

# Un Algorithme d'Appariement des Compétences pour l'Affectation des Modules d'un Parcours Informatique aux Enseignants

Imane Boughedda<sup>1</sup>, Samia Djamai<sup>2</sup>, Saloua Chettibi<sup>3</sup>

<sup>1,2</sup> Département d'Informatique, Université de Jijel.

<sup>3</sup> Equipe SCAL, Laboratoire MISC, Université Mentouri Constantine.

E-mail : <sup>3</sup>sa.chettibi@yahoo.com

## Résumé

Dans l'objectif d'automatiser la procédure d'affectation des modules aux enseignants dans un parcours de spécialité informatique, nous proposons une approche qui combine l'annotation sémantique basée ontologie des CVs et des descriptifs des modules avec un algorithme d'appariement entre compétences acquises par les enseignants et de celles requises pour enseigner un module. En effet, ces compétences (acquises et requises) correspondent à des concepts de l'ontologie du domaine Informatique dont on peut calculer la similarité. L'algorithme proposé exploite essentiellement les mesures de Wu&Palmer et de Dice.

Les ontologies construites avec l'algorithme proposé sont implémentés au sein du système SyGAM (Système de Gestion de l'Affectation des Modules) que nous avons développé.

**Mots clés :** Affectation des modules, ontologie, compétence, annotation sémantique, appariement de concepts, mesure de Wu&Palmer, mesure de Dice.

## 1. Introduction

Le terme ontologie trouve son premier sens en philosophie où l'ontologie est l'étude des propriétés générales de ce qui existe. En informatique, d'après Borst [1], une ontologie est une spécification formelle et explicite d'une conceptualisation partagée. Cette définition désigne qu'une ontologie repose sur des définitions explicites de concepts d'un domaine et de leurs relations. Ces définitions doivent être partagées et acceptées par les différents utilisateurs et exploitables par des machines. Les ontologies ont des immenses possibilités d'utilisations et dans des domaines très variés, on cite entre autre: la communication, l'aide à la spécification des systèmes d'information, l'indexation et la recherche d'information, etc.

Le travail présenté dans cet article est particulièrement basé sur les ontologies et sur le calcul de similarité entre concepts d'ontologie pour traiter le problème d'affectation des modules aux

enseignants dans un parcours de spécialité informatique. Souvent les deux critères de grade et d'ancienneté de l'enseignant sont utilisés pour surmonter les difficultés de prise de décision qui se manifestent, notamment, quand plusieurs enseignants veulent se charger d'un même module, quand il y a des nouveaux enseignants recrutés ou même quand des nouvelles options de formation sont ouvertes. Toutefois, les deux critères évoqués ci-avant restent insuffisants pour assurer l'affectation optimale des modules en termes des compétences ce qui risque de diminuer le rendement pédagogique.

Ce travail a pour objectif l'automatisation de la procédure d'affectation des modules aux enseignants tout en tenant compte des compétences/spécialités de ces derniers mais aussi de leurs préférences (exprimées sous forme de fiches de vœux), ça d'une part, et d'autre part, des compétences requises pour chaque module.

Le reste de cet article est organisé en cinq sections. Comme nous nous intéressons au calcul de similarité entre concepts d'ontologie, nous nous focalisons sur les mesures de similarité dans la section 2. La section 3 est réservée à la description des ontologies utilisées dans notre travail, à savoir : l'ontologie «Compétences\_Enseignant» et l'ontologie «OntInformatique». La section 4 présente l'algorithme d'appariement proposé suivi par une étude de cas illustratif. Nous donnons un aperçu du système de gestion de l'affectation des modules que nous avons développé dans la section 5. Enfin, la section 6 conclut l'article.

