche adoptée pour mesurer la pertinence d'une ressource pédagogique est la suivante :

- 1- Extraire la notion courante de l'ACM ;
- 2- Extraire les notions de la ressource pédagogique (les notions qui représentent des instances du concept *LearningTopic* et qui sont les co-domaines de la propriété *learningResourceTopic*) ;
- 3- Extraire les notions accessibles de l'ACM ;
- 4- Calculer le poids relatif de chaque notion extraite de la ressource en se basant sur les relations sémantiques qui relient cette notions à la notion courante dans l'ACM;
- 5- Calculer la pertinence sémantique de la ressource pédagogique pour l'ACM.

3.1.2. Calcul du poids d'une notion pédagogique

Nous attribuons un poids à chaque notion de la ressource pédagogique. Le poids d'une notion n'est pas absolu mais dépend de la notion courante qui intéresse l'apprenant et des relations qui relient cette notion pédagogique à la notion courante dans l'ACM. Soit le chemin de notions $\langle t_1, t_2, t_3, \dots, t_n \rangle$ de longueur n qui relie une notion t_n à une notion courante t_1 . Soit W_{i/t_1} (tel que $i > 1$ et $i < n+1$) le poids relatif de la notion t_i par rapport à la notion courante t_1 . Nous définissons le poids relatif d'une notion t_i comme suit :

$$\begin{aligned}
 W_{1/t_1} &= N \\
 W_{2/t_1} &= 1/a \\
 W_{i/t_1} &= (1/a) W_{i-1/t_1}, \quad i > 2
 \end{aligned}$$

'a' est une variable qui prend ses valeurs dans l'ensemble des nombre premiers comme suit :

$a=2$ si la relation entre t_{i-1} et t_i est *subTypeOf* ou sa relation inverse.

$a=3$ si la relation entre t_{i-1} et t_i est *prerequisiteOf* ou sa relation inverse.

$a=5$ si la relation entre t_{i-1} et t_i est *agregateOf* ou sa relation inverse.

Pour chaque nouvelle relation dans le modèle du domaine nous définissons une nouvelle valeur pour 'a' (a=7, a=11, a=13, etc.).

N est le poids de la notion courante et est considéré comme grand. Ceci permet d'attribuer une grande importance à la notion courante relativement aux autres notions de l'annotation conceptuelle.

La valeur inverse de 'a' représente le poids de la relation qui relie deux notions voisines dans l'ACM. Elle représente la force du lien sémantique entre deux notions voisines. Une seule contrainte est imposée : la variable 'a' prends ses valeurs exclusivement dans l'ensemble des nombres premiers. Cette contrainte nous permet d'identifier de manière unique à partir de la valeur du poids W_{i/t_j} les relations qui composent le chemin qui relie t_i à t_j dans l'ACM. Par exemple, le poids $W_{i/t_j} = 1/2 * 1/3 * 1/2 * 1/5$ signifie que le chemin qui relie t_i à t_j est composé de deux relations *subTypeOf* (de poids 1/2), d'une relation *prerequisiteOf* (de poids 1/3) et d'une relation *aggregateOf* (de poids 1/5). Cette caractéristique est vérifiée uniquement dans l'ensemble des nombres premiers. Notre objectif de l'utilisation des nombres premiers dans le calcul des poids est d'améliorer l'efficacité des tâches de sauvegarde et de restauration des ACM d'un apprenant (contexte de formation) (voir chapitre 6).

La correspondance entre les relations du modèle du domaine et les valeurs de 'a' a été définie par des experts du domaine qui s'appuie à la fois sur leurs connaissances du domaine et sur des expériences réalisées avec des personnes initiées dans le domaine.

Ces expériences s'inspirent des expériences réalisées par (Gandon et al., 2008) où des participants organisent et regroupent spatialement des notions selon leur proximité intuitive. Pour (Gandon et al., 2008), l'objectif de ces expériences était d'étudier les distances dans leur milieu naturel et de les comparer avec leurs simulations informatiques. Dans notre travail, nous avons réalisé des expériences similaires pour tenter de déterminer la force intuitive des liens qui peuvent relier les notions d'un domaine. Concrètement, nous avons réalisé deux expériences avec 10 participants à chaque fois : une sur une vingtaine de notions de l'algèbre et une autre sur une vingtaine de notions de l'algorithmique. Nous avons constaté que la plupart des participants ont positionné les notions reliées par une relation de subsomption (*subTypeOf*) de manière plus rapprochée que les notions reliées par une relation de prérequis (*prerequisiteOf*) (voir figure 22). Cependant, ces résultats ne peuvent être généralisés et dépendent énormément du domaine de la formation.

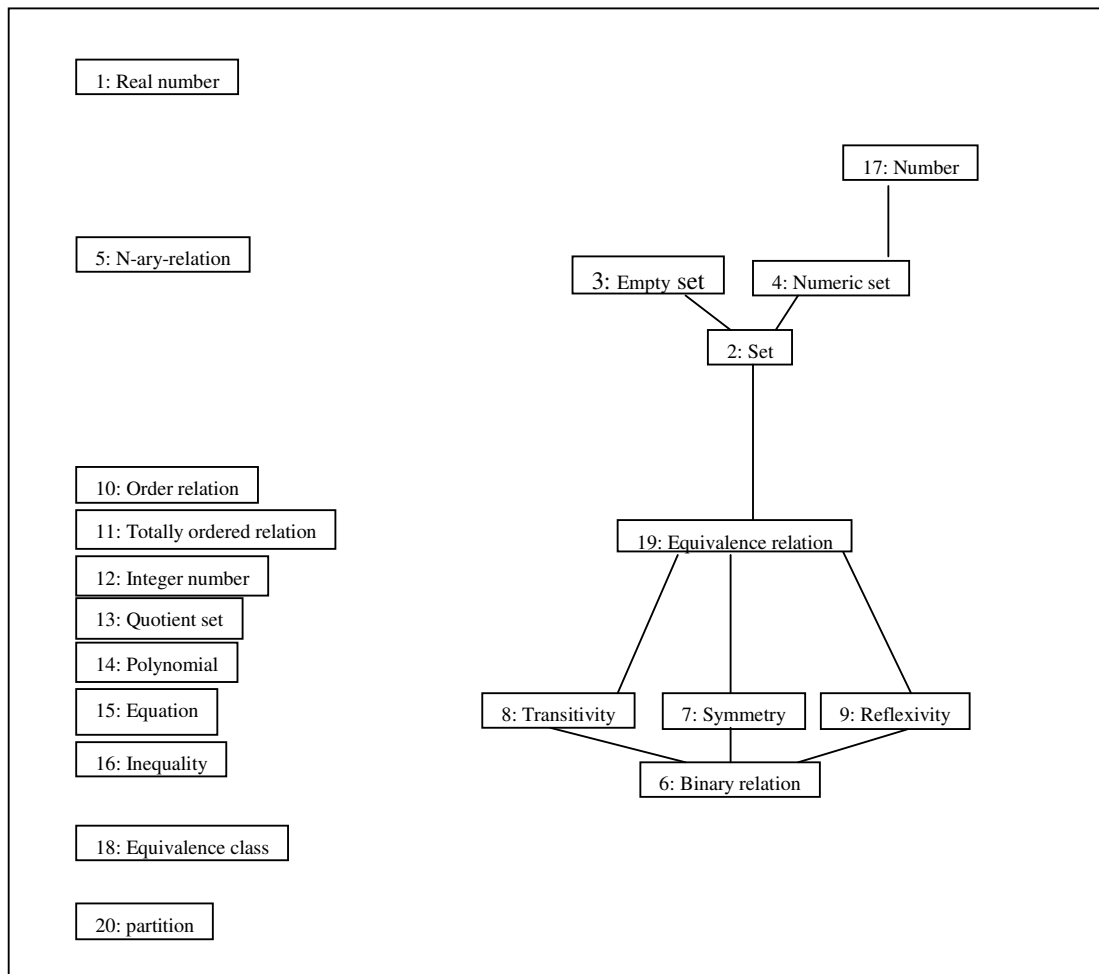


Figure 22 : Exercice de placement de notions de l'algèbre

Le modèle du domaine peut comporter des cycles et donc il est possible d'avoir plusieurs poids relatifs pour une même notion pédagogique. Dans ce cas de figure, nous retenons le plus petit poids relatif.

3.1.3. Calcul de la pertinence sémantique (Semantic relevance ou SR)

Une fois que les poids relatifs des notions de la ressource pédagogique sont calculés, nous pouvons calculer la pertinence sémantique (SR) de la ressource pédagogique pour le contexte de formation comme suit :

Soit E l'ensemble des notions communes à la ressource pédagogique et à l'ensemble des notions accessibles de l'ACM. Soit F l'ensemble des notions qui appartiennent à la ressource pédagogiques et n'appartiennent pas à l'ensemble des notions accessibles de l'ACM. Soit t la notion courante de l'ACM.

$$SR = \frac{\sum_{x \in E} w_{x/t}}{1 + \sum_{y \in F} w_{y/t}}$$

Cette mesure de la pertinence sémantique reflète les faits suivant :

- Une notion commune (qui appartient à E) augmente la pertinence d'une ressource pédagogique et son poids relatif représente son importance pour le contexte de formation.
- Une notion de la ressource pédagogique et qui n'est pas une notion accessible de l'ACM diminue la pertinence de la ressource pour le contexte de formation.
- Une ressource pédagogique avec un maximum de notions similaires à celles de l'ACM et avec des poids relatifs importants est une ressource pédagogique, à priori, pertinente.
- La pertinence sémantique d'une ressource pédagogique dépend du contexte de formation et de la notion courante.

Exemple:

Dans la figure 23, nous assignons un poids relatif à chacune des notions d'une ACM du domaine de l'algèbre. Par exemple, le poids de la notion *BinaryRelation* relativement à la notion courante *OrderRelation* est égal à 1/6 parce que le chemin qui relie ces deux notions est composé d'une relation *prerequisiteOf* (a=3) et d'une relation *subTypeOf* (a=2):
 $W_{BinaryRelation/OrderRelation} = 1/2 * W_{Transitivity/OrderRelation}$ and $W_{Transitivity/OrderRelation} = 1/3$.

En prenant l'exemple de la ressource pédagogique de la figure 21 identifiée par l'URI <http://www-sop.inria.fr/.../Course1> et annotée par les notions *Reflexivity*, *Symmetry* et *Transitivity*, la pertinence de cette ressource pédagogique pour l'ACM de la figure 23 où nous supposons que toutes les notions sont accessibles est calculée comme suit :

$E = \{Reflexivity, Symmetry, Transitivity\}$

$F = \emptyset$

$$SR = \frac{1/3 + 1/27 + 1/3}{1 + 0} = 0.703$$

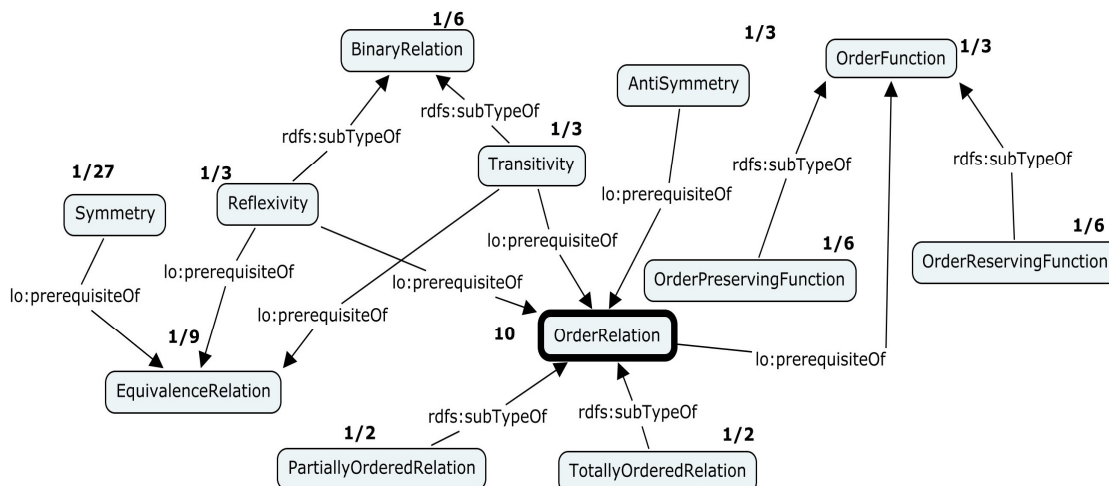


Figure 23: Assignement de poids relatifs aux notions reliées à la notion OrderRelation dans le domaine de l'algèbre

Le problème qui consiste à chercher du contenu adaptatif pour l'apprenant n'est pas encore complètement résolu : la pertinence sémantique SR calcule la similarité sémantique entre les notions de la ressource pédagogique et les notions de l'ACM mais d'autres facteurs autres que la connaissance de l'apprenant peuvent être pris en compte dans la définition du contenu adaptatif, tel que le type pédagogique de la ressource ou les préférences de l'apprenant. L'analyse de personnalisation a pour objectif de vérifier que les ressources pédagogiques considérées comme pertinentes à l'issue de l'analyse sémantique sont également adaptées aux préférences de l'apprenant et au modèle pédagogique.

3.2. Analyse de personnalisation

Nous appliquons l'analyse de personnalisation sur les ressources pédagogiques évaluées comme pertinentes à l'issue de l'analyse sémantique. Cette analyse est composée de deux tâches :

- 1- calculer la pertinence pédagogique de la ressource pédagogique ;
- 2- vérifier qu'une ressource pédagogique est adaptée aux préférences de l'apprenant, c'est-à-dire, qu'elle est de tel ou tel format, qu'elle est créée par tel ou tel auteur, etc.

3.2.1. Pertinence pédagogique d'une ressource pédagogique (PR) :

Les ressources pédagogiques pertinentes à l'issue de l'analyse sémantique, c'est-à-dire, les ressources dont la pertinence était supérieure à un seuil de similarité donné sont soumises à un

calcul pour déterminer leur pertinence pédagogique PR. Le calcul de la pertinence pédagogique est basé sur un calcul de similarité entre l'activité pédagogique de la ressource (représentée par la propriété *learningResourceRole*) et l'activité pédagogique courante du modèle pédagogique.

A l'instar de l'analyse sémantique, pour calculer la pertinence pédagogique d'une ressource, nous attribuons un poids relatif à son activité pédagogique. Ce poids relatif indique la similarité de cette activité pédagogique notée a_i avec l'activité pédagogique courante a_1 de l'apprenant.

Soit dans le modèle pédagogique le chemin d'activités pédagogiques $\langle a_1, a_2, a_3, \dots, a_n \rangle$ de longueur n qui relie une activité pédagogique a_n à une activité pédagogique courante a_1 . Soit $W_{ai/a1}$ (tel que $i > 1$ et $i < n+1$) le poids relatif de l'activité a_i par rapport à l'activité courante a_1 . Nous définissons le poids relatif d'une activité a_i comme suit :

$$\begin{aligned} W_{a1/a1} &= N \\ W_{a2/a1} &= 1/b \\ W_{ai/a1} &= (1/b) W_{ai-1/a1}, i > 2 \end{aligned}$$

Avec :

$b=2$ si la relation entre a_{i-1} et a_i est *alternativeStrategy* ou sa relation inverse.

$b=3$ si la relation entre a_{i-1} et a_i est *sequencingStrategy* ou sa relation inverse.

N est le poids de l'activité courante et est considéré comme grand. Cela permet d'attribuer une grande importance à l'activité pédagogique courante.

La pertinence pédagogique d'une ressource pédagogique est calculée comme suit :

Soit a_i l'activité pédagogique de la ressource pédagogique (le co-domaine de la propriété *learningResourceRole* et instance du concept *PedagogicalActivity*) et a_1 l'activité pédagogique courante du contexte de formation.

$$PR = SR * W_{ai/a1}$$

C'est seulement à l'issue du calcul de la PR que les ressources pédagogiques sémantiquement et pédagogiquement pertinentes sont triées pour être recommandées à l'apprenant dans l'ordre décroissant de leur PR.

Exemple :

Dans la figure 24, nous assignons un poids relatif à chacune des activités pédagogiques du modèle pédagogique présenté dans le chapitre précédent. Le type pédagogique de la ressource <http://www-sop.inria.fr/.../Course1> (voir figure 21) est *FormalDefinition*. Si l'apprenant cherche des exercices sur la notion *OrderRelation* (voir figure 23) alors la pertinence pédagogique de cette ressource est égale à $PR=SR * W_{FormalDefinition/Exercice} = 0.703 * 1/8 = 0.08$

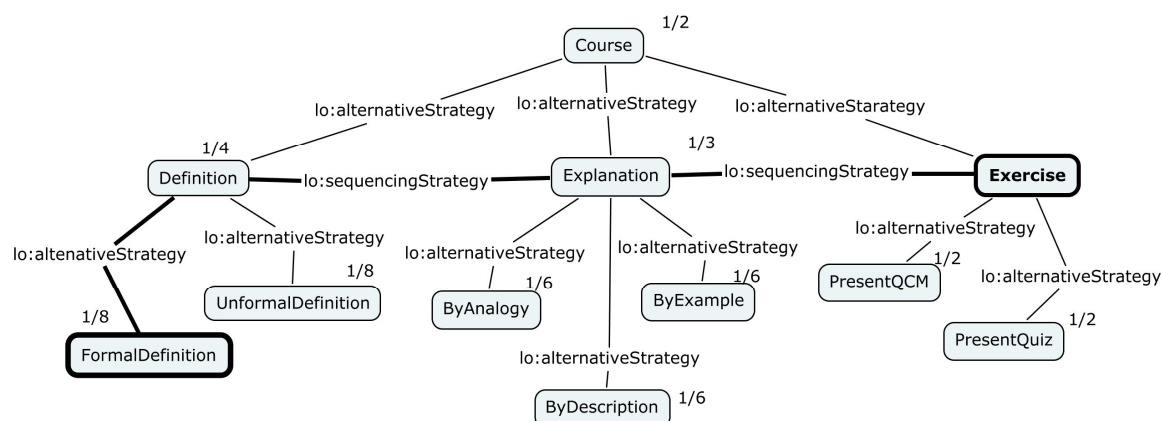


Figure 24: Assignement de poids relatifs aux activités pédagogiques reliés à l'activité pédagogique *Exercice* dans le modèle pédagogique

A l'instar du modèle du domaine, le modèle pédagogique peut comporter des cycles alors il est possible d'avoir plusieurs poids relatifs pour une même activité pédagogique (par exemple l'activité *Definition* dans la figure 24). Dans ce cas de figure, nous retenons le plus petit poids relatif.

3.2.2. Adaptation d'une ressource pédagogique aux préférences de l'apprenant

Cette tâche consiste à vérifier l'adaptation de la ressource pédagogique sémantiquement et pédagogiquement pertinente aux préférences de l'apprenant. Il s'agit, par exemple, de vérifier que l'auteur de la ressource et l'auteur préféré de l'apprenant, que le format de la ressource est le format préféré de l'apprenant ou que la langue de la ressource et la langue préférée de l'apprenant, etc.

Nous construisons des requêtes SPARQL qui permettent d'interroger les annotations conceptuelles des ressources pertinentes pour vérifier si elles s'adaptent aux préférences de l'apprenant (lm pour learner model).