## 2. Calcul de Similarité

La similarité [2] se rapporte à la comparaison des éléments d'ontologies. Elle renvoie une valeur numérique indiquant si les deux éléments ont un degré élevé ou bas de similitude. Formellement, une mesure de similarité [3]  $\sigma : O \times O \rightarrow R$  est une fonction qui exprime la similarité entre deux entités tel que :

$$-\forall x, y \in O, \sigma(x, y) \geq 0.$$

$$-x \in O, \forall y, z \in O, \sigma(x, x) \geq \sigma(y, z)$$

$$-\forall x, y \in O, \sigma(x, y) = \sigma(y, x).$$

Une similarité  $d$  est dite *sémantique* si et seulement si elle vérifie les propriétés suivantes [4] :

- si  $A \equiv B$  alors  $d(A, B) = 0$ .
- si  $A \subseteq B \subseteq C$  alors  $d(A, B) \leq d(A, C)$ .
- si  $A \cap B \subseteq \perp$  alors  $d(A, B) = \infty$ .

Une telle distance sémantique doit être *mathématique* en vérifiant les axiomes suivants:

- $d(A, B) = 0 \Leftrightarrow A = B$ .
- $d(A, B) = d(B, A)$ .
- $d(A, C) \leq d(A, B) + d(B, C)$ .

### 2.1. Travaux existants

Parmi les techniques de calcul de similarité existantes dans la littérature nous nous intéressons, dans ce travail, à deux approches ; à savoir : l'approche basée sur les arcs et l'approche basée sur l'espace vectoriel.

#### 2.1.1 Approche basée sur les arcs

Ce type de mesure s'appuie sur la structure de la ressource sémantique en proposant un comptage plus ou moins élaboré du nombre d'arcs séparant deux concepts. Ces mesures se servent de la structure hiérarchique de l'ontologie pour déterminer la similarité sémantique entre les concepts. Parmi les travaux classifiés sous cette bannière on peut citer : la mesure de Wu-Plamer [5], la mesure de Zargayouna [6] et la mesure de Resnik [7].

#### 2.1.2 Approche basée sur l'espace vectoriel

Cette approche utilise un vecteur caractéristique, dans un espace dimensionnel, pour représenter chaque objet. Le modèle de l'espace vectoriel est employé pour un arrangement des objets complexes en les représentant comme des vecteurs de  $k$ -dimensions. La définition de la similarité entre deux vecteurs d'objets est obtenue par leurs contenus internes [8]. Parmi les mesures appartenant à cette famille on trouve [9] : la mesure de Jaccard, la mesure de Cosine, la mesure Euclidienne et la mesure de Dice.

### 2.2 Les mesures utilisées

Nous nous limitons dans cette section à la description des mesures que nous avons exploitées, à savoir: Wu-Palmer et Dice.

#### 2.2.1 La mesure de Wu-Palmer

Dans une ontologie, le principe pour calculer la similarité de Wu-Palmer [5] est basé sur les distances  $D1$  et  $D2$  qui séparent les concepts  $C1$  et  $C2$  du concept subsumant (le père le plus spécifique) (PPS) et la distance  $D$  qui sépare PPS au concept racine.

La similarité entre  $C1$  et  $C2$  est donnée par:

$$Sim_{WuPalmer}(C1, C2) = \frac{2 * D}{D1 + D2 + 2 * D}$$

#### 2.2.2 Similarité de Dice

La similarité de Dice est définie par le nombre des objets communs multipliés par 2 sur le nombre total d'objets. La mesure de Dice est donc définie par la formule suivante :

$$Sim_{Dice}(x, y) = \frac{2xy}{\|x\|_2^2 + \|y\|_2^2}$$

### 3. Les Ontologies Construites

Certes les ontologies représentent un modèle conceptuel sémantiquement riche offrant un cadre de représentation de connaissances très expressif, mais la conception des ontologies reste une tâche fastidieuse et de longue haleine. En matière de méthodologies de construction des ontologies, on distingue deux grandes classes : la première classe compte les méthodes qui construisent des ontologies à « partir de zéro » ou par réutilisation d'autres ontologies où la conception est réalisée manuellement ; La deuxième classe englobe les méthodes de conception des ontologies à l'aide des outils de traitement de langue naturelle (construction semi-automatique). Bien qu'aucune méthode ne soit reconnue comme standard, la méthode METHONTOLOGIE [10] est considérée parmi les méthodes les plus complètes.