Par exemple, nous recherchons parmi les ressources pertinentes les ressources *r* qui s'adaptent au format de ressources préféré de *Maria* et à l'auteur préféré de *Maria*. La requête SPARQL correspondante s'écrit comme suit:

```
SELECT ?r WHERE
{?r rdf:type lo:LearningResource.
 ?r lo:learningResourceFormat ?f
 ?r lo:learningResourceAuthor ?a
 "Maria" rdf:type lo:Learner
 "Maria" lm:preferredResourceFormat ?f.
 "Maria" lm:preferredResourceAuthor ?a. }
```

4. Annotation des tests

Un test est une ressource pédagogique spécifique. Dans l'environnement pédagogique adaptatif OrPAF, un test concerne une seule notion du domaine et il est de type Quiz ou QCM. Le test de type quiz est associé à une chaîne de caractère représentant la réponse correct au Quiz.

Pour annoter les ressources de type test, nous étendons l'ontologie du e-Learning en définissant un sous-concept du concept *LearningResource* appelée *LearningTest*. Ce dernier est utilisé pour construire les tests destinés aux apprenants et est spécialisé en deux sous-concepts appelés *LearningQuiz* et *LearningQCM*. Nous définissons également une sous-propriété à la propriété *learningResourceProperty* appelée *learningTestProperty* pour représenter les différentes propriétés d'un test qui sont : la propriété *learningQuizSolution* qui permet d'associer à un quiz une chaîne de caractères représentant la solution du quiz et la propriété *learningQCMSolution* qui permet d'associer à un QCM une liste de choix de solutions pour le QCM (voir figure 25).

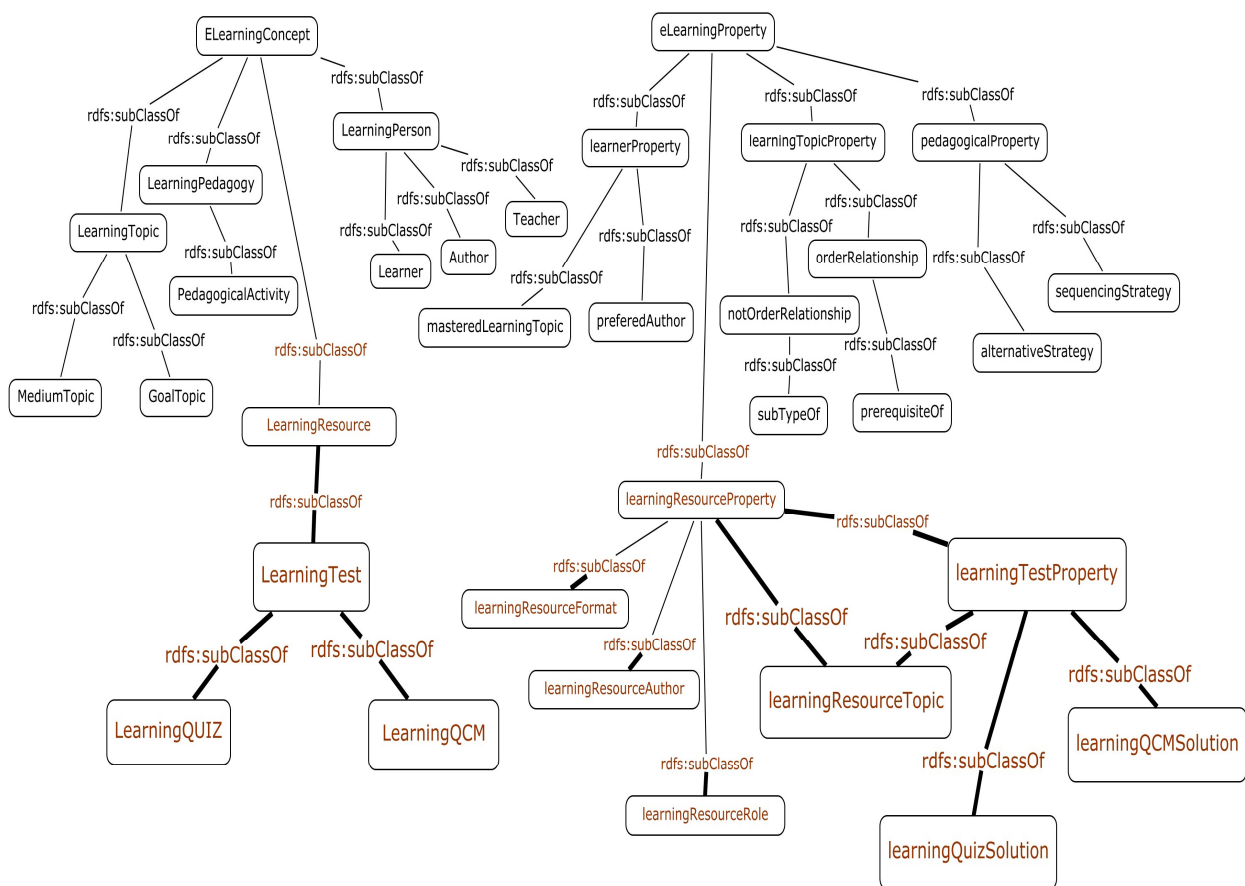


Figure 25: Extension de l'ontologie du e-Learning pour définir un schéma pour l'annotation des tests pédagogiques

Conclusion : Dans ce chapitre, nous avons présenté les approches utilisées pour la génération d'un cours hypermédia adapté aux caractéristiques de l'apprenant. Cette adaptation est nécessaire pour assurer l'efficacité de l'apprentissage. Cependant, un travail d'ingénierie pédagogique très coûteux en termes d'efforts humain, de temps et d'argent est nécessaire pour rendre les ressources du web réutilisables et adaptées dans le contexte d'une formation donnée. L'idée de notre travail de thèse est de proposer une automatisation de la génération de cours adaptatifs qui met en œuvre une partie de ce travail d'ingénierie pédagogique et assiste ainsi les responsables de formation et les enseignants qui veulent monter efficacement des formations destinées à des apprenants autonomes.

Nous sommes partis de l'idée qu'un cours est la combinaison de deux objets : une structure et un contenu. Pour que le cours généré soit adaptatif, nous avons conçu et implémenté un composant `AdaptationCourse` qui génère à la fois une structure adaptative et un contenu adaptatif pour l'apprenant et ceci en opérationnalisant la structure ontologique des modèles de connaissances représentés dans OrPAF.

La structure du cours est présentée sous la forme d'une carte conceptuelle construite en filtrant le modèle du domaine par des règles basées sur l'objectif de l'apprenant et ses contraintes temporelles. Les notions de la carte conceptuelle sont annotées suivant le modèle des connaissances de l'apprenant représenté dans son modèle apprenant.

L'apprentissage d'une notion particulière nécessite l'accès par l'apprenant à un contenu pertinent et adaptatif. Le composant `AdaptationCourse` fait appel à un composant `AdaptationContent` qui recherche, annoté et évalue la pertinence des ressources pédagogiques du web. D'abord une recherche distante est effectuée en exploitant les métadonnées LOM des ressources pédagogiques. La description LOM des ressources est ensuite enrichie en utilisant le vocabulaire fournis par les différentes ontologies du système OrPAF. Finalement, d'une part, des mesures de pertinence sémantique et pédagogique sont proposées pour détecter la pertinence d'une ressource pédagogique pour un contexte de formation donnée et d'autre part nous vérifions que les ressources pertinentes sont adaptées aux préférences de l'apprenant.

6. Réalisation de l'environnement pédagogique adaptatif OrPAF

Dans les deux chapitres précédents, nous avons présenté les choix conceptuels que nous avons adoptés pour la mise en œuvre du générateur de cours hypermédias adaptatifs. Ces choix concernent d'une part la représentation des connaissances qui servent à décrire des domaines de formation, des stratégies pédagogiques, des apprenants, des ressources pédagogiques et des tests et d'autre part l'utilisation de ces connaissances pour construire à la fois une structure adaptative du cours sous la forme d'une ACM et un contenu adaptatif composé de ressources pédagogiques pertinentes. Nous avons détaillé les approches qui nous permettent de construire une ACM adaptée aux caractéristiques de l'apprenant et celles qui nous permettent d'identifier la pertinence des ressources pédagogiques.

Dans ce chapitre, nous présentons les choix d'implémentation des différents composants de l'environnement pédagogique adaptatif OrPAF. Nous détaillons les outils et les API utilisés pour réaliser nos approches et choix conceptuels et qui sont intégrés dans les composants java du système OrPAF.

Le principe de mise en œuvre que nous avons choisi pour le développement de notre environnement logiciel est le principe Modèle-Vue-Contrôleur (MVC).

1. Architecture de l'environnement OrPAF

L'architecture de l'environnement OrPAF est une architecture organisée suivant le principe MVC avec une partie modèle (les objets persistants), une partie vue (les objets de l'interface graphique) et une partie contrôle qui fait la liaison entre la vue et le modèle. Nous présentons succinctement ces trois parties que nous détaillons dans les sections suivantes en abordant les usages de l'apprenant.

Le modèle décrit les données et les connaissances manipulées par le système OrPAF. Il assure la gestion de ces données et connaissances et garantit leur intégrité et leur cohérence. Dans notre cas, les différentes ontologies de formation (le modèle du domaine, le modèle apprenant, le modèle pédagogique), les annotations conceptuelles des ressources pédagogiques et les annotations des tests destinés aux apprenants doivent être cohérentes avec l'ontologie du e-Learning. Quant à la base de données que nous utilisons pour sauvegarder les sessions des apprenants et leurs identifiants (login et mot de passe), elle obéit à un schéma de données bien précis (voir section 6). Le modèle offre des méthodes pour accéder à ces données et les mettre à jour (insertion, suppression, changement de valeur).

Les résultats renvoyés par le modèle sont dénués de toute présentation. Nous verrons dans la suite que nous utilisons l'API du moteur sémantique Corese (Corby et al. 2006) pour interagir avec les modèles du e-Learning décrits en RDF/OWL et l'API JDBC pour interagir avec notre base de données.

La vue est l'interface avec laquelle l'apprenant interagit. Sa première tâche est de présenter les résultats renvoyés par le modèle. Sa seconde tâche est de recevoir toutes les actions de l'apprenant (clic de souris, sélection d'une entrée, boutons, ...). Ces différents événements sont envoyés au contrôleur. La vue n'effectue aucun traitement, elle se contente d'afficher les résultats des traitements effectués par le modèle. La vue qui est mise en œuvre dans OrPAF a pour objectif d'afficher un cours hypermédia adaptatif composé d'une structure interactive et évolutive (une ACM) et d'un contenu (une liste de ressources pédagogiques pertinentes consultable via l'interface de l'apprenant), de passer d'un contexte de formation à un autre et de passer des tests sur les notions accessibles de l'ACM. Notre vue est développée en Java avec des appels à des APIs JUNG pour l'affichage de l'ACM (voir section 4.2).

Le contrôleur prend en charge la gestion des événements de synchronisation pour mettre à jour la vue ou le modèle et les synchroniser. Il reçoit tous les événements de l'apprenant sur la vue et enclenche les actions à effectuer. Si une action nécessite un changement dans les modèles du e-Learning ou la base de données, le contrôleur demande la modification des

données au modèle et ensuite avertit la vue que les données ont changé pour qu'elle se mette à jour. Par exemple, la base de données qui sauvegarde les sessions des apprenants est mise à jour lorsqu'un contexte de formation est créé, modifié ou supprimé. Certains événements de l'apprenant ne concernent pas les données mais la vue, par exemple, lorsque l'apprenant passe momentanément de son contexte de formation à une activité de test. Dans ce cas, le contrôleur demande à la vue de se modifier. Cependant, la réussite de l'apprenant dans une activité de test implique la mise à jour du modèle (précisément la mise à jour du modèle apprenant). Le contrôleur n'effectue aucun traitement, ne modifie aucune donnée. Il analyse la requête de l'apprenant et se contente d'appeler le modèle adéquat et de renvoyer la vue correspondante à la demande.

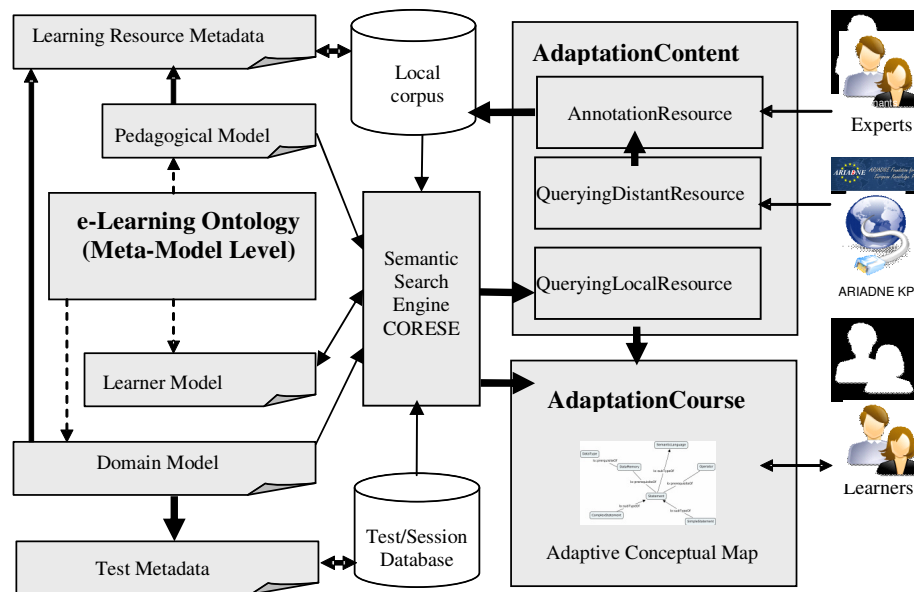


Figure 26 : Architecture de l'environnement OrPAF

La suite de ce chapitre est structurée suivant les scénarios d'utilisation du système OrPAF (voir figure 27):

- créer un contexte de formation, le consulter, le faire évoluer, le sauvegarder et le restaurer ultérieurement ;
- consulter des ressources pédagogiques pertinentes pour le contexte de formation de l'apprenant;
- passer des tests pour évoluer dans la formation.

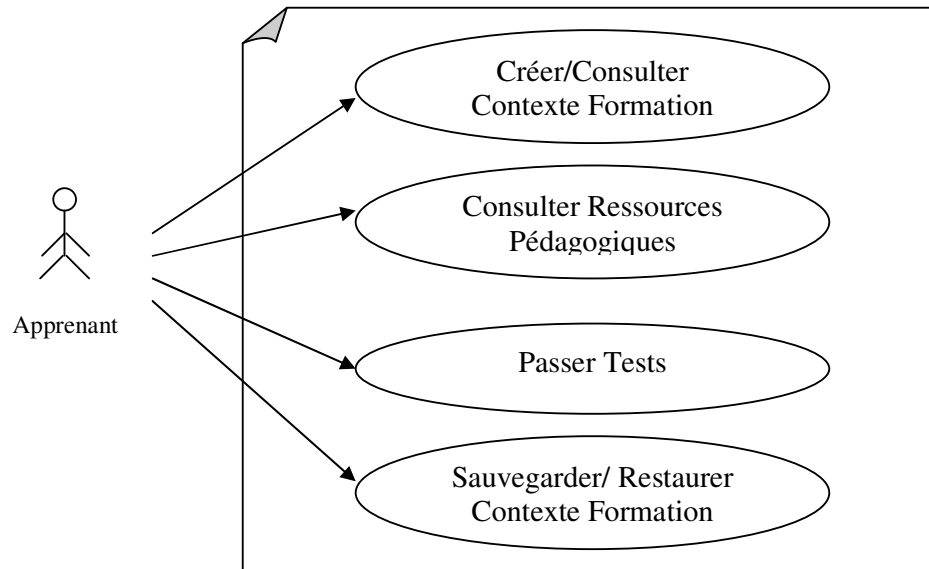


Figure 27 : Scénarios d'utilisation du système OrPAF par l'apprenant

2. Environnement proposé à l'apprenant

L'environnement pédagogique adaptatif OrPAF est utilisé par des apprenants pour atteindre des objectifs d'apprentissage en explorant un cours hypermédia adaptatif. Chaque apprenant a besoin de s'identifier pour accéder à sa session où il peut créer un nouveau contexte de formation, parcourir un contexte de formation en cours ou supprimer un contexte de formation dont l'objectif d'apprentissage est atteint. Un contexte de formation est une ACM construite à partir d'une notion cible de l'apprenant, de ses contraintes temporelles et de son état de connaissances. L'apprenant atteint son objectif lorsqu'il arrive à maîtriser toutes les notions de l'ACM : dans le modèle de l'apprenant, les notions de l'ACM sont des valeurs de la propriété *masteredLearningTopic* (définie dans l'ontologie du e-Learning).

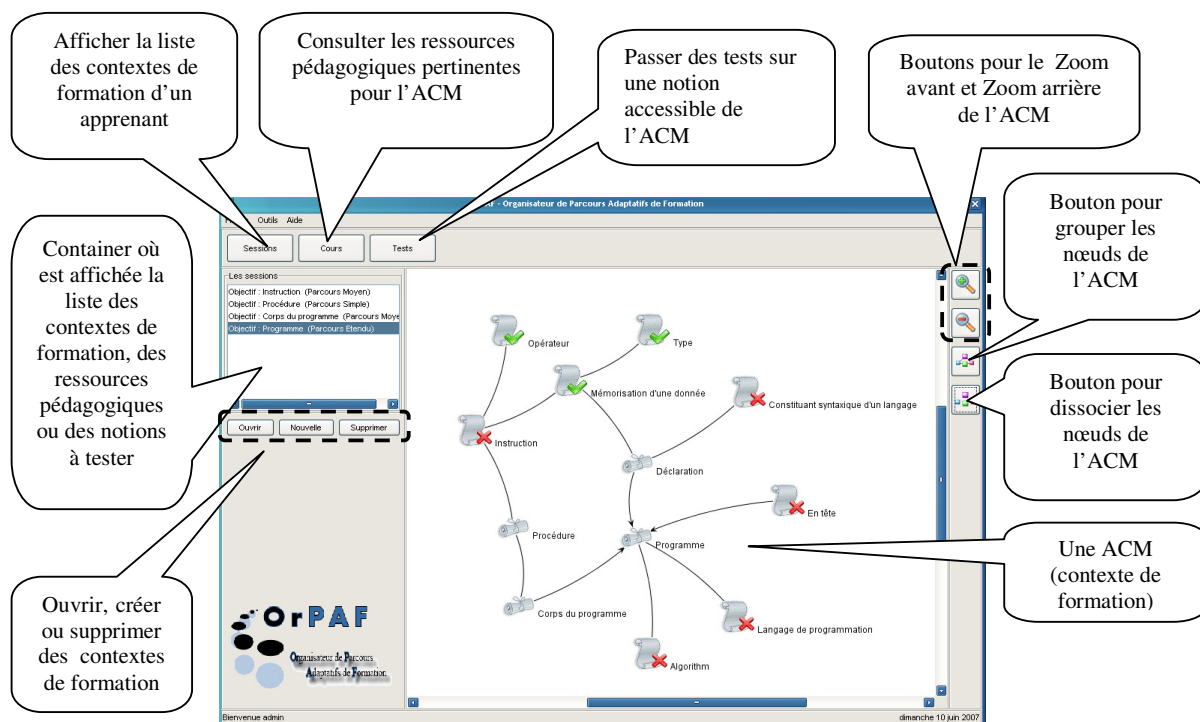


Figure 28 : Environnement graphique du système OrPAF

3. Extraction de données RDF

Dans le système OrPAF, les connaissances qui décrivent le domaine, la pédagogie, les apprenants, les ressources pédagogiques et les tests sont représentées par des annotations RDF quiinstancient l'ontologie du e-Learning représentée en OWL. Nous utilisons le moteur de recherche sémantique Corese (Corby et al. 2006) développé au sein de l'équipe Edelweiss-INRIA pour la recherche et l'exploitation de ces différentes connaissances.

3.1. Moteur de recherche sémantique Corese

Corese repose sur le formalisme des graphes conceptuels (Sowa, 1984) pour la représentation des connaissances ontologiques et des annotations. Rechercher des données revient à faire une recherche dans une base de graphes conceptuels (en anglais Conceptual Graph ou CG). Le moteur de recherche repose sur l'opération de projection des GC qui permet de retrouver des occurrences d'un graphe requête dans des graphes cibles. Une occurrence est soit le graphe recherché, avec des concepts instanciés, soit une spécialisation du graphe recherché au sens des hiérarchies de concepts et de relations.

L'originalité de Corese repose d'une part dans l'intégration des composants. D'autre part, Corese est compatible avec les formalismes RDF¹² et OWL¹³ développés au sein du W3C. L'utilisation de Corese garantit l'ouverture de l'environnement OrPAF aux données/services du Web sémantique.

Le moteur de recherche sémantique Corese est écrit en Java, il est constitué des composants logiciels principaux suivants (Corby et al., 2006) (voir figure 29):

- un serveur de graphes conceptuels,
- un traducteur RDF/CG et CG/RDF,
- un parser de langage de requêtes et de règles,
- un moteur de recherche basé sur la projection de graphes,
- un moteur de règles d'inférence en chaînage avant,
- un serveur Web sémantique.

L'architecture du serveur Web sémantique (SWS) intégrant Corese et basé sur Tomcat est celle d'une couche supplémentaire au-dessus du moteur de recherche gérant les échanges avec le Web et avec les ressources locales. L'application Web implantant le SWS permet à un Web Master de concevoir un site reposant, entièrement ou en partie, sur des requêtes posées au moteur de recherche Corese et sur la transformation du résultat de ces requêtes en pages HTML, JSP ou SVG en utilisant des feuilles de style XSLT.

Afin d'assurer une indépendance maximale de l'application au domaine, l'architecture du serveur implante une forme de MVC (Model-View-Controller) pour séparer les aspects données, les aspects calculs et les aspects rendu. Pour ce faire, l'application Web se décompose en plusieurs composants réutilisables, notamment:

- Des servlets implantant des tâches récurrentes dans les scénarios d'utilisation d'un serveur Web sémantique: résolution d'une requête personnalisée, modification d'une annotation, traitement d'un formulaire, gestion des gabarits des pages, etc.
- Des feuilles de style implantant des transformations XSLT dans un langage déclaratif et donc modifiable par l'utilisateur. Elles sont utilisées pour des tâches de rendu (format d'affichage d'une réponse, génération du JSP d'un formulaire, génération d'une vue de l'ontologie, etc.) et des tâches de modification de fichiers XML (édition d'une annotation, d'une classe, etc.). Corese est utilisé pour étendre les feuilles de style XSLT en fournissant des fonctions sémantiques supplémentaires. De plus les feuilles XSLT sont aussi utilisées

¹² Resource Description Framework

¹³ Ontology Web Language

pour décrire des patrons d'extractions de données pour des bases de données légataires par exemple.

- Des Tags JSP permettent de faire appel à Corese à l'intérieur de pages JSP et permettent ainsi de générer des éléments à la demande comme, par exemple, un menu, une liste de requêtes contextuelles ou la présentation d'une sous-partie de l'ontologie.

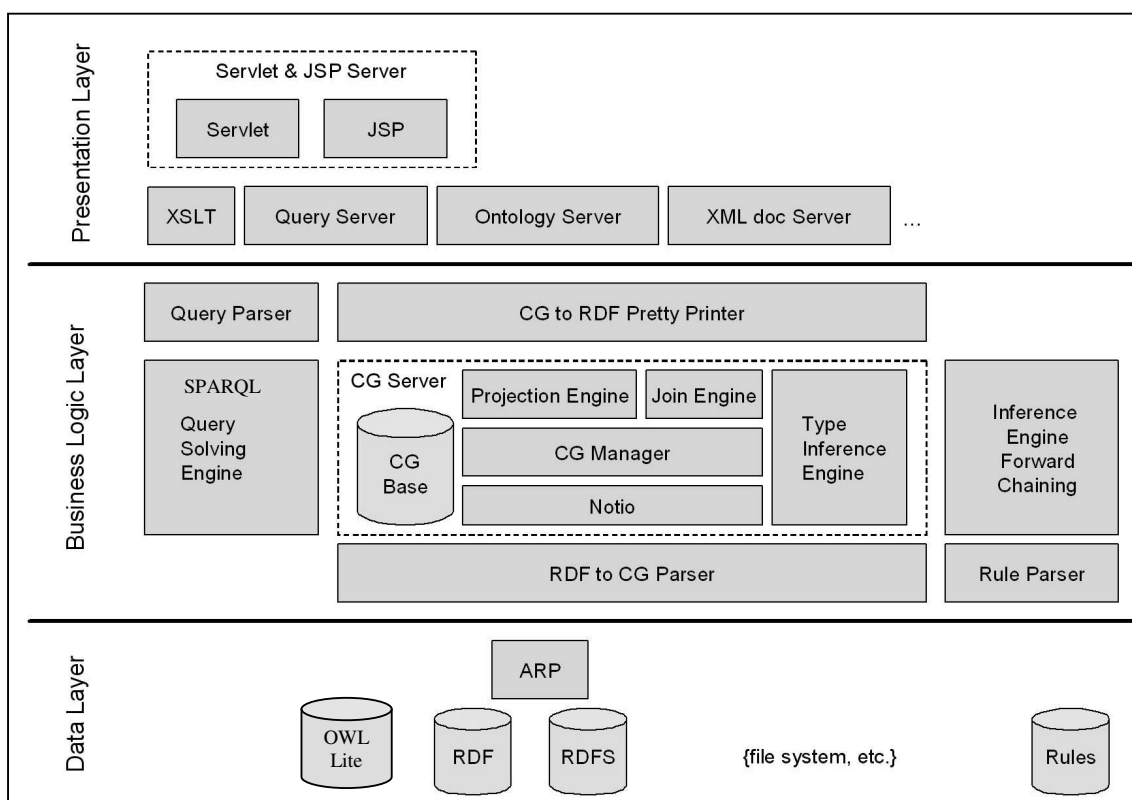


Figure 29: Architecture 3-tiers de Corese (Corby et al., 2006)

3.2. Utilisation de Corese dans le système OrPAF

Dans notre travail, nous utilisons Corese dans la réalisation des fonctions suivantes :

- Recherche des notions du domaine et de leurs relations dans le but de construire l'ACM du cours ;
- Recherche des connaissances sur l'apprenant, dans le but par exemple, d'annoter les notions de l'ACM avec les différentes icônes d'état ;
- Recherche des connaissances du modèle pédagogique, pour par exemple, calculer la pertinence pédagogique d'une ressource pédagogique ;
- Recherche des connaissances à partir des annotations conceptuelles d'une ressource, pour par exemple, calculer la pertinence sémantique de la ressource ;

- Recherche des connaissances à partir des annotations de tests pour sélectionner des tests adaptés à l'apprenant ;
- Mettre à jour le modèle de l'apprenant suite à un test.

Afin d'extraire les annotations RDF qui nous intéressent, nous soumettons des requêtes SPARQL au moteur sémantique Corese. Pour chercher la correspondance entre la requête SPARQL et les annotations RDF en conformité avec l'ontologie du e-Learning, Corese transforme la requête SPARQL, les annotations RDF et l'ontologie OWL en graphes conceptuels et les apparie. Corese renvoie le résultat de la requête sous la forme d'un graphe RDF, XML ou SPARQL result (Corby et al. 2006).

Dans la sortie suivante, Corese nous renvoie en XML deux projections possibles (formatées en gras) de la requête (formatée en gras italique) qui cherche les prérequis de la notion *Statement* dans le modèle du domaine de l'algorithmique:

```

<?xml version='1.0' encoding='Cp1252'?>
<cos:result xmlns:cos='http://www.inria.fr/acacia/corese# '>
<cos:tquery>
<![CDATA[
prefix lo:
<http://www.orpaf.com/learningontologies/learningOntology.rdfs#>
prefix instance:
<http://www.orpaf.com/learningontologies/learningOntology.rdfs-
instances#>

select list display xml distinct ?x
where
{ { ?x lo:prerequisiteOf instance:Statement . } union { ?x
rdf:type lo:LearningTopic . } }
]]>
</cos:tquery>

<sparql xmlns='http://www.w3.org/2005/sparql-results#'
xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#' >
<head>
<variable name='x' />
</head>
<results ordered='false' distinct='true' >
<result>
<binding
name='x'><uri>http://www.orpaf.com/learningontologies/learningOntology.rdf
fs-instances#DataMemory</uri></binding>
</result>
<result>
<binding
name='x'><uri>http://www.orpaf.com/learningontologies/learningOntology.rdf
fs-instances#Operator</uri></binding>
</result>
<result>
</results>
</sparql>
</cos:result>

```

Cette sortie de Corese est renvoyée dans un objet java de la classe **IResults** implémentée par l'API de Corese. C'est cet objet qui est manipulé par les objets de l'environnement OrPAF.

4. Session d'apprentissage d'un apprenant

La session d'un apprenant est composée de son modèle apprenant, de l'ensemble de ses contextes de formation et des ressources pédagogiques pertinentes à chaque contexte de formation. Un apprenant peut avoir plusieurs contextes de formation dans le système. Un apprenant peut interrompre sa session d'apprentissage et la reprendre plus tard. Les scénarios d'utilisation d'un contexte de formation dans la session d'un apprenant sont (voir figure 30) :

- La création d'un nouveau contexte de formation en insérant dans le système la notion cible à apprendre et les contraintes temporelles de l'apprenant. Le système utilise ces

- deux informations ainsi que le niveau de connaissances de l'apprenant pour filtrer le modèle du domaine et créer une ACM (la fonction createACM);
- La sauvegarde d'un contexte de formation en cours (la fonction storeACM) ;
 - La restauration d'un contexte de formation sauvegardé dans une session précédente (la fonction restoreACM) ;
 - L'affichage d'un nouveau contexte de formation ou d'un contexte de formation existant (la fonction displayACM).

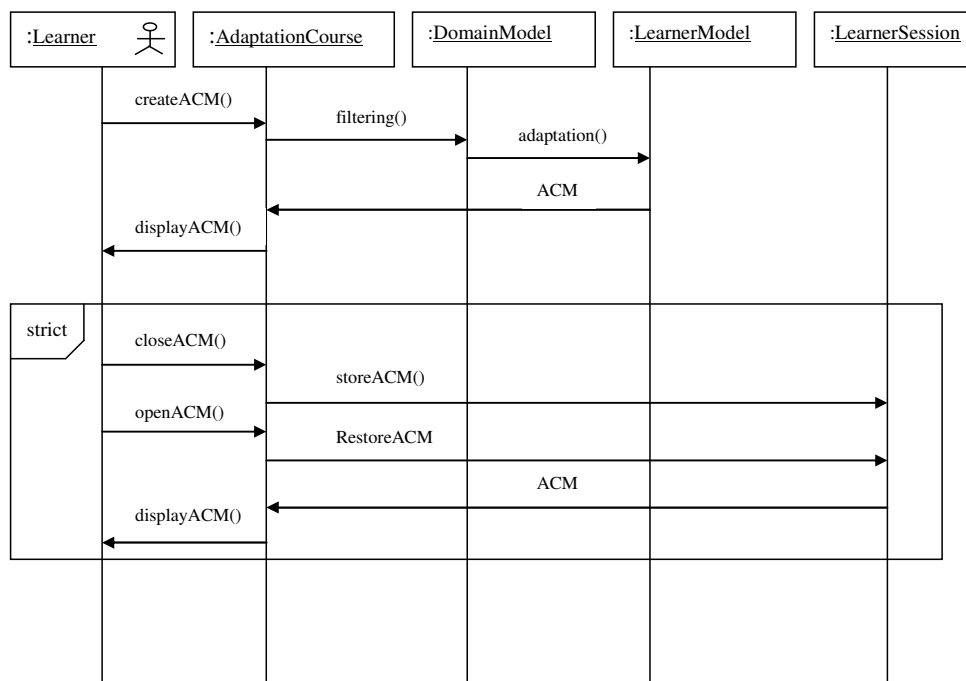


Figure 30: Le diagramme de séquence des scénarios d'utilisation d'une ACM.

4.1. Construction d'une ACM

Une ACM est une carte de notions reliées et appartenant à un domaine de connaissances. Elle est construite en filtrant le modèle du domaine (décrit en RDF dans le système OrPAF) en se basant sur la notion cible de l'apprenant (l'objectif de l'apprenant), ses contraintes temporelles et son état de connaissances. La notion cible et les contraintes temporelles sont fournies par l'apprenant alors que son état de connaissances est décrit dans son modèle apprenant où les notions connues par l'apprenant sont représentées en tant que valeurs de la propriété *masteredLearningTopic*.

La construction d'une ACM consiste d'abord à faire appel à Corese pour extraire des données RDF du modèle du domaine et ensuite à afficher ces données sous la forme d'un graphe.

Dans l'exemple suivant, nous décrivons une requête SPARQL envoyée à Corese et utilisée dans le filtre hiérarchique du modèle du domaine, c'est-à-dire, dans l'extraction des notions ascendantes et des notions descendantes d'une notion cible:

```
// Requête SPARQL pour renvoyer les ascendants y et les descendants x
// d'une notion cible contenue dans la variable goalTopic

String queryString=
"PREFIX
lo:<http://www.orpaf.com/learningontologies/learningOntology.rdfs#>
PREFIX
instance:<http://www.orpaf.com/learningontologies/learningOntology.rdfs-
instances#>
select DISTINCT ?x ?y
where
{UNION { ?x lo:subTypeOf instance:"+goalTopic+" }
UNION { instance:"+goalTopic+" lo:subTypeOf ?y }
}";

// Exécution de la requête SPARQL sur le données RDF
try {
    IResults resultRDF = engine.SPARQLQuery(queryString);
    return resultRDF;
}

catch (EngineException e) {
    e.printStackTrace();
    return null;
}
```

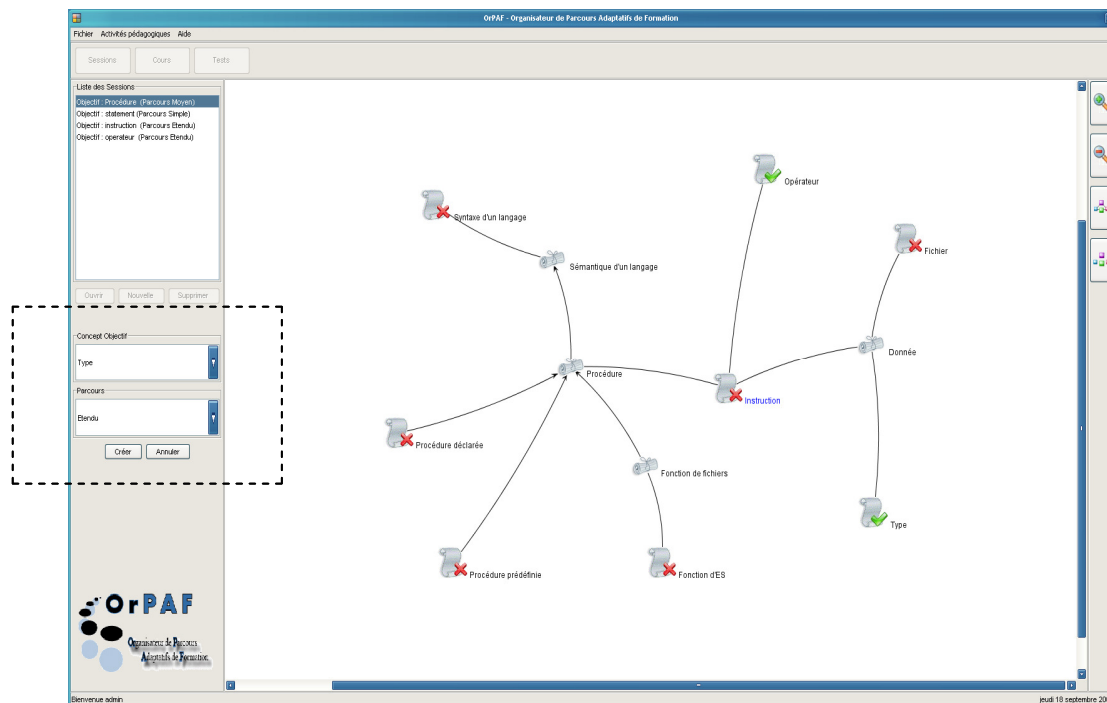


Figure 31: Création d'une ACM relationnelle (un parcours étendu) avec la notion cible *Type* (Le contexte de formation affiché est celui de la notion cible *Procédure* (ACM Hiérarchique))

4.2. Affichage d'une ACM

Une ACM est un graphe non orienté composé de nœuds et de relations entre les nœuds (appelées arêtes dans la terminologie des graphes). Pour l'affichage interactif de l'ACM, nous avons choisis d'utiliser JUNG (the Java Universal Network/Graph Framework), une bibliothèque libre pour la modélisation, l'analyse et la visualisation des données qui peuvent être représentées sous forme d'un graphe ou d'un réseau (voir figure 32). L'architecture de JUNG est conçue pour supporter différentes représentations des nœuds et de relations, telles que les graphes orientés, les graphes multi-modaux, les graphes avec des nœuds parallèles et les hypergraphes.

JUNG implémente des algorithmes de la théorie des graphes, du data mining, des réseaux sociaux tels que les routines de classification, de décomposition, d'optimisation, de génération aléatoire de graphes, etc.

JUNG offre également des algorithmes d'affichage de graphes (Layout algorithms) qui peuvent être utilisés pour faire des affichages personnalisés de graphes. Dans notre implémentation, nous avons utilisé l'algorithme KKLayout (The Kamada-Kawai algorithm for node layout) qui permet d'afficher des graphes aérés et ergonomiques.

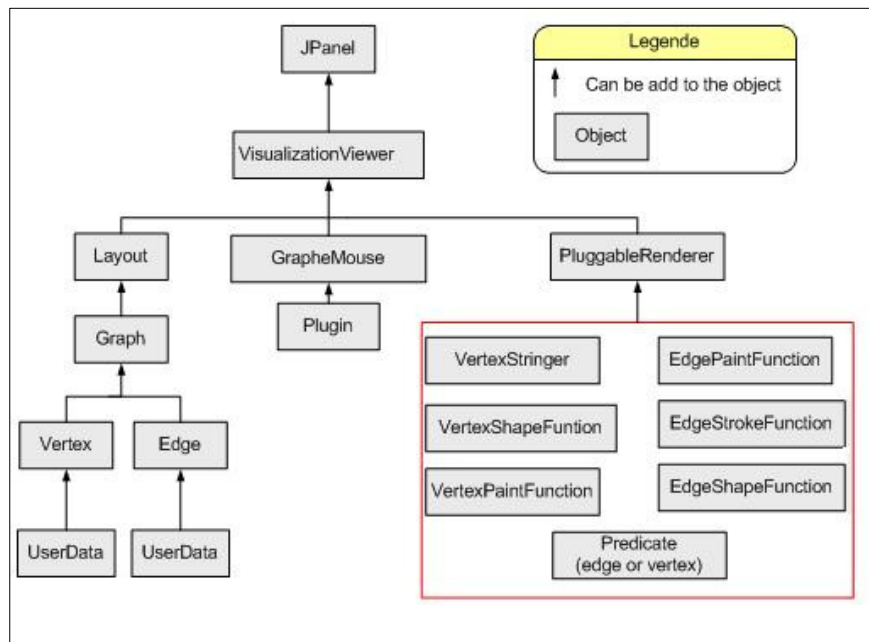


Figure 32 : Architecture de la bibliothèque JUNG

Les nœuds de l'ACM représentent des notions du domaine et sont annotés par différentes icônes qui représentent l'état des notions pour un apprenant donné (voir figure 33).