En se guidant par les étapes de METHONTOLOGIE, nous avons construit deux ontologies ; à savoir : l'ontologie « Compétences\_Enseignant » et l'ontologie « OntInformatique ». On s'est servi de la première ontologie pour modéliser les compétences d'un enseignant en termes de : modules enseignés, mémoires et thèses soutenus ou encadrés, etc. (voir Fig. 1 et Fig. 2).

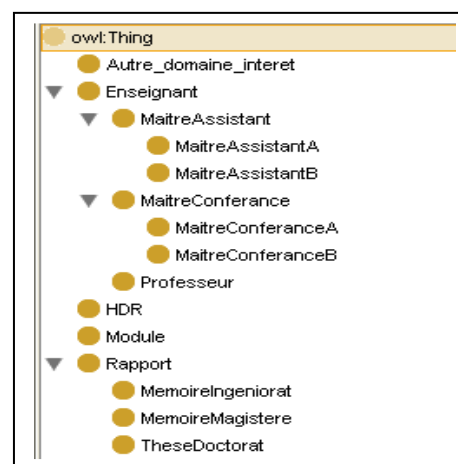


Figure 1- Hiérarchie des concepts de l'ontologie Compétences-Enseignant

La deuxième ontologie modélise les différents domaines de la discipline Informatique (voir Fig. 4). Dans la conception de cette ontologie, nous avons passé tout d'abord par une ontologie générique (voir Fig. 3) qui récapitule les différentes relations qui

peuvent exister entre différents domaines de connaissance indépendamment d'un contexte particulier. L'ontologie du domaine informatique, nous a servi de référentiel pour l'appariement des compétences acquises des enseignants et de celles requises pour enseigner un module.