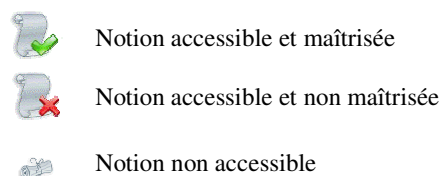
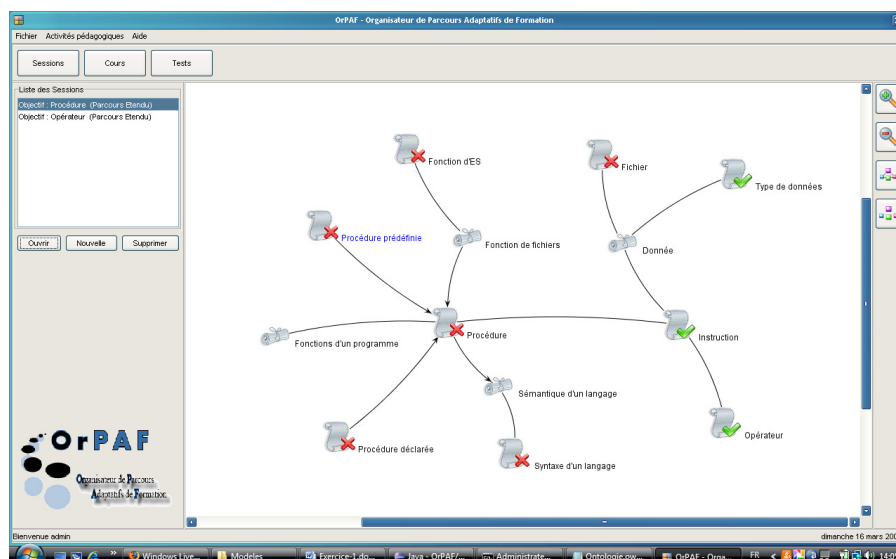


Figure 33 : Annotation graphique des notions utilisées dans une ACM

4.3. Evolution d'une ACM

L'évolution de l'ACM d'un apprenant est liée à sa réussite ou son échec à l'issue des tests. Un apprenant peut choisir de passer un test sur n'importe quelle notion accessible de l'ACM. Les notions inaccessibles sont exclues du test car, pour le système, elles ne sont pas connues de l'apprenant et donc il n'est pas censé être testé dessus. Un test est une séquence de QCM ou de Quiz que nous construisons comme suit :

- 60% des QCM/Quiz sont sur la notion choisie par l'apprenant ;
- 40% des QCM/Quiz sont sur les prérequis de cette notion.

Cette approche qui consiste à intégrer dans le test d'une notion des QCM/Quiz sur ses notions prérequis vise à s'assurer que l'apprenant maîtrise ces notions. En effet, étant donné que la connaissance que possède le système sur le niveau d'un apprenant n'est pas toujours fiable, cette approche est un moyen de remettre en cause cette connaissance et ainsi obtenir un profil apprenant plus fiable. En effet, si un apprenant échoue dans les QCM/Quiz qui concernent les prérequis d'une notion alors cette notion devient inaccessible dans l'ACM (NotAccessible) et ses notions prérequis deviennent non maîtrisées (AccessibleAndNotMastered). Par contre si

l'apprenant réussit le test d'une notion alors cette notion devient maîtrisée dans l'ACM (AccessibleAndMastered).

4.4. Sauvegarde/Restauration d'une ACM

La taille du fichier RDF du modèle du domaine étant souvent très grande, le filtrage systématique de ce modèle à chaque fois qu'un apprenant demande à afficher son contexte de formation implique des coûts de traitement importants. Ainsi, nous proposons une autre approche qui consiste à sauvegarder le contexte de formation d'un apprenant dans une table de données numériques et de le restaurer à partir de cette structure de sauvegarde numérique.

Un algorithme de restauration du contexte de formation d'un apprenant à partir de cette structure de sauvegarde numérique a été conçu et implémenté dans OrPAF afin d'éviter le chargement et le filtrage systématiques et coûteux du modèle du domaine.

Structure de sauvegarde :

La sauvegarde de la session d'un apprenant consiste à sauvegarder dans une table d'une base de données les contextes de formation de la session. Nous avons proposé pour cette structure de sauvegarde le schéma suivant:

LearnerSession (LearnerID, LearningContextID, TopicName, SignedTopicWight)

Cette structure attribuée à chaque notion de chaque contexte de formation de chaque apprenant un poids signé. Ce dernier est calculé comme suit :

Soit le chemin de notions $\langle t_1, t_2, t_3, \dots, t_n \rangle$ de longueur n qui relie dans une ACM une notion cible t_1 à une notion t_n . Le poids de la notion t_n est calculé récursivement par rapport à la notion cible t_1 comme suit :

$$W_{t_n/t_1} = (1/b)^{n-1} W_{t_{n-1}/t_1}, n > 1$$

$$W_{t_1/t_1} = 1$$

A l'instar du facteur 'a' utilisé dans le calcul de la relevance sémantique SR d'une ressource pédagogique, le facteur 'b' est une variable qui prend ses valeurs dans l'ensemble des nombres premiers comme suit :

$b=2$ si la relation entre t_{i-1} et t_i est *subTypeOf* ou sa relation inverse.

$b=3$ si la relation entre t_{i-1} et t_i est *prerequisiteOf* ou sa relation inverse.

$b=5$ si la relation entre t_{i-1} et t_i est *agregateOf* ou sa relation inverse.

Nous calculons la valeur du poids signé PoidsNotionSigné d'une notion t_n à partir du poids de cette notion comme suit :

$$\begin{array}{ll} \text{PoidsNotionSigné}(t_n) = -1/W_{t_n/t_1} & \text{si } t_n \text{ est à l'extrémité d'un chemin de l'ACM,} \\ \text{PoidsNotionSigné}(t_n) = 1/W_{t_n/t_1} & \text{sinon.} \end{array}$$

Au niveau du poids signé, le facteur b est élevé à la puissance $i-1$ ($n \geq i \geq 2$) pour représenter la position de la notion t_i dans le chemin des notions $\langle t_1, t_2, t_3, \dots, t_n \rangle$ et permet d'identifier la séquence exacte des relations qui constituent le chemin reliant la notion cible t_1 à la notion t_i . En effet, nous nous basons sur la propriété mathématique qu'à tout nombre de pouvoir être écrit de manière unique sous la forme d'une multiplication de nombres premiers élevés à des puissances entières. Cette caractéristique nous permet de reconstituer les chemins de l'ACM en utilisant uniquement les poids signés $PoidsNotionSigné$ de ses différentes notions.

Le signe du poids signé d'une notion exprime le fait qu'elle est à l'extrémité d'un chemin ou pas. Cette information est utile dans l'algorithme de restauration pour réduire le nombre de combinaisons de chemins (ACM) à reconstituer et donc optimiser le temps de calcul. Par exemple, les valeurs de $PoidsNotionSigné$ des notions de l'ACM de la figure 34 sont sauvegardées comme suit :

Table 2 : Structure de sauvegarde de l'ACM de la figure 34

| Notions t_i (la notion cible est t_1) | $PoidsNotionSigné(t_i)$ | $PoidsNotionSigné(t_i)$ (Ecriture normalisée de t_1 à t_i). (A lire de gauche à droite) |
|--|-------------------------|--|
| t_2 | 2 | 2 |
| t_3 | 8 | 2×2^2 |
| t_4 | -3 | -3 |
| t_5 | 3 | 3 |
| t_6 | -6 | -3×2 |
| t_7 | -6 | -3×2 |
| t_8 | -64 | $-2 \times 2^2 \times 2^3$ |

Si nous analysons la valeur du $PoidsNotionSigné$ de la notion t_8 , nous pouvons déduire que cette notion est positionnée à l'extrémité d'un chemin ($PoidsNotionSigné(t_8) < 0$) et qu'à partir de la notion cible t_1 ce chemin est composé dans l'ordre de trois relations *subTypeOf* consécutives. Le $PoidsNotionSigné$ de la notion t_7 indique que cette notion est à l'extrémité d'un chemin composé à partir de la notion cible t_1 d'une relation *prerequisiteOf* suivie d'une relation *subTypeOf*. Enfin, la valeur du $PoidsNotionSigné$ de la notion t_2 indique que cette notion se trouve à l'intérieur d'un chemin ($PoidsNotionSigné(t_2) > 0$) et que son sommet dans l'ACM est adjacent au sommet de la notion cible t_1 via une relation *subTypeOf*.

Nous pouvons également affirmer que la notion t_6 ne peut être reliée à la notion t_4 mais peut être reliée à la notion t_5 car t_4 se trouve à l'extrémité d'un chemin ($\text{PoidsNotionSigné}<0$) et pas à l'intérieur d'un chemin.

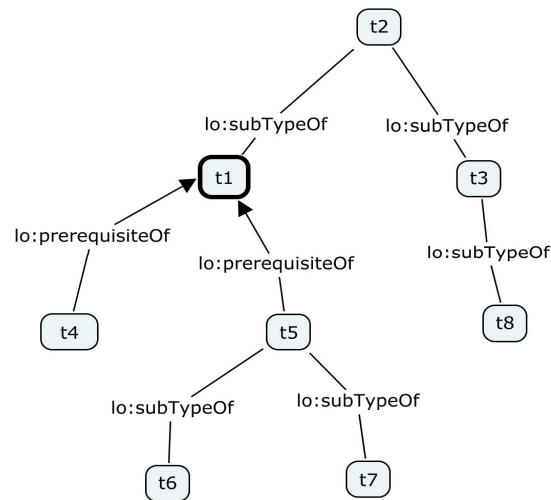


Figure 34 : une ACM

Dans certains cas l'algorithme de restauration proposé calcule à partir de la structure de sauvegarde et pour un même contexte de formation, plusieurs ACM possibles et donc exige des accès au modèle du domaine pour décider de la solution à sélectionner. Ces cas apparaissent généralement pour les ACM relationnelles avec des longueurs de chemin $m>2$.

Algorithme de restauration

L'algorithme de restauration sert à générer l'ACM de l'apprenant à partir de sa structure de sauvegarde en identifiant les notions composant les chemins de l'ACM. Le principe de cet algorithme est de commencer par identifier récursivement les notions des chemins les plus longs jusqu'aux chemins les plus courts. Pour un chemin donné, le calcul est récursif pour chacune des notions du chemin et commence par la notion la plus éloignée de la notion cible t_1 jusqu'à la notion adjacente à celle-ci.

Pour réaliser notre algorithme de restauration, nous avons définis les fonctions suivantes :

- la fonction longueur (t_i) qui renvoie la longueur du chemin qui relie la notion cible t_1 à t_i . Par exemple longueur (t_8)=3 ;
- la fonction path(t_i) qui renvoie la séquence des relations qui relie la notion cible t_1 à t_i . Par exemple path(t_7)= (prerequisiteOf, subTypeOf) ou path(t_7)=(3,2) ;
- la fonction Matching(path(t_j), path(t_i)) qui vérifie l'appariement entre deux chemins path(t_j) et path(t_i) et renvoie une valeur booléenne. Cette fonction vérifie qu'un chemin (le plus court) est identique à une partie située à l'extrémité d'un autre chemin (le

chemin le plus long) et la comparaison commence à la notion cible t_1 . Par exemple $\text{path}(t_8)=(2,2,2)$ s'apparie avec $\text{path}(t_3)=(2,2)$ ou $\text{path}(t_2)=(2)$ et ne s'apparie pas avec $\text{path}(t_5)=(3)$ ou $\text{path}(t_5)=(3,2)$.

Les étapes de l'algorithme sont :

- 1- Pour toute valeur absolue $|\text{PoidsNotionSigné}(t_i)| = \alpha_1 \times \alpha_2^2 \times \dots \times \alpha_j^{j-1}$ d'une notion t_i ($i > 1$) de l'ACM, construire un chemin de relations $\text{path}(t_i) = \langle t_1, \dots, t_i \rangle$ de longueur j de la notion cible t_1 à la notion t_i ;
- 2- Trier les chemins par ordre décroissant de longueur(t_i) ;
- 3- Appeler la procédure `RestaureACM()` qui appelle itérativement, pour tous les chemins de l'ACM, la procédure `RestaureChemin()` dont la fonction est de retrouver les notions qui composent un chemin de l'ACM.

```

RestaureACM (Entier NumApprenant, Entier NumContexteFormation) {
long :Entier ;
Debut
long :=max (longueur( $t_i$ ))
Tant que long>0 faire {
Pour chaque  $t_i$  tel que longueur ( $t_i$ )=long et PoidsNotionSigné ( $t_i$ ) < 0 faire {
RestaureChemin (path( $t_i$ )) ; }
long :=long-1 ;
}
Fin; }

```

```

RestaureChemin (tableauEntiers path( $t_i$ )) {
Debut
Si longueur( $t_i$ )=1 alors
 $t_i$  et  $t_1$  sont deux sommets adjacents dans l'ACM ;
Sinon {
Pour chaque path( $t_j$ ) tel que longueur( $t_j$ )= longueur( $t_i$ )-1 et PoidsNotionSigné ( $t_j$ ) >0 faire
{
Si Matching(path( $t_j$ ), path( $t_i$ )) alors {
 $t_j$  et  $t_i$  sont deux sommets adjacents dans l'ACM ;
Restore (path( $t_j$ )) ; }
}
}
Fin ; }

```

La procédure `RestaureChemin()` calcule de manière récursive les notions composant un chemin en partant de la notion t_i à l'extrémité du chemin jusqu'à la notion cible t_1 .

5. Apprentissage centré sur l'apprenant

Nous avons voulu concevoir une interface centrée sur l'apprenant où celui-ci est libre de créer, de parcourir ou de supprimer des contextes de formation, d'accéder à des ressources pédagogiques pertinentes pour son contexte de formation (page HTML, documents PDF, simulations, vidéos, etc.) et de passer des tests. Les icônes utilisées pour annoter les nœuds de l'ACM permettent à l'apprenant de distinguer les différents états d'une notion (AccessibleAndMastered, AccessibleAndNotMastered et NotAccessible). Le changement d'état d'une notion s'accompagne d'un changement d'icône qui informe l'apprenant sur sa progression ou sa régression dans le processus d'apprentissage.

Dans la conception de l'interface, nous avons clairement séparé l'ACM (la structure du cours) des ressources pédagogiques (le contenu du cours) et des tests sur les différentes notions de l'ACM. L'exploration de l'ACM, l'accès aux ressources et l'exécution des tests sont des activités d'apprentissage qui représentent des cas d'utilisation du système OrPAF. L'apprenant est libre de choisir l'ordre de ses activités d'apprentissage. Le recours aux tests est nécessaire si l'apprenant désire faire évoluer son modèle et avoir ainsi accès à d'autres notions de l'ACM. C'est l'accessibilité des notions dans l'ACM qui pousse l'apprenant soit à passer des tests pour faire évoluer son modèle et atteindre son objectif soit à reporter ces tests. En effet, il est le seul responsable de son niveau de connaissances et du temps qu'il passe à apprendre. Par exemple, les apprenants avec des contraintes temporelles strictes ont plutôt tendance à aller très tôt vers les tests pour vérifier leur degré de maîtrise des notions de l'ACM et valider leurs connaissances.

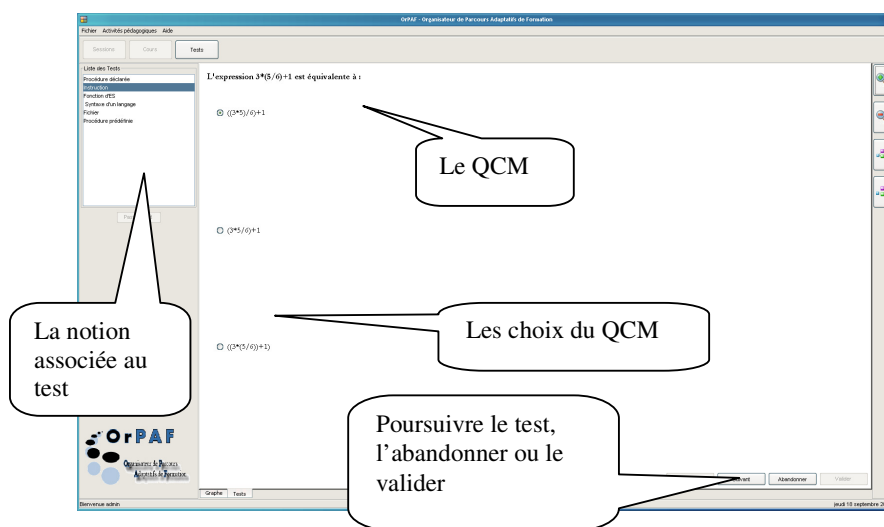


Figure 35 : Tests associés à la notion *Statement* du modèle du domaine de l'algorithmique

6. Base de données du système OrPAF

Dans le système OrPAF, nous avons conçu et construit une base de données qui nous permet de sauvegarder les contextes de formation des différents apprenants et leurs identifiants (login et password). Le schéma de cette base de données est construit en normalisant le graphe des dépendances fonctionnelles suivant (voir figure 36) :

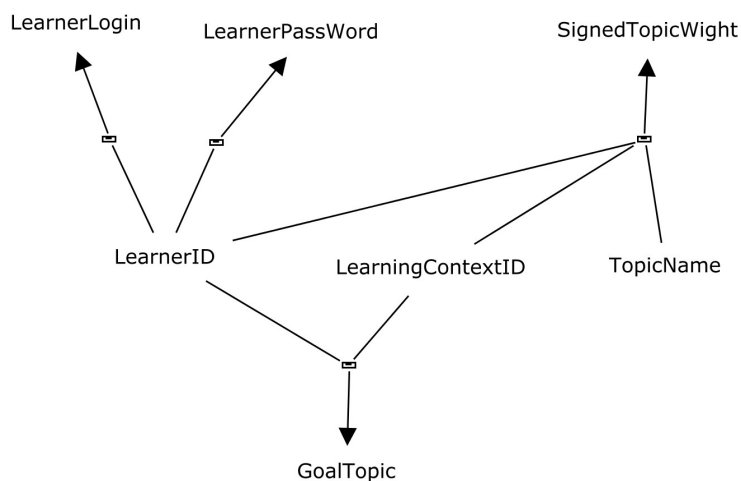


Figure 36 : Le GDF des sessions des apprenants

Après la normalisation du GDF de la figure 36 graphe, nous obtenons les relations suivantes : LearnerSession (LearnerID, LearningContextID, TopicName, SignedTopicWight) qui sauvegarde le poids signé de chaque notion de chaque contexte de formation de chaque apprenant.

LearningContextGoal (LearnerID, LearningContextID, GoalTopic) qui sauvegarde pour chaque contexte de formation, la notion cible qui a servi à le créer.

Learner (LearnerID, LearnerPassWord, LearnerLogin) qui sauvegarde le mot de passe et le login d'un apprenant.

Nous réalisons cette base de données à l'aide du SGBD Access que nous interrogeons et manipulons à l'aide de la passerelle JDBC/ODBC. L'API JDBC permet la manipulation de la base de données à partir des objets java du système OrPAF et le pilote ODBC permet la connexion et l'interaction physiques avec le SGBD Access.

7. Recherche des ressources pédagogiques distantes

Dans le système OrPAF, les enseignants peuvent interroger de manière manuelle ou automatique des corpus distants du Web. Particulièrement, nous avons mis en œuvre

l'interrogation du corpus ARIADNE (le KPS¹⁴ ARIADNE) qui héberge des ressources pédagogiques annotées avec le standard LOM. Le corpus ARIADNE joue également le rôle de portail pour de nombreux autres corpus du Web (EducaNext, Merlot, Prolearn deliverables, etc.). Nous importons dans le composant QueryingDistantResource du système OrPAF le package ARIADNE KPS Client, une API java pour accéder et télécharger des ressources pédagogiques du corpus ARIADNE. Ce package constitue un kit Web services qui rend le protocole SOAP (Simple Object Access Protocol) transparent pour OrPAF, simplifiant ainsi la manipulation des métadonnées LOM.

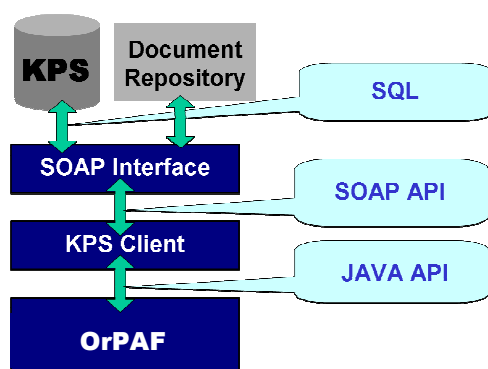


Figure 37 : KPS Client dans l'architecture d'ARIADNE

7.1. Interrogation manuelle du corpus ARIADNE

L'objectif des enseignants est de rechercher des ressources pédagogiques relatives aux notions du domaine de formation visé. Avec le corpus ARIADNE, ils ont la possibilité de construire manuellement des requêtes pour effectuer :

- une recherche simple par mots clés en introduisant des termes qui font référence à des notions du modèle du domaine ;
- une recherche avancée en introduisant la notion qui les intéresse (voir le champ « Main Concept » dans la capture d'écran de l'interface du KPS ARIADNE), le titre de la ressource, l'auteur de la ressource, la langue de la ressource, etc. (voir figure 38) ;
- une recherche fédérée pour des ressources pédagogiques hébergées dans d'autres corpus du Web (voir figure 39).

¹⁴ Knowledge Pool System

Figure 38 : Recherche avancée de ressources pédagogiques dans ARIADNE

Figure 39 : Recherche fédérée de ressources pédagogiques dans ARIADNE

Les enseignants peuvent également rechercher visuellement des ressources pédagogiques dans le corpus ARIADNE. Une classification visuelle des ressources du corpus selon leur domaine d'appartenance permet d'avoir une idée globale sur la répartition des ressources dans le corpus.

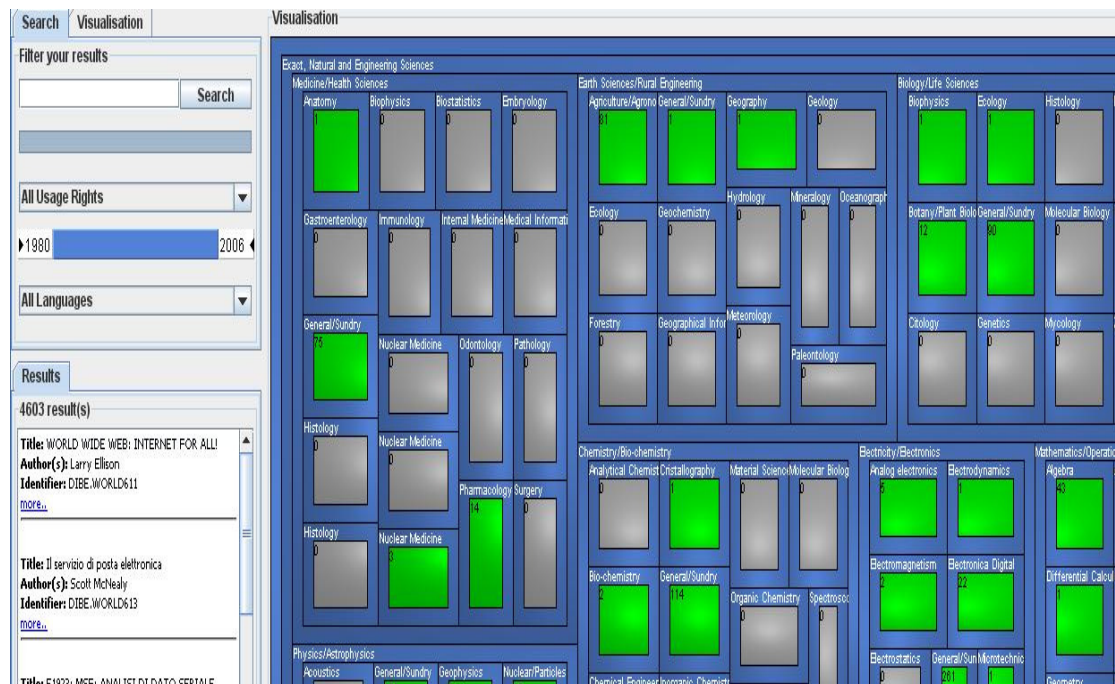


Figure 40 : Classification visuelle des ressources pédagogiques par domaine d'appartenance dans ARIADNE

7.2. Interrogation automatique du corpus ARIADNE

Tous les modes de recherche manuelle de ressources pédagogiques distantes, présentés ci-dessus, sont également possibles par programmation en important l'API KPS Client. Nous avons utilisé cette API dans le composant QueryingDistantResource du système OrPAF (voir figure 26). Ainsi, nous avons construits, par exemple, une requête pour rechercher des ressources en langue française sur une notion contenue dans la variable *topic* comme suit :

```
//Créer un objet requête
QueryClient requeteClient = new QueryClient();

// Définir les critères de la requête avancée : topic et language
AdvancedQueryItem concept = new AdvancedQueryItem (
"ariadneDocument/mainConcepts/mainConcept" , "contains" , topic ) ;

AdvancedQueryItem language = new AdvancedQueryItem (
"ariadneDocument/documentLanguage" , "equals" , "fr" ) ;

//Composer la requête avancée avec la conjonction And
AdvancedQuery requeteAvancee = new AdvancedQueryAnd( concept, language)
;

//Soumettre la requête avancée
requeteClient.advancedQuery (requeteAvancee);
```

8. Recherche des ressources pédagogiques locales

A chaque fois qu'un apprenant veut visualiser des contenus pédagogiques pertinents pour son contexte de formation, c'est-à-dire, des ressources pédagogiques dont la pertinence sémantique SR est supérieure à un seuil prédéfini par les experts du domaine, le composant `QueryingLocalResource` est invoqué. Ce dernier accède aux annotations conceptuelles des ressources pédagogiques du corpus local pour calculer la pertinence sémantique de ces ressources. Nous utilisons l'API de Corese pour rechercher les connaissances nécessaires à ce calcul. Cette API extrait les notions relatives à la ressource pédagogique : les valeurs de la propriété `learningResourceTopic` utilisée dans l'annotation conceptuelle de la ressource pédagogique. Une fois les ressources pédagogiques pertinentes identifiées, elles sont triées par ordre décroissant de pertinence pédagogique de sorte que celles qui s'adaptent le plus avec l'activité pédagogique courante soient affichées en haut de la liste des ressources pédagogiques.

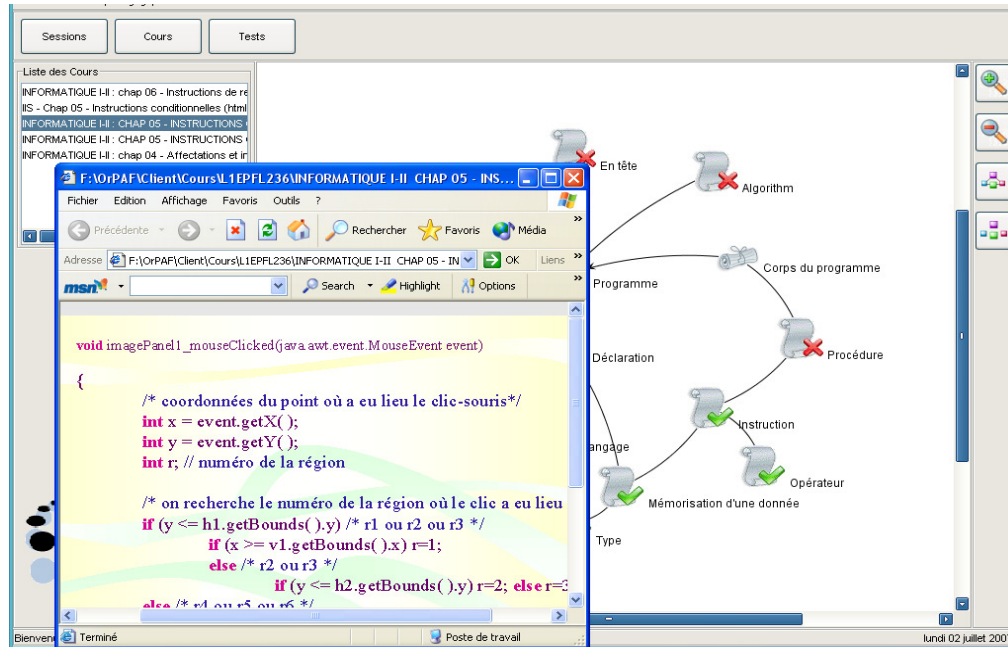


Figure 41: Visualisation d'une ressource pédagogique par l'apprenant

Conclusion : Dans ce chapitre, nous avons décrit la réalisation de l'environnement OrPAF où plusieurs outils et API de développement ont été intégrés pour réaliser les fonctionnalités et les choix conceptuels décrits dans les chapitres 4 et 5. L'environnement OrPAF est développé en java sous Eclipse SDK 3.3.2 et utilise :

- L'API du moteur de recherche Corese pour extraire des connaissances des différents modèles du e-Learning, des annotations des ressources pédagogiques et des des tests ;
- L'API KPS Client pour interroger, accéder et télécharger les ressources pédagogiques du corpus ARIADNE ;
- L'API JDBC et le pilote ODBC pour accéder et manipuler la base de données des sessions d'apprentissage des apprenants ;
- L'API de JUNG pour l'affichage des ACM sous la forme de graphes dynamiques et interactifs.

Dans le chapitre suivant, nous présentons les expérimentations que nous avons effectuées sur l'environnement pédagogique adaptatif OrPAF.

7. Evaluation et retour d'expérience

Nous avons implémenté l'environnement pédagogique adaptatif OrPAF pour permettre à des apprenants d'apprendre des notions du domaine de l'algorithmique et des langages de programmation et nous l'avons évalué avec un groupe d'apprenants de profils différents et des experts du domaine. Ces apprenants sont des étudiants qui ont déjà été dans un cours d'algorithmique. Ce choix s'explique par notre volonté de travailler et d'évaluer le système OrPAF avec des apprenants initiés et autonomes.

Nous avons mené notre expérimentation selon deux directions orthogonales. La première direction vise à expérimenter l'environnement OrPAF sur les aspects utilisabilité et utilité perçues par les apprenants. La deuxième direction a pour objectif l'évaluation de la mesure SR qui calcule la pertinence sémantique des ressources pédagogiques annotées au sein du système OrPAF. Nous avons essayé lors de nos expérimentations d'une part de dégager la proximité de cette mesure mathématique de la perception humaine et d'autre part d'évaluer ses performances sur un corpus réel.

1. Evaluation du système OrPAF

1.1. Test d'utilisabilité de l'environnement par des apprenants

1.1.1. Principe et critères du test d'utilisabilité

Le test d'utilisabilité¹⁵ est une méthode efficace pour évaluer un logiciel car il permet d'observer directement la façon dont l'utilisateur se sert de l'application et ainsi d'identifier concrètement les véritables problèmes qu'il rencontre en situation d'utilisation.

Dans notre cas, il s'agit de mettre l'apprenant en situation effective d'utilisation de l'environnement OrPAF et de lui demander d'acquérir des notions cibles. Ces dernières sont choisies par nos experts de façon à ce que les contextes de formation obtenus ne se chevauchent pas ou très peu. Ceci assure que les notions cibles soient très différentes les unes des autres pour une bonne couverture du modèle du domaine.

L'intérêt du test d'utilisabilité est de déterminer de façon objective l'utilisabilité de l'environnement en mesurant la performance de l'apprenant, par exemple, a-t-il pu atteindre son objectif dans le temps exprimé par ses contraintes temporelles ? Le test d'utilisabilité est aussi et surtout l'occasion de voir l'apprenant en situation et d'observer les problèmes qu'il rencontre, les questions qu'il se pose, ainsi que les fonctionnalités qu'il apprécie ou pas. Nous pouvons ainsi recueillir des éléments précieux sur la façon de rendre l'environnement OrPAF plus facile à utiliser car mieux adapté aux besoins de l'apprenant.

La norme ISO 9241-11 définit l'utilisabilité de la manière suivante : « Un système est utilisable lorsqu'il permet à l'utilisateur de réaliser sa tâche avec efficacité, efficience et satisfaction dans le contexte d'utilisation spécifié ». En d'autres termes, on considère qu'un logiciel est utilisable lorsque l'utilisateur peut réaliser sa tâche (efficacité), qu'il consomme un minimum de ressources pour le faire (efficience) et que le système est agréable à utiliser (satisfaction). Tester l'utilisabilité d'un système consiste donc à valider ces trois critères :

Efficacité : Vérifier que les objectifs visés par l'utilisateur sont atteints.

Efficience : Mesurer les ressources nécessaires pour atteindre ces objectifs, par exemple le temps que l'utilisateur met pour réaliser son objectif.

¹⁵ Test d'utilisabilité (<http://www.wpostal.com/testlogiciel.htm>)

Satisfaction : Déterminer si le système est agréable à utiliser, par exemple, le critère de satisfaction peut être inversement proportionnel au nombre de remarques négatives émises par les utilisateurs.

La norme ISO 9241-11 définit l'utilisabilité sur la base de ces 3 caractéristiques. Il s'agit effectivement des points les plus représentatifs dans le cas général. Cependant, il peut être utile, selon les spécificités de l'application, d'évaluer d'autres aspects du logiciel :

Sécurité : Nombre d'erreurs commises par l'utilisateur et rapidité de correction des erreurs.

Facilité d'apprentissage : Compréhension correcte et assimilation rapide du mode de fonctionnement du logiciel.

Dans le test d'utilisabilité du système OrPAF, nous avons évalué les trois critères suivants :

- l'efficacité de l'apprenant à atteindre son objectif. Nous avons aussi vérifié si le choix du filtre du modèle du domaine influe réellement sur le temps que passe l'apprenant dans son contexte de formation avant d'atteindre son objectif ;
- la satisfaction de l'apprenant de l'interface du système OrPAF ;
- la facilité qu'a l'apprenant à apprendre l'interface du système OrPAF.

1.1.2. Protocole du test d'utilisabilité :

Profils des apprenants:

J. Nielsen a montré qu'un test d'utilisabilité mené avec 5 utilisateurs permet de lever au moins 80 % des problèmes d'utilisabilité. En effet, ce n'est pas en augmentant le nombre d'utilisateurs que l'on trouve plus de problèmes. Les problèmes sont liés au logiciel, pas aux utilisateurs ! Tester avec un plus grand nombre d'utilisateurs augmente le coût du test, pas la pertinence des résultats.

Ainsi, pour tester les trois critères décrits ci-dessus nous avons travaillé avec 15 apprenants qui préparent le diplôme d'ingénieurs d'état en Informatique à l'université 20 Août 1955 (Algérie) et qui appartiennent à deux niveaux d'études différents : 8 étudiants en 3^{ème} année et 7 étudiants en 4^{ème} année. Etant donné que le système OrPAF est destiné principalement à des apprenants avertis et autonomes, nous avons choisi des étudiants initiés avec des niveaux d'études avancés en Informatique. Ces 15 étudiants ont été divisés en 3 groupes de 5 étudiants chacun. Chaque groupe nous sert à évaluer un des trois critères d'utilisabilité.

Matériel utilisé:

Le test d'utilisabilité se passe dans une salle de travaux pratiques équipée de 15 postes de travail connectés à Internet. Chaque apprenant est affecté à un poste de travail sur lequel on a installé l'environnement OrPAF.

Déroulement du test:

En fonction du critère à évaluer, nous avons imposé un déroulement spécifique pour le test d'utilisabilité:

| Critère à évaluer | Déroulement du test |
|--|---|
| Efficacité et temps d'apprentissage de l'apprenant | <p><u>Groupe 1</u></p> <p>Le test se passe comme suit :</p> <ul style="list-style-type: none"> - présentation de l'environnement graphique et de l'intérêt du système OrPAF en utilisant un data show ; - les 5 apprenants explorent pendant une heure 6 contextes de formation sans aucune consigne sur l'ordre d'exploration ; - les 5 apprenants sont invités à créer chacun une ACM simple et à l'explorer pendant 10 minutes ; ensuite une ACM hiérarchique et à l'explorer pendant 20 minutes et enfin une ACM relationnelle et à l'explorer pendant 40 minutes. |
| Satisfaction de l'apprenant | <p><u>Groupe 2</u></p> <p>Le test se passe comme suit :</p> <ul style="list-style-type: none"> - les 5 apprenants sont invités à se documenter sur Internet à propos de notions cibles proposées ; - présentation de l'environnement graphique et de l'intérêt du système OrPAF en utilisant un data show ; - les 5 apprenants sont invités à explorer librement l'environnement OrPAF pour apprendre d'autres notions cibles. |
| Facilité d'apprentissage de l'interface | <p><u>Groupe 3</u></p> <p>Le test se passe comme suit :</p> <ul style="list-style-type: none"> - les apprenants sont invités à utiliser OrPAF sans autres indications que celles fournies par le « help » et le « tutorial » du système. |

Nous avons utilisé une feuille d'observation pour relever les problèmes auxquels ont été confrontés les apprenants et certains usages atypiques de ceux-ci.

A l'issue des trois protocoles de test, nous avons soumis aux trois groupes d'apprenants trois questionnaires :

Groupe 1 :

Avez-vous souvent changé de contexte de formation ?

Avez-vous évolué de manière satisfaisante dans vos contextes de formation ?

Le graphe des notions vous a-t-il permis de comprendre votre évolution vers la notion cible?

Le temps accordé à chaque parcours était-il approprié ?

Existe-il un lien entre la taille du graphe de notions et la complexité de la notion cible ?

Existe-il un lien entre la taille du graphe de notions et le temps nécessaire à parcourir toutes les notions du graphe ?

Groups 2 :

Avez-vous aimé la charte graphique de l'interface ?

L'interface est-elle encombrée ?

Avez-vous été aidés par le système dans les parcours de formation ?

Le graphe des notions est-il facilement explorable ?

L'accès aux tests est-il évident ?

Groupe 3 :

Avez-vous compris le rôle des tests dans OrPAF?

Ya-il un lien entre le graphe de notions et les tests affichés ?

Ya-il un lien entre le graphe de notions et les ressources affichées ?

Avez-vous été gênés par l'exploration par notions du cours?

Avez-vous compris le sens des liens entre les notions ?

Avez-vous compris le sens des icônes représentant les notions ?

1.1.3. Résultats et dépouillement

Les tables 1, 2 et 3 résument les résultats obtenus respectivement du groupe 1, 2 et 3.

Table 3 : Synthèse des résultats du test d'utilisabilité du groupe 1

| Questionnaire : Efficacité et temps d'apprentissage de l'apprenant | Réponses | Nombre d'apprenants | % |
|--|----------|---------------------|-----------|
| Avez-vous souvent changé de contexte de formation ? | Oui | 3 | 60 |
| | Non | 1 | 20 |
| | Un peu | 1 | 20 |
| Avez-vous évolué de manière satisfaisante dans vos contextes de formation ? | Oui | 2 | 40 |
| | Non | 1 | 20 |
| | Un peu | 2 | 40 |
| Le graphe des notions vous a-t-il permis de comprendre votre évolution vers l'objectif ? | Oui | 3 | 60 |
| | Non | 1 | 20 |
| | Un peu | 1 | 20 |
| Le temps accordé à chaque parcours était-il approprié ? | Oui | 4 | 80 |
| | Non | 1 | 20 |
| | Un peu | 0 | 0 |
| Existe-il un lien entre la taille du graphe de notions et la complexité de la notion cible ? | Oui | 0 | 0 |
| | Non | 4 | 80 |
| | Un peu | 1 | 20 |
| Existe-il un lien entre la taille du graphe de notions et le temps nécessaire à parcourir toutes les notions du graphe ? | Oui | 3 | 60 |
| | Non | 0 | 0 |
| | Un peu | 2 | 40 |

Les résultats du questionnaire qui correspond à l'évaluation de l'efficacité et du temps d'apprentissage de l'apprenant (voir table 2) montrent clairement que les apprenants constatent un lien entre le temps nécessaire pour atteindre un objectif et la taille de l'ACM (le graphe des notions). Les résultats montrent également une certaine instabilité des apprenants qui commutent fréquemment entre leurs différents contextes de formation. Cette manière de travailler les empêche d'atteindre efficacement leurs objectifs. Cependant, ils arrivent globalement à assimiler leur évolution dans chaque ACM.

Ces résultats ainsi que nos observations reflètent la difficulté qu'a l'apprenant à gérer parallèlement plusieurs contextes de formation ouverts. Ainsi, pour réduire la désorientation des apprenants, nous avons limité le nombre de contextes de formation dans une session donnée à un maximum de 3 contextes de formation.

Table 4 : Synthèse des résultats du test d'utilisabilité du groupe 2

| Questionnaire : Satisfaction de l'apprenant | Réponses | Nombre d'apprenants | % |
|---|----------|---------------------|-----|
| Avez-vous aimé la charte graphique de l'interface ? | Oui | 3 | 60 |
| | Non | 1 | 20 |
| | Un peu | 1 | 20 |
| L'interface est-elle encombrée ? | Oui | 5 | 100 |
| | Non | 0 | 0 |
| | Un peu | 0 | 0 |
| Avez-vous été aidés par le système dans les parcours de formation ? | Oui | 4 | 60 |
| | Non | 0 | 0 |
| | Un peu | 1 | 40 |
| Le graphe des notions est-il facilement explorable ? | Oui | 5 | 100 |
| | Non | 0 | 0 |
| | Un peu | 0 | 0 |
| L'accès aux tests est-il évident ? | Oui | 2 | 40 |
| | Non | 2 | 40 |
| | Un peu | 1 | 20 |







Les résultats du questionnaire sur la satisfaction des apprenants (voir table 4) montrent que les apprenants ont une prise en main très naturelle des ACM et que le mode d'exploration graphique des notions ne les gêne pas. Aussi, les apprenants ont apprécié pour la plupart l'assistance apporté par OrPAF dans leur parcours d'apprentissage et ceci en comparaison avec la méthode classique qui consiste à chercher des documents sur Internet via un moteur de recherche. Par contre, certains apprenants n'ont pas eu recours aux tests pour rendre accessibles des notions inaccessibles des ACM. Ils n'ont pas compris que les tests leur sont proposés pour les faire évoluer dans leurs contextes de formation. Nous avons récupéré ce problème par l'affichage régulier d'un message invitant l'apprenant à passer les tests si l'ACM contient des notions accessibles et non maîtrisées.

Table 5 : Synthèse des résultats du test d'utilisabilité du groupe 3

| Questionnaire : Facilité d'apprentissage de l'interface | Réponses | Nombre des d'apprenants | % |
|---|----------|-------------------------|-----------|
| Avez-vous compris le rôle des tests dans OrPAF? | Oui | 3 | 60 |
| | Non | 2 | 40 |
| | Un peu | 0 | 0 |
| Ya-il un lien entre le graphe de notions et les tests affichés ? | Oui | 4 | 80 |
| | Non | 1 | 20 |
| | Un peu | 0 | 0 |
| Ya-il un lien entre le graphe de notions et les ressources affichées ? | Oui | 4 | 60 |
| | Non | 0 | 0 |
| | Un peu | 1 | 40 |
| Avez-vous été gênés par l'exploration par notions du cours? | Oui | 0 | 0 |
| | Non | 4 | 80 |
| | Un peu | 1 | 20 |
| Avez-vous compris le sens des liens entre les notions ? | Oui | 1 | 20 |
| | Non | 2 | 40 |
| | Un peu | 2 | 40 |
| Avez-vous compris naturellement le sens des icônes représentant les notions ? | Oui | 2 | 60 |
| | Non | 3 | 40 |
| | Un peu | 0 | 0 |

Les résultats du questionnaire sur l'apprentissage de l'interface (voir table 5) reflètent clairement la compréhension par la plupart des apprenants des liens qui relient les tests et les ressources pédagogiques aux notions de l'ACM. Cependant, ces résultats montrent que la sémantique des liens qui relient les notions d'une ACM ainsi que le sens des icônes utilisées pour les notions d'une ACM ne sont pas correctement compris par les apprenants. Ainsi, nous avons réexaminé les icônes utilisées pour annoter les notions des ACM. En effet, nous avons consulté un ergonome qui nous a aidés à choisir une autre annotation pour représenter les notions qui soit plus facile à comprendre et plus intuitive pour les apprenants (voir la table 6).

Table 6 : Nouvelle annotation des notions d'une ACM

| Etats d'une notion dans une ACM | Ancienne Annotation | Nouvelle Annotation |
|------------------------------------|---|---|
| Notion accessible et maîtrisée |  |  |
| Notion accessible et non maîtrisée |  |  |
| Notion non accessible |  |  |

1.2. Test sur le développement des aptitudes conceptuelles chez les apprenants et la pertinence des ressources pédagogiques ressentie par les apprenants

Nous avons mené un autre test pour évaluer l'impact de l'utilisation du système OrPAF sur le développement des aptitudes conceptuelles des apprenants. L'intérêt de ce test est de vérifier si le mode de transmission des connaissances via un hypermédia adaptatif structuré en une ACM améliore ou pas les connaissances conceptuelles d'un apprenant dans un domaine donné en comparaison avec un cours où la structure est de type hiérarchique (chapitre, section, sous section, etc). Contrairement à une structure hiérarchique, une ACM permet une visibilité globale de la structure du cours. Nous avons également questionné les apprenants sur la pertinence des ressources pédagogiques proposées dans le système OrPAF.

Nous avons travaillé avec 14 apprenants. Nous les avons divisé en deux groupes A et B. Le test s'est déroulé en deux étapes :

Etape 1 : les apprenants du groupe A ont travaillé sur le système OrPAF où nous leur avons demandés de créer deux contextes de formation : une ACM relationnelle avec la notion cible *File* (en français Fichier) et une autre ACM relationnelle avec la notion cible *Procedure* (en français Procédure). Pendant ce temps, nous avons demandé aux apprenants du groupe B de travailler sur des documents numériques recherchés sur Internet pour réaliser les mêmes objectifs : la notion *File* et la notion *Procedure*.

Nous avons accordé aux deux groupes une heure de travail et ensuite nous leur avons faits passer :

- un exercice qui consiste à organiser en graphe un ensemble de notions avec deux relations : la relation de prérequis entre notions et la relation de spécialisation entre notions. Les notions en question étaient déjà présentées

aussi bien dans les documents numériques proposés au groupe B que dans le système OrPAF utilisé par le groupe A ;

- un questionnaire sur la pertinence d'une part des ressources pédagogiques recherchées sur Internet et d'autre part de celles proposées dans le système OrPAF.

Etape 2 : les apprenants du groupe B ont été invité à travailler sur OrPAF pour apprendre deux autres notions cibles : *Array* (en français Tableau) et *Pointer* (en français Pointeur).

Nous avons accordé aux apprenants de ce groupe une heure de travail et ensuite nous leur avons faits passer un autre exercice de conceptualisation et le questionnaire sur la pertinence des ressources pédagogique proposées dans le système OrPAF.

L'histogramme de la figure 42 Synthétise les résultats obtenus.

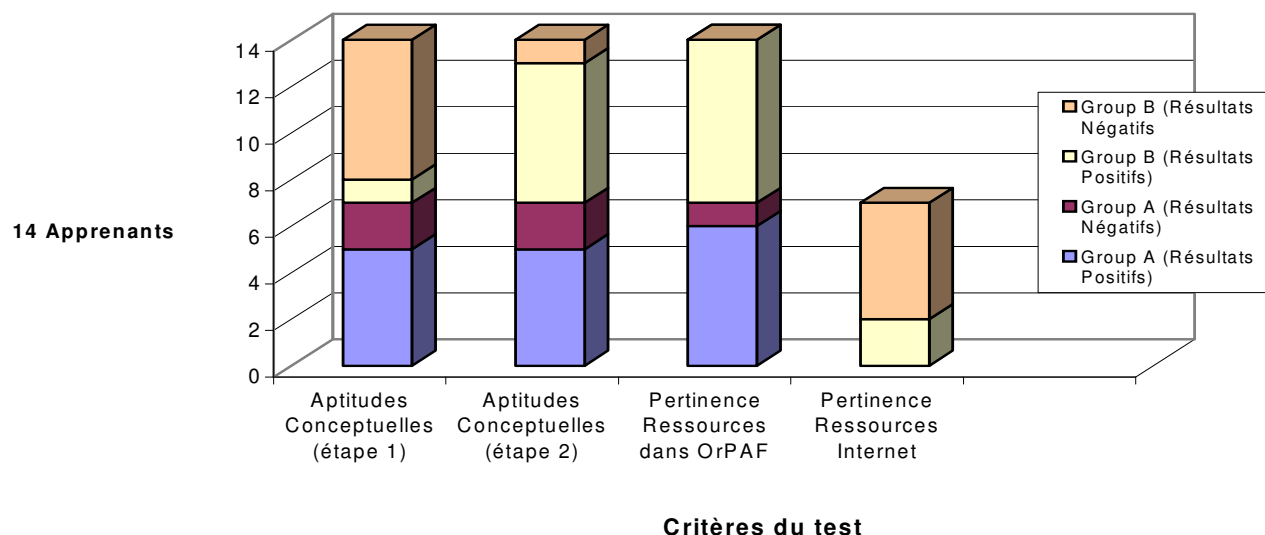


Figure 42: Test des Aptitudes Conceptuelles des Apprenants- Pertinence des Ressources Pédagogiques

L'analyse de cet histogramme nous permet de tirer les conclusions suivantes :

- La plupart des apprenants considèrent que les ressources pédagogiques proposées par le système OrPAF sont pertinentes et adaptées au contexte de la formation. Par contre les ressources pédagogiques recherchées du Web sont considérées peu pertinentes par les apprenants du groupe B. ces derniers ont investis beaucoup de leur temps et de leurs efforts

pour trier les ressources pédagogiques renvoyées par le moteur de recherche. D'ailleurs, nous avons observé une chute importante d'efficacité pour ces apprenants ;

- Le taux de réussite des apprenants du groupe A dans l'exercice de conceptualisation est élevé à l'issue de l'étape 1 alors que les apprenants du groupe B ont pour la plupart échoué dans cet exercice. Ces mêmes étudiants ont réussi l'exercice à l'issue de l'étape 2. Ainsi, nous pouvons conclure que l'utilisation du système OrPAF, où la structure du cours est explicitement et globalement affichée aux apprenants à travers une ACM, développe les aptitudes conceptuelles de l'apprenant, c'est-à-dire, sa capacité à organiser les concepts d'un domaine de connaissances donné.

2. Evaluation de la mesure de calcul de la pertinence sémantique d'une ressource pédagogique

Nous avons adopté deux approches complémentaires pour évaluer la mesure de calcul SR qui sert à calculer la pertinence sémantique d'une ressource pédagogique dans OrPAF.

La première approche consiste à comparer les résultats obtenus par des experts humains et les résultats obtenus par la mesure SR. Nous avons demandé à des experts humains d'annoter un ensemble de ressources pédagogiques et ensuite de juger de la pertinence de chacune de ces ressources pédagogiques pour 6 contextes de formation différents et appartenant au domaine de l'algorithmique et des langages de programmation (3 ACM simples et 3 ACM relationnelles). Chaque expert humain attribue une note de 0 à 10 à chacune des ressources pédagogiques. La note 0 exprime qu'une ressource n'est pas pertinente pour le contexte de la formation et la note 10 exprime qu'elle est pertinente et complètement adaptée au contexte de la formation. Nous étudions la corrélation entre les valeurs moyennes des notes données par les experts et les valeurs obtenues en utilisant la mesure SR et ceci pour chacun des six contextes de formation. Pour pousser la comparaison encore plus loin, nous avons étudié la corrélation entre les données des experts et les valeurs obtenues par la mesure de similarité sémantique entre des graphes conceptuels SoG proposée dans (Zhong et al., 2002). Cette étude nous a permis de vérifier quelle mesure entre SR et SoG est la plus proche de l'évaluation des experts humains. Le choix de la mesure SoG n'est pas aléatoire car cette mesure comporte des similarités avec la mesure SR. A l'instar de SR, elle calcule l'appariement entre deux graphes conceptuels : un graphe requête (the query graph) et un graphe représentant l'annotation conceptuelle d'une ressource (the resource graph). Un autre point commun entre SR et SoG est l'utilisation d'un point de départ pour le calcul de la

similarité : le concept d'entrée (the entry concept) pour SoG et la notion courante pour SR ; Cependant, nous relevons un certain nombre de divergences entre la mesure SR et la mesure SoG :

- SoG dépend de la profondeur des concepts dans la hiérarchie des concepts alors que SR dépend du contexte de la formation, c'est-à-dire, de toutes les relations qui relient les notions du modèle du domaine.
- Le poids d'un concept dans SoG est défini par les utilisateurs de la mesure à travers l'IHM alors que le poids d'une notion dans SR est calculé en fonction des contextes de la formation.

La seconde approche d'évaluation de la mesure SR consiste à mesurer sa performance à extraire des ressources pédagogiques pertinentes et à empêcher l'extraction des ressources inadaptées au contexte de formation.

2.1. Comparaison avec les résultats des experts humains

2.1.1. Données de l'expérience

Nous avons demandé à trois experts de noter sur une échelle de 0 à 10 la pertinence de 20 ressources pédagogiques (20 documents textuels au format PDF) pour six contextes de formation différents (trois ACM simples et trois ACM relationnelles) appartenant au domaine de l'algorithmique et des langages de programmation. Dans chacun des contextes de formation, la notion courante était explicitement représentée.

2.1.2. Déroulement de l'expérience

- Contexte de l'expérience

Afin de mettre les experts dans le contexte de l'expérience, nous leur avons présentés le système OrPAF comme suit:

OrPAF est un système de e-Learning qui a été développé pour permettre à des apprenants d'apprendre de manière efficace (c'est-à-dire apprendre dans un délai minimum). C'est un système qui fournit aux apprenants un parcours d'apprentissage guidé par leurs objectifs. Il suffit que l'apprenant choisisse la notion cible qui l'intéresse pour qu'un parcours d'apprentissage lui soit généré automatiquement. Des ressources sont également fournies à l'apprenant pour lui permettre d'appréhender les notions du parcours. Le parcours est visualisé par l'apprenant sous forme d'un graphe de notions. Ces notions appartiennent au domaine de l'algorithmique et des langages de programmation (par exemple les notions : Instruction, Type de données, Procédure, Opérateur, etc.).

- Objectif de l'expérience :

L'objectif de cette expérience est de collecter les avis des experts sur la pertinence d'une ressource pédagogique pour un contexte de formation matérialisé par un graphe de notions et une notion courante.

- Démarche à suivre :

Pour chaque contexte de formation, nous avons préconisé aux experts les étapes suivantes:

1. Observer le graphe des notions (ACM) où les notions sont représentées différemment:
 - Certaines notions sont accessibles mais ne sont pas encore maîtrisées par l'apprenant,
 - D'autres notions sont accessibles et déjà maîtrisées par l'apprenant,
 - D'autres notions ne sont pas accessibles.

Une notion accessible signifie que l'apprenant peut appréhender cette notion en consultant des ressources pertinentes. En d'autres termes, le niveau de connaissances de l'apprenant lui permet d'apprendre cette notion.

2. Consulter, attentivement, les 20 ressources pédagogiques proposées ;
3. Extraire les notions traitées par chaque ressource pédagogique (Vous devez utiliser les notions décrites dans le modèle du domaine).
4. Noter la pertinence de chaque ressource pour le graphe de notions proposé. La note N doit être comprise entre 0 et 10 avec:
 - N=0 si la ressource n'est pas pertinente;
 - N=10 si la ressource est pertinente.

Une ressource peut être jugée « pertinente » pour l'apprenant même si elle ne concerne pas directement la notion courante du graphe. Elle peut concerner d'autres notions accessibles du graphe. Par contre une ressource qui concerne uniquement des notions absentes ou inaccessibles dans le graphe est forcément impertinente pour l'apprenant.

5. Remplir le tableau (voir table 7).

Table 7 : Tableau à remplir par les experts humains pour annoter et noter 20 ressources pédagogiques

| | Liste des notions de R_i | ACM ₁ | ACM ₂ | ACM ₃ | ACM ₄ | ACM ₅ | ACM ₆ |
|--|----------------------------|--|------------------|------------------|------------------|------------------|------------------|
| La ressource R_i ($1 \leq i \leq 20$) | | Pertinence N de la ressource R_i ($1 \leq N \leq 10$) | | | | | |

2.1.3. Résultats et discussion

La table 8 présente les coefficients de corrélation entre les résultats obtenus par les experts humains et les valeurs de similarité sémantique calculées par les mesures SR et SoG.

La corrélation entre les valeurs de pertinence données par les trois experts et les valeurs de pertinence sémantique calculées par la mesure SR varie entre 0.627 et 0.964 avec une forte convergence entre les résultats des experts et de la SR (de 0.883 à 0.964) pour les ACM simples. Etant donné que le calcul de la SR est le même pour toutes les ACM, nous pouvons expliquer cette convergence par le fait que les ACM relationnelles ont des tailles importantes et peuvent désorienter les experts humains. Les experts ont plus de difficultés à évaluer la pertinence des ressources pédagogiques pour ces ACM que pour les ACM simples. D'ailleurs, les trois experts convergent moins entre eux pour les ACM relationnelles que pour les ACM simples. Nous pouvons conclure que plus la taille de l'ACM est petite plus les évaluations des experts convergent avec les valeurs de pertinence sémantique calculées par la SR. Ces résultats montrent également que le principe de calcul de la mesure SR peut être considéré naturel car très proche des évaluations humaines.

Tandis que les résultats obtenus par la mesure SR sont bonnes et même très bonnes pour les ACM simples, les résultats obtenus par la mesure SoG sont très mauvais et divergent énormément avec les évaluations des experts humains. Ceci pourrait s'expliquer par le fait que la mesure SoG est uniquement adaptée et donnent de bonnes résultats pour les ontologies linguistiques. Nous pouvons également énoncer qu'une mesure de similarité sémantique dépend énormément de son contexte applicatif. La mesure SR n'est pas meilleure ou pire qu'une autre mesure mais simplement elle est bien adaptée au contexte d'utilisation du système OrPAF.

Table 8 : Coefficients de corrélation entre les résultats des experts humains et les mesures de similarité SR et SoG

| Mesures de Similarité | ACM simples | | | ACM relationnelles | | |
|-----------------------|------------------|------------------|------------------|--------------------|------------------|------------------|
| | ACM ₁ | ACM ₂ | ACM ₃ | ACM ₄ | ACM ₅ | ACM ₆ |
| Mesure SR | 0.883 | 0.945 | 0.964 | 0.627 | 0.690 | 0.691 |
| Mesure SoG | -0.124 | 0.008 | -0.678 | -0.495 | 0.003 | -0.097 |

2.2. Performances de la mesure SR d'une ressource pédagogique

2.2.1. Données de l'expérience

Nous travaillons sur un corpus de 36 ressources pédagogiques appartenant au domaine de l'algorithmique et des langages de programmation. Seulement 15 ressources de ces ressources pédagogiques sont considérées pertinentes pour le contexte de formation de la figure 43.

D'un point de vue documentaire, nous pouvons considérer que la recherche locale des ressources pédagogiques est un processus de recherche documentaire basé sur le calcul de la pertinence sémantique. A partir de l'ensemble des réponses obtenus et au regard de l'attente de l'utilisateur nous pouvons mesurer les performances de l'algorithme de recherche mis en œuvre pour retrouver une ressource pédagogique pertinente. Les critères de performances que nous utilisons dans cette expérience sont la précision P (precision), le rappel R (recall) et la F-Mesure F (F-measure) définis comme suit :

Précision :

La précision P est le nombre de ressources pertinentes retrouvées rapportée au nombre de ressources total retourné par l'algorithme de recherche suite une requête donnée.

Le principe est le suivant: quand un utilisateur interroge le corpus, il souhaite que les ressources proposées en réponse à son interrogation correspondent à son attente. Toutes les ressources retournées superflues ou non pertinentes constituent du bruit. La précision s'oppose à ce bruit documentaire. Si elle est élevée, cela signifie que peu de ressources inutiles sont proposées par le système et que ce dernier peut être considéré comme "précis".

On calcule la précision P avec la formule suivante:

$$P = \frac{|\{ \text{ressources pertinentes} \} \cap \{ \text{ressources retournées} \}|}{|\{ \text{ressources retournées} \}|}$$

Rappel :

Le rappel R est défini par le nombre de ressources pertinentes retrouvées au regard du nombre de ressources pertinentes que possède le corpus. Cela signifie que lorsque l'utilisateur interroge le corpus il souhaite voir apparaître toutes les ressources qui pourraient répondre à son besoin d'information. Si cette adéquation entre le questionnement de l'utilisateur et le nombre de documents présentés est importante alors le taux de rappel est élevé. A l'inverse si le système possède de nombreuses ressources intéressantes mais que celles-ci n'apparaissent pas on parle de silence. Le silence s'oppose au rappel.

On calcule le rappel R avec la formule suivante

$$R = \frac{|\{ \text{ressources pertinentes} \} \cap \{ \text{ressources retournées} \}|}{|\{ \text{ressources pertinentes} \}|}$$

F-mesure :

La F-mesure est une mesure qui combine la précision et le rappel. On calcule la F-mesure avec la formule suivante :

$$F = 2 * (P * R) / (P + R)$$

Un système de recherche documentaire parfait fournira des réponses dont la précision et le rappel sont égaux à 1 (l'algorithme trouve la totalité des documents pertinents - rappel - et ne fait aucune erreur - précision). Dans la réalité, les algorithmes de recherche sont plus ou moins précis, et plus ou moins pertinents. Il sera possible d'obtenir un système très précis (par exemple un score de précision de 0,99), mais peu performant (par exemple avec un rappel de 0.10, qui signifiera qu'il n'a trouvé que 10% des réponses possibles). Dans le même ordre d'idée, un algorithme dont le rappel est fort (par exemple 0.99 soit la quasi totalité des ressources pertinentes), mais la précision faible (par exemple 0.10) fournira en guise de réponse de nombreuses ressources erronées en plus de celles pertinentes: il sera donc difficilement exploitable.

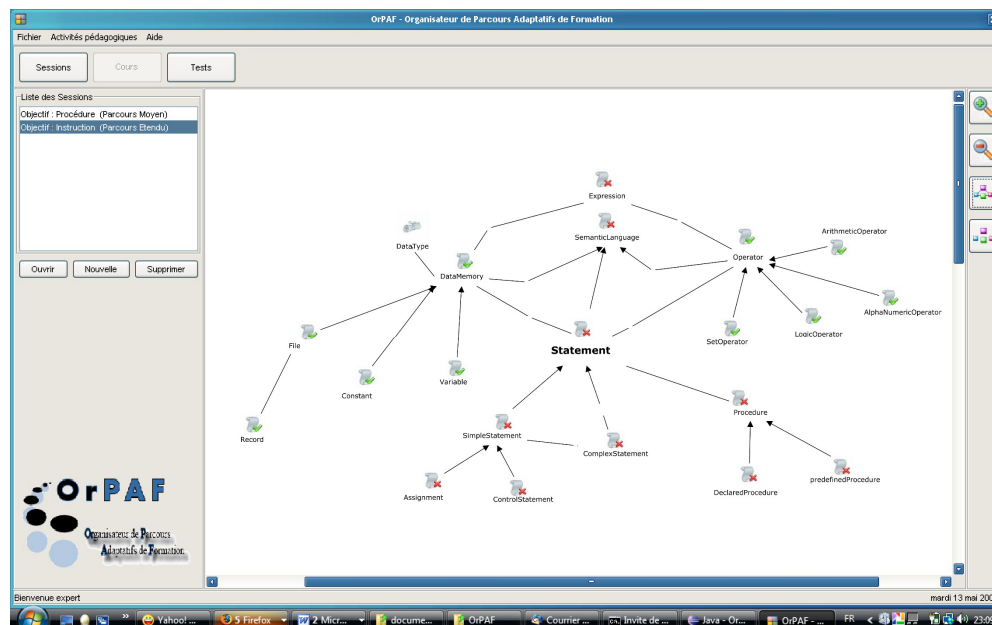


Figure 43: Un contexte de formation du domaine de l'algorithmique

En plus de ces trois critères de performance P, R et F-mesure, nous avons définis un paramètre pour l'expérimentation que nous avons appelé la portée de la requête ou QS (en anglais Query Scope). La QS est le nombre maximum autorisé de ressources retournées suite à l'exécution d'une requête de recherche et représente ainsi le nombre maximum de

ressources pédagogiques accessibles simultanément à l'apprenant. La valeur de ce paramètre est à priori définie par les experts de la pédagogie afin d'éviter la surcharge cognitive de l'apprenant en l'exposant à un nombre trop important de ressources pédagogiques. Notre idée était d'étudier les performances de la mesure SR en variant la QS. Nous avons voulu connaître la courbe de variation des performances en fonction de la QS et il nous a paru utile d'identifier la valeur de QS qui donne les meilleures performances. Sur un plan purement ergonomique, cette valeur nous a permis de définir la taille du container graphique qui sert à afficher dans l'interface du système OrPAF la liste des ressources pédagogiques pertinentes.

2.2.2. Résultats et discussion

Les résultats présentés dans la table 9 montrent les valeurs de la SR pour les 36 ressources pédagogiques du corpus triées par ordre décroissant. Nous pouvons constater que quelques ressources pertinentes (représentées en gras) n'occupent pas le haut du tableau ce qui réduit le rappel de la recherche. Par contre, certains bruits ou ressources pédagogiques non pertinentes occupent le haut du tableau ce qui implique une diminution de la précision de la recherche.

Dans la figure 44, nous pouvons constater que la précision de la SR varie de 66.6% à 84.6% et devient stable (presque constante) lorsque $QS \geq 13$. Ces valeurs expriment le fait que les ressources retournées par la mesure de pertinence sont pour la plupart pertinentes avec des erreurs qui représentent 15.4% à 33.4% des cas.

Nous pouvons également remarquer que la détection des ressources pertinentes augmente et que l'apparition des erreurs diminue avec l'augmentation de la QS. Aussi, la valeur $QS=13$ donne à la fois une bonne précision et un bon rappel pour la mesure SR.

En conclusion, nous pouvons dire que la mesure de similarité sémantique SR est performante car elle constitue un filtre fiable et imperméable pour les ressources non pertinentes.

Table 9: Valeurs de pertinence SR des ressources pédagogiques pour le contexte de formation de la figure 43

| Ressource pédagogique | Annotation conceptuelle (liste des notions) | SR |
|---------------------------------|---|--------------|
| Learning resource 7 | Operator, arithmetic operator, Alphanumeric operator, logic operator | 0.833 |
| Ressource pédagogique 28 | Data type, data memory | 0.444 |
| <i>Ressource pédagogique 15</i> | <i>Data, date type, data declaration,</i> | <i>0.410</i> |
| Ressource pédagogique 5 | Variable, constant | 0.333 |
| Ressource pédagogique 32 | Operator | 0.333 |
| Ressource pédagogique 34 | Data type, set operator | 0.277 |
| Ressource pédagogique 23 | Record, variable | 0.222 |
| <i>Ressource pédagogique 19</i> | <i>File, record, file function</i> | <i>0.205</i> |
| Ressource pédagogique 16 | Logic operator | 0.166 |
| Ressource pédagogique 21 | Set operator | 0.166 |
| Ressource pédagogique 24 | Arithmetic operator | 0.166 |
| Ressource pédagogique 33 | Variable | 0.166 |
| Ressource pédagogique 35 | Constant | 0.166 |
| <i>Ressource pédagogique 11</i> | <i>Choice statement, logic operator</i> | <i>0.148</i> |
| <i>Ressource pédagogique 27</i> | <i>Control statement, logic operator</i> | <i>0.133</i> |
| Ressource pédagogique 9 | Data type | 0.111 |
| Ressource pédagogique 4 | Data type, array | 0.108 |
| <i>Ressource pédagogique 20</i> | <i>Data type, simple data, structured data, array</i> | <i>0.097</i> |
| Ressource pédagogique 31 | Record | 0.055 |
| Ressource pédagogique 36 | Record | 0.055 |
| <i>Ressource pédagogique 1</i> | <i>Algorithm</i> | <i>0</i> |
| <i>Ressource pédagogique 2</i> | <i>Machine code</i> | <i>0</i> |
| <i>Ressource pédagogique 3</i> | <i>Iterative statement</i> | <i>0</i> |
| <i>Ressource pédagogique 6</i> | <i>IO statement</i> | <i>0</i> |
| <i>Ressource pédagogique 8</i> | <i>Assignment statement</i> | <i>0</i> |
| <i>Ressource pédagogique 10</i> | <i>Functional algorithm</i> | <i>0</i> |
| <i>Ressource pédagogique 12</i> | <i>procedure, parameter, local parameter, global parameter</i> | <i>0</i> |
| <i>Ressource pédagogique 13</i> | <i>Algorithm, program, programming language</i> | <i>0</i> |
| <i>Ressource pédagogique 14</i> | <i>Structured program, non structured program, recursive program</i> | <i>0</i> |
| <i>Ressource pédagogique 17</i> | <i>Array sorting</i> | <i>0</i> |
| <i>Ressource pédagogique 18</i> | <i>Function, predefined function</i> | <i>0</i> |
| <i>Ressource pédagogique 22</i> | <i>Statement</i> | <i>0</i> |
| <i>Ressource pédagogique 25</i> | <i>Expression</i> | <i>0</i> |
| <i>Ressource pédagogique 26</i> | <i>Control statement, choice statement</i> | <i>0</i> |
| <i>Ressource pédagogique 29</i> | <i>Semantic language</i> | <i>0</i> |
| <i>Ressource pédagogique 30</i> | <i>Syntax language</i> | <i>0</i> |



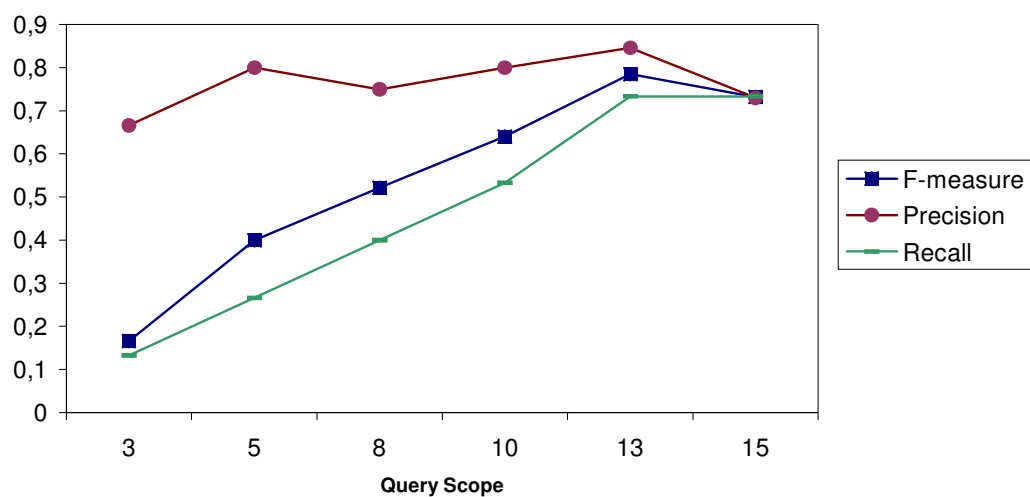


Figure 44: Variation des performances de la mesure SR en fonction du paramètre QS

Conclusion : Dans ce chapitre, nous avons présenté les résultats des expérimentations faites sur l'environnement d'apprentissage personnalisé OrPAF.

D'une part, ces expérimentations avaient pour objectifs de tester l'utilisabilité de l'environnement OrPAF, le développement des aptitudes conceptuelles des apprenants et leur ressenti vis-à-vis des ressources pédagogiques qui leur sont présentés lors du processus d'apprentissage dans OrPAF. Nous avons travaillé sur une ontologie du domaine de l'algorithmique et des langages de programmation avec un groupe d'étudiants initiés dans ce domaine. Les retours obtenus sont satisfaisants :

- La plupart des étudiants ont apprécié l'ergonomie de l'interface d'OrPAF et ont compris son fonctionnement ;
- La structure graphique du cours a permis aux apprenants de mieux conceptualiser le domaine des connaissances ;
- La représentation globale de la structure du cours a permis aux apprenants de mieux comprendre leur évolution vers leurs objectifs ;
- La plupart des apprenants ont su gérer leurs objectifs en fonction de leurs contraintes temporelles puisque ils ont parfaitement compris le lien entre la taille d'une ACM et le temps nécessaire pour la parcourir;
- L'apprentissage guidé par les notions du domaine n'a pas gêné les apprenants qui ont trouvé naturelle cette approche d'exploration de cours.

Néanmoins, certaines lacunes ont été constatées :

- Très peu d'apprenants ont compris l'intérêt des tests proposés dans OrPAF ;
- Le manque de compréhension par la majorité des apprenants des icônes utilisées pour représenter les états des notions d'une ACM ;
- La désorientation des apprenants à cause du nombre élevé des contextes de formation ouverts parallèlement ;

Nous avons essayé de combler ces lacunes par les actions suivantes :

- Recommander régulièrement par des messages aux apprenants de passer des tests sur les notions accessibles et non maîtrisées de l'ACM ;
- Proposer à l'aide d'un ergonome une annotation plus expressive des notions d'une ACM
- Limiter le nombre des ACM que l'apprenant peut créer en parallèle dans sa session.

D'autre part, des expérimentations ont été menées pour évaluer la mesure SR qui calcule la pertinence sémantique d'une ressource pédagogique.

Nous avons d'abord comparé les valeurs obtenues par la mesure SR et les évaluations exprimées par des experts humaines sur un corpus de ressources pédagogiques. Nous avons constaté que la philosophie derrière la mesure mathématique SR est très proche du raisonnement humain (les données étaient fortement corrélées). Nous avons également comparé les évaluations des experts humains avec les données de la mesure SoG (Zhong et al., 2000) sur le même corpus. Les résultats obtenus étaient mauvais et montrent clairement l'inadéquation de la mesure SoG avec le contexte applicatif du système OrPAF. Ceci démontre que la fiabilité des distances ou des liens de similarité entre concepts ou graphes de concepts dépend fortement des applications.

Ensuite, nous avons évalué les performances de la mesure SR en utilisant les critères fournies par le domaine de la recherche documentaire : la Précision, le Rappel et la F-mesure. En effet, nous pouvons considérer la mesure SR comme un algorithme de recherche documentaire qui a pour objectif l'extraction de documents pertinents pour les utilisateurs. Les résultats obtenus ont montré clairement les performances positives de la mesure SR.

Conclusion Générale

Bilan

Gestion des connaissances

Nous avons proposé l'environnement pédagogique adaptatif OrPAF (Organisateur de Parcours Adaptatifs de Formation) qui permet de monter efficacement des formations en s'adaptant avec un minimum d'effort, de coût et de temps à de nouveaux profils d'apprenants, à de nouveaux domaines de connaissances et à de nouvelles stratégies pédagogiques.

Nos choix conceptuels pour la description et la gestion des connaissances dans OrPAF se sont orientés vers le formalisme des ontologies, noyau central et fondateur du Web sémantique. Ces ontologies permettent d'avoir des descriptions partagées et communes et donc réutilisables et aisément adaptables pour répondre à différents besoins de formation.

L'architecture des connaissances que nous avons proposée est organisée à deux niveaux : d'une part, une ontologie du e-Learning qui décrit des concepts et des propriétés génériques et communes à n'importe quel type de formation. D'autre part, des ontologies de formation (modèle apprenant, modèle du domaine et modèle pédagogique) et des annotations conceptuelles spécifiques à une formation donnée.

Proposition d'un environnement pédagogique en ligne

Nous avons développé un environnement pédagogique ouvert sur le Web et particulièrement sur la partie annotée du Web où les ressources pédagogiques sont déjà décrites en utilisant le standard de métadonnées LOM (Learning Object Metadata). Dans le prototype réalisé, un composant de recherche distante interroge le corpus ARIADNE à la recherche de ressources pédagogiques relatives à une formation donnée. Des annotations conceptuelles sont ensuite construites de manière semi-automatique pour décrire plus finement les ressources pédagogiques recherchées du corpus ARIADNE. Ces annotations conceptuellesinstancient des connaissances définies dans l'ontologie du e-Learning et les ontologies de formation. Ces mêmes connaissances sont exploitées par le composant de recherche locale pour identifier les pertinences sémantique et pédagogique d'une ressource pédagogique pour un contexte de formation donné.

Génération de cours hypermédias adaptatifs

La description des connaissances du système OrPAF nous a permis de générer des cours adaptés aux caractéristiques individuelles de l'apprenant. Nous considérons un cours adaptatif comme une entité composée d'une structure adaptative et d'un contenu adaptatif. La structure

adaptative du cours est un graphe de notions (Conceptual Map ou CM) construit par filtrage des notions du modèle du domaine en fonction de l'objectif et des contraintes temporelles de l'apprenant. Les notions de la structure sont annotées en se basant sur le modèle de l'apprenant, i.e. certaines notions sont accessibles et d'autres ne le sont pas. Une CM annotée est appelée ACM (Adaptive Conceptual Map). La structure du cours ainsi construite constitue un moyen de guidage global pour assister les apprenants dans leurs parcours d'apprentissage. Le contenu du cours est un ensemble de ressources pédagogiques recherchées du Web. Ces ressources sont annotées par les enseignants en utilisant les différentes ontologies du système OrPAF. Des mesures de similarité sont proposées pour évaluer les pertinences sémantique et pédagogique de ces ressources pédagogiques pour un contexte de formation donné.

Par ailleurs, l'apprenant évolue dans son parcours d'apprentissage en exécutant librement et à sa guise des tests associés aux notions de l'ACM. La réussite ou l'échec de l'apprenant dans un test donne lieu à une mise à jour de son modèle apprenant et à un changement de l'annotation de l'ACM, i.e. le niveau de connaissances de l'apprenant évolue ou régresse.

Expérimentation

Nous avons expérimenté différents aspects de l'environnement OrPAF :

Un test d'utilisabilité de l'environnement nous a montré une bonne satisfaction et une bonne compréhension de l'interface par les apprenants. Les apprenants n'ont pas été gênés par la structure graphique du cours ni par le mode d'exploration par notions du cours. Les apprenants ont au contraire su se positionner dans leurs parcours d'apprentissage et se construire une structuration conceptuelle du domaine en comparaison avec les apprenants qui ont travaillé sans l'assistance du système OrPAF. Le retour des apprenants sur la pertinence des ressources pédagogiques proposées dans OrPAF est positif car ils ont constaté, pour la plupart, une bonne corrélation entre le contexte de formation et les ressources pédagogiques proposées.

Néanmoins, nous avons relevé un certain nombre de lacunes :

L'apprenant est désorienté lorsque le nombre de contextes de formation ouverts dans sa session est élevé. En réponse à ce problème, nous avons limité à trois le nombre maximal de contextes de formation que l'apprenant a le droit de créer en parallèle dans sa session.

L'apprenant n'a pas recours systématiquement aux tests pour évoluer dans son parcours d'apprentissage. En réponse à ce problème, nous avons implémenté des notifications périodiques à l'apprenant pour lui suggérer de passer les tests.

Nous avons également expérimenté la mesure SR de pertinence sémantique des ressources pédagogiques. Nous avons d'abord demandé à des experts d'estimer la pertinence de ressources pédagogiques pour des contextes de formation choisis. Nous avons étudié la corrélation des estimations des experts avec les données obtenues par la mesure SR et également avec les données obtenues par la mesure SoG (Zhong et al., 2002). Les résultats nous ont clairement montrés une forte similarité entre l'estimation humaine et le principe de calcul qui sous-tend la mesure SR. Cette dernière s'est avérée très adaptée au contexte de l'environnement OrPAF contrairement à la mesure SoG qui donne de bons résultats dans d'autres contextes d'application mais qui n'est pas adaptée pour l'environnement OrPAF.

Ensuite, nous avons étudié la performance de la mesure SR en termes de précision et de rappel. Globalement SR a montré de bonnes performances quand à la recherche de ressources pertinentes.

Perspectives de recherche

Les possibilités d'enrichissement de l'environnement OrPAF sont multiples :

Concernant l'évaluation des apprenant, les tests de type QCM ou Quiz sont considérés comme très peu fiables par les spécialistes en éducation, une amélioration de l'environnement OrPAF serait d'évaluer les apprenants en mesurant la similarité ou la distance sémantique entre les ACM qu'ils construisent et celle construite par un expert (calculée à partir du modèle du domaine).

Une autre perspective serait d'intégrer l'outil TextToOnto dans l'environnement OrPAF. OrPAF offrirait ainsi une interface unique aux experts, aux ingénieurs de la connaissance et aux enseignants qui faciliterait et améliorerait la communication entre les différents acteurs d'OrPAF.

Une autre perspective serait de proposer un système de veille technologique. OrPAF pourrait être utilisé comme système de veille technologique en implémentant un composant de recherche de liens et de ressources susceptibles d'informer les utilisateurs sur les innovations d'un domaine particulier. On rechercherait toutes sortes de ressources et pas simplement des ressources pédagogiques et les ACM contiendraient des notions qui intéressent les profils des utilisateurs.

Références bibliographiques

- ADL (2001). Sharable Content Object Reference Model (SCORM) 2001. www.adlnet.org.
- Allert, H., Dhraief, H., & Nejdil, W. (2002). How are learning Objects Used in learning Processes? Conference on Educational Multimedia, Hypermedia & Telecommunications.
- Aroyo, L., & Dicheva, D. (2004). The New Challenges for E-learning: The Educational Semantic Web. In *Journal of Educational Technology & Society*, Special Issue on Ontologies and the Semantic Web for E-learning, vol. 7(4), pp. 59-69.
- Benyon D. (1993). Adaptive systems: from intelligent tutoring to autonomous agents. www.dcs.napier.ac.uk/~dbenyon/IITpaper.pdf.
- Berners-Lee, T., Hendler, J., Lassila, O. (2001, Mai). The Semantic Web, *Scientific American*, pp. 28-37. <http://www.scientificamerican.com/2001/0501issue/0501berners-lee.html>.
- Boehm, B.W. (1981). *Software Engineering Economics*, Prentice Hall, New Jersey.
- Bolanos-Mora, G., Alvarado-Blanco, J., D'Agostino-Santoro, G. & Segura-Loaiza, M. (1992). Modèle pour contrôler la qualité académique des textes de l'instruction à distance. *Revue de l'enseignement à distance*, 7 (3), pp. 101-120.
- Brown, J.S., Collins, A. et Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, 18 (1), pp. 32-42.
- Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. *User Modeling and User Adapted Interaction*, vol. 6 (2-3), pp. 87-129.
- Brusilovsky, P., Eklund, J., & Schwarz, E. (1998, Avril 14-18). Web-based education for all: A tool for developing adaptive courseware. In *Proceedings of Seventh International World Wide Web Conference, Computer Networks and ISDN Systems*, 30(1-7), 291-300.
- Brusilovsky, P. (2001). "Adaptive and Intelligent Technologies for Web-based Education", <http://www.kuenstliche-Intelligenz.de/archive/>
- Buitelaar, P., Olejnik, D., Sintek, M. (2004, Mai). A Protégé Plug-In for Ontology Extraction from Text Based on Linguistic Analysis In: *Proceedings of the 1st European Semantic Web Symposium (ESWS)*, Heraklion, Greece.
- Chiou, G.-F. (1992). Situated learning, metaphors, and computer-based learning environments. *Educational technology*, 32 (8), pp. 7-11.
- Clancey, W.J. (1991). The frame of reference problem in the design of intelligent machines. K. VanLehn (ed.), *Architectures for intelligence, The twenty-second Carnegie Mellon symposium on cognition*, pp. 357-423. Hillsdale, N.J. : Erlbaum.
- Corby, O., Dieng-Kuntz, R., Faron-Zucker, C., & Gandon, F. (2006). Searching the semantic web: Approximate query processing based on ontologies. *IEEE Intelligent Systems Journal*, 21(1), pp. 20-27.
- Crozat, S. (2002). *Eléments pour la conception industrialisée des supports pédagogiques numériques*, Thèse en informatique de l'Université de Technologie de Compiègne.

- De Bra, P., & Calvi, L. (1998). AHA! An open Adaptive Hypermedia Architecture. *The New Review of Hypermedia and Multimedia*, vol. 4, pp. 115-139.
- De Bra, P., Brusilovsky, P., & Houben, G.-J. (1999). Adaptive hypermedia: From systems to framework. *ACM Computing Surveys*, 31(4).
- De Bra, P., Stash, N., Smits, D., Romero, C., Ventura, S. (2007). Authoring and Management Tools for Adaptive Educational Hypermedia Systems: The AHA! Case Study, in: *Studies in Computational Intelligence (SCI) nr. 62*, pp. 285-308, Springer Verlag.
- Dehors, S., Faron-Zucker, C., & Dieng-Kuntz, R. (2006). Reusing Learning Resources based on Semantic Web Technologies. In *Proceedings of 6th IEEE International Conference on Advanced Learning Technologies, ICALT 2006*, Kerkrade, Netherlands.
- Delestre, N., Pécuchet, J.-P., Barry-Gréboval, C. (2000). Why to use a dynamic adaptive hypermedia for teaching, and how to design it? <http://psimail.insa-rouen.fr/MnemosNet/psipubli.nsf/>
- Deschênes, A.-J. (1991). Autonomie et enseignement à distance. *Revue canadienne pour l'étude de l'éducation des adultes*, 5 (1), pp. 32-54.
- Deschênes, A.J., Bilodeau, H., Bourdages, L., Dionne, M., Gagné, P., Lebel, C., & Rada-Donath, A. (1996). Constructivisme et formation à distance. *Distances*, 1 (1), pp. 9-25.
- Dolog, P., Henze, N., Nedjl, W., & Sintek, M. (2004). "The personal Reader: Personalizing and Enriching Learning Resources using Semantic Web Technologies". *International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*.
- Duval, E. (2002). Normalisation des technologies éducatives : à quoi bon ? *Cinquième Colloque Hypermédiat et apprentissage*.
- Faure D., Nédellec C., & Rouveiro C. (1998). Acquisition of Semantic Knowledge using Machine learning methods: The System ASIUM Technical report number ICS-TR-88-16.
- Frederiksen, C.H., & Donin, J. (1994). Contextes fonctionnels de l'expertise discursive : implications pour l'évaluation. J.-Y. Boyer, J.-P. Dionne et P. Raymond (dir.), *Évaluer le savoir-lire*, pp. 41-67. Montréal : Logiques.
- Galeev, I., Tararina, L., & Kolosov, O. (2004, Août 30-Septembre 1). Adaptation on the basis of the skills overlay model. *Proceedings of IEEE International Conference on Advanced Learning Technologies*, pp. 648-650. Joensuu, Finland: IEEE Computer Society.
- Gandon F. (2002, Novembre 7). *Distributed Artificial Intelligence and Knowledge Management: Ontologies and Multi-Agent Systems for a Corporate Semantic Web*. These PhD. INRIA et University of Nice - Sophia Antipolis
- Gandon, F., Corby, O., Diop, I., Lo, M. (2008). Distances sémantiques dans des applications de gestion d'information utilisant le web sémantique. *Semantic similarity workshop à EGC, Sophia Antipolis*.
- Garrison, D.R. (1993). A cognitive constructivist view of distance education : an analysis of teaching-learning assumptions. *Distance education*, 14 (2), pp. 199-211.
- Grüninger, M., Lee, J. (2002). *Ontology Applications and Design*. *Communication of the ACM*, 45(2), pp. 39-41.

- Halasz, F., Schwartz, M. (1994). The Dexter Hypertext Reference Model. *Communications of the ACM*, 37(2), pp. 30-39.
- Henri, F., & Kaye, A. (1985). *Le savoir à domicile, pédagogie et problématique de la formation à distance*. Québec: Presses de l'Université du Québec, Téléuniversité.
- Henze, N., Nejd, W. (1999). Adaptivity in the KBS Hyperbook System. *Proceedings of the 2nd Workshop on Adaptive System and User Modeling on the WWW*.
- Henze, N., Dolog, P., and Nejd, W. (2004). Reasoning and Ontologies for Personalized E-Learning in the Semantic Web. *Educational Technology & Society*, 7(4), pp. 82-97
- Höök, K. (1997, Mars). Steps to take before IUI becomes real. *The reality of intelligent interface technology*, Edinburgh.
- Horton, P.B., McConney, A.A., Gallo, M., Woods, A.L., Senn, G.J., Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education* 77(1), pp. 95-111.
- IMS (2002). IMS Global Learning Consortium Inc.: IMS Learning Resources Metadata Specification vol. 1.2.1. www.imsproject.org/metadata/index.html
- Jacquinet, G. (1993). Apprivoiser la distance et supprimer l'absence ? ou les défis de la formation à distance. *Revue française de pédagogie*, 102, pp. 55-67.
- Jegede, O. (1992). Constructivist epistemology and its implications for contemporary research in distance learning. T. Evans et P. Juler (eds), *Research in distance education 2* (21-29). Deakin: Institute of distance education.
- Jonassen D., Marra R. (2001, Juin). Epistemological development: an implicit entailment of constructivist learning environments. *ED-MEDIA*, 13th World conference on educational multimedia, hypermedia and telecommunications. Tampere, Finland.
- Jonassen, D.H., Campbell, J.P., & Davidson, M.E. (1994). Learning with media : restructuring the debate. *Educational technology research and development*, 42 (2), pp. 31-39.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension : a construction-integration model. *Psychological review*, 95 (2), pp. 163-182.
- Lave, J. (1990). The culture of acquisition and practice of understanding. J.W. Stigler, R.A. Shweder et G. Herdt (eds), *Essays on comparative human development*, pp. 309-327. Cambridge : Cambridge university press.
- Larochelle, M., & Bednarz, N. (1994). À propos du constructivisme et de l'éducation. *Revue des sciences de l'éducation*, 20 (1), pp. 5-20.
- Lebow, D. (1993). Constructivist values for instructional systems design : five principles toward a new midset. *Educational technology research and development*, 41 (3), pp. 3-16.
- Lester, G., Sim, Y.W., & Wang, C. (2005, Juillet 5 - 8). An e-Learning Systems Engineering Methodology. *The 5th IEEE International Conference on Advanced Learning Technologies*, Kaohsiung, Taiwan.

- LOM (2002). LOM IEEE Learning Technology Standard Committee (LTSC), Learning Object metadata Group. <http://ltsc.ieee.org/wg12/>
- Maedche, A., & Staab, S. (2000). Semi-automatic engineering of ontologies from text. In Proceedings of the 12th International Conference on Software and Knowledge Engineering .
- Mizoguchi, R., & Ikeda, M. (1996). Towards Ontological Engineering (AI-TR-96-1). Osaka: ISIR, Osaka.
- Mizoguchi, R., & Bourdeau, J. (2000). Using Ontological Engineering to Overcome Common AI-ED Problems. In International Journal of Artificial Intelligence and Education, vol.11 (Special Issue on AIED 2010), pp. 107-121.
- Moore, M. (1977). A model of independent study. *Epistologidaktika*, Vol. 1, pp. 6-40.
- Moore, M., & Kearsley, G. (1996). *Distance Education. A Systems View*. Wadsworth: Belmont.
- Navigli R., Velardi P., Gangemi A. (2003, Janvier-Février). Ontology Learning and its application to automated terminology translation *IEEE Intelligent Systems*, 18(1).
- Paquette, G. (2002). *L'ingénierie pédagogique : pour construire l'apprentissage en réseau*. Québec: Presses de l'Université du Québec.
- Pépin, Y. (1994). Savoirs pratiques et savoirs scolaires: une représentation constructiviste de l'éducation. *Revue des sciences de l'éducation*, 20 (1), pp. 63-86.
- Peraya, D., & McCluskey, A. (1996). Enseignement à distance (EAD) au niveau universitaire. Genève, Berne : Université de Genève, TECFA, Office fédéral pour l'éducation et la science.
- Perkins, D. (1991). Technology meets constructivism: Do they make a marriage? *Educational Technology*, Vol. 31, pp. 18-23.
- Piaget, J. (1970). Piaget's theory. P. H. Mussen (Ed). *Carmichael's manual of child psychology*, 1(3), pp. 703-732. New York: wiley.
- Psyché, V., Mendes, O., & Bourdeau, J. (2003). Apport de l'ingénierie ontologique aux environnements de formation à distance. *STICEF*, Vol. 10, ISSN : 1764-7223.
- Reigeluth, C.M. (1999). *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*. Lawrence Erlbaum Associates, Mahwah NJ.
- Rhéaume, J. (1991). *Hypermédiat et stratégies pédagogiques*. Premier colloque Hypermédiat et Apprentissages.
- Russ, T., Valente, A., MacGregor, R., & Swartout, W. (1999). Experiences in Trading Off Ontology Usability and Reusability. The Twelfth Banff Knowledge Acquisition for Knowledge based Systems Workshop, Alberta, Canada.
- Schneider, D. (2003). Conception et implémentation de scénarios pédagogiques riches avec des portails communautaires. Second Colloque de Guéret, Faculté de Psychologie et des Sciences de l'éducation.
- Schubauer-Leoni, M.L., & Ntamakiliro, L. (1994). La construction de réponses à des problèmes impossibles. *Revue des sciences de l'éducation*, 20(1), pp. 87-113.
- Sintek, M., Decker, S. (2002). TRIPLE – an rdf query, inference, and transformation language. Horrocks, I. and Hendler, J. edidors, *Int. Semantic Web Conference (ISWC)*, pp. 364-378.

- Sowa, J. (1984). *Conceptual Structures: information processing in mind and machine*. The System Programming Series, Addison Wesley.
- Staab, S., Maedche, A. (2001). *Axioms are Objects, too – Ontology Engineering beyond the Modeling of Concepts and Relations*, In Internal Report 399, Institute AIFB, Karlsruhe University.
- Stojanovic, L., Staab, S., & Studer R. (2001). *E-Learning based on the Semantic Web*. WebNet2001, World Conference on the WWW and Internet, October 23-27, Orlando, Florida - USA.
- Tardif, J. (1998). *Intégrer les nouvelles technologies de l'information : quel cadre pédagogique ?* Paris: ESF, 1998.
- Tochon, F. (1992). *Didactique du français : de la planification à ses organisateurs cognitifs*. Paris : ESF, 1990.
- Uschold, M., & Gruninger, M. (1996, Juin). *Ontologies : Principales, Methodes and Applications* Knowledge, Engineering Review, 11(2).
- Vargas-Vera, M., Motta, E., Domingue, J., Lanzoni, M., Stutt, A., & Ciravegna, F. (2002). *MnM: Ontology Driven Semi-Automatic and Automatic Support for Semantic Markup*. Proceedings of the 13th International Conference on Knowledge Engineering and Management (EKAW 2002), pp. 379-391, Spain.
- Vargas-Vera, M., & Lytras, M.D. (2008). *Personalized Learning Using Ontologies and Semantic Web Technologies*. WSKS (1) pp. 177-186
- Vassileva, J. (1997). *Dynamic Courseware Generation on the www*. Proceedings of the workshop: Adaptive Systems and User Modeling on the World Wide Web. Sixth International Conference On User Modeling.
- Wagner, E.D., & McCombs, B.L. (1995). *Learner centered psychological principles in practice : Designs for distance education*. Educational technology, 35(2), pp. 32-35.
- Yessad, A., & Laskri, M.T. (2005a, 9-11 Mai). *Construction d'Hypermédias adaptatifs Ouverts : Représentation et Recherche des Ressources Pédagogiques à partir du Web Sémantique*. In Proceedings of ISPS, Alger, Algérie.
- Yessad, A., Hérin, D., & Laskri, M.T. (2005b, Mai 27-29). *Utilisation du Web Sémantique pour la construction d'Hypermédias Educationnels Adaptatifs*. Environnement Informatique pour l'Apprentissage Humain (EIAH), Montpellier, France.
- Yessad, A., & Laskri, M.T. (2006, Decembre 7-9). *Ontology-driven dynamic course generation for web-based education*. Proceedings of MCSEAI, Agadir, Morocco.
- Yessad, A., Faron-Zucker, C., Dieng-Kuntz, R., Laskri, M.T. (2008a, Juin 30-Juillet 4). *Adaptive Course Generation in Semantic Web Context*. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Vienna, Austria,
- Yessad, A., Faron-Zucker, C., Dieng-Kuntz, R., & Laskri, M.T. (2008b, Octobre-Décembre). *Adaptive Learning Organizer for Web-based Education*. International Journal of Web-based Learning and Teaching Technologies, 3(4), pp. 57-73.
- Yessad, A., Faron-Zucker, C., & Laskri, M.T. (2008c, Novembre). *Ontology-based Personalization of Hypermedia Courses*. International Review on Computers and Software, to appear.

- Wang, Y., Völker, J., Haase, P. (2006, Octobre). Towards Semi-automatic Ontology Building Supported by Large-scale Knowledge Acquisition. AAAI Fall Symposium On Semantic Web for Collaborative Knowledge Acquisition, vol. FS-06-06, pp. 70-77. AAAI, AAAI Press, Arlington, VA, USA.
- Zhong, J., Zhu, H., Li, J., & Yu, Y. (2002, Juillet 15-19). Conceptual graph matching for semantic search. In Proceedings of 10th International Conference on Conceptual Structures, ICCS, pp. 92-106. LNCS 2393, Springer Verlag, Borovets, Bulgaria.