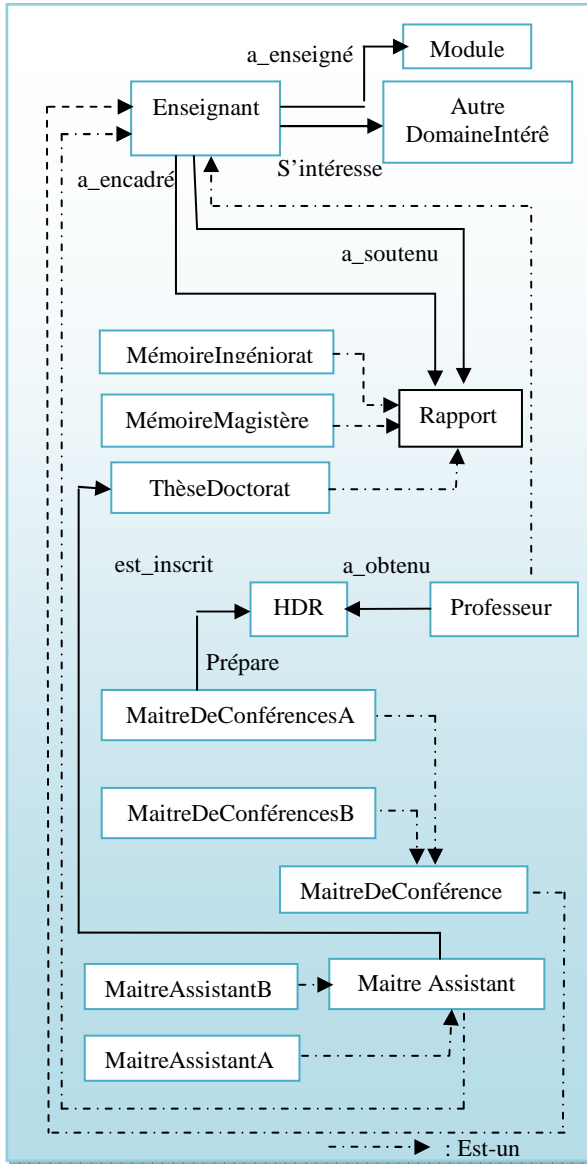


Figure 2- Diagramme de relations binaires

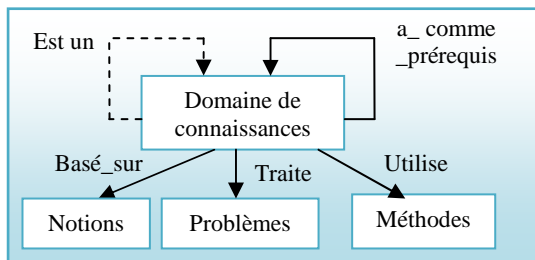


Figure 3- Ontologie Générique



Figure 4- Hiérarchie des concepts de l'ontologie Ont-Informatique

#### 4. L'Algorithme d'Appariement

Notre objectif principal est l'affectation d'un module à l'enseignant le plus adéquat en termes de ses compétences. Pour ce faire, nous proposons un algorithme qui combine la distance de Wu-Palmer avec celle de Dice. Cette combinaison est naturellement justifiée par le fait que l'ontologie du domaine informatique construite contient des relations de subsumption «est-un» et d'autres associatives «basé\_sur, utilise, a\_comme\_pré-requis, traite». Dans cette combinaison, les distances de Wu-Palmer et de Dice exploitent les relations de subsumption et les relations associatives, respectivement, dans le calcul de similarité entre les compétences d'un enseignant et les compétences requises d'un module.

Par compétences requises nous désignons tous les concepts reliés avec le concept «Module» dans l'ontologie OntInformatique.

En effet, nous avons choisi la distance de Wu-Palmer pour sa simplicité et pour les performances qu'elle présente tout en restant aussi expressive que les autres distances basées sur les arcs. Le choix de Dice est justifié par son adaptation aux cas des concepts complexes.

A part le calcul de similarité, nous avons fait recours également au principe de pondération en associant à chaque classe des compétences acquises un coefficient reflétant son importance dans l'affectation (voir Tab.1). De même, nous avons attribué plus d'importance aux compétences requises issues des relations «basé\_sur», «traite» et

« utilise » par rapport à celles issues de la relation de subsomption. Le poids le plus faible a été donné aux compétences venant des relations et « a\_comme\_pré-requis » (voir Tab.2).

**Table 1-** Les poids de chaque classe de compétence

Compétences acquises	Poids
Modules enseignés	0,5
Thèse de doctorat	0,5
Mémoire magistère	0,4
Mémoire ingéniorat	0,3
Travail d'encadrement	0,2
Autre domaine d'intérêt	0,2

**Table 2-** Les poids associés à chaque relation

Relation	Poids
Basé_sur	0,5
Utilise	0,5
traite	0,5
Est-un	0,4
A_comme_prérequis	0,3

Avant de donner l'algorithme d'appariement des compétences, nous définissons la distance de Wu\_Palmer comme suit :

**Distance\_WPalmer** (concept1, concept2)  
 $D \leftarrow 0$ ;  $D1 \leftarrow 0$ ;  $D2 \leftarrow 0$ ;  
 $PPS \leftarrow$  chercher\_père (concept1, concept2) ;  
 $D \leftarrow$  distance (racine, PPS) ;  
 $D1 \leftarrow$  distance (concept1, PPS) ;  
 $D2 \leftarrow$  distance (concept2, PPS)  
 $Distance\_Palmer = (2*d)/(d1+d2+(2*d))$  ;

**Figure 5-** La fonction calculant la distance de Wu-Palmer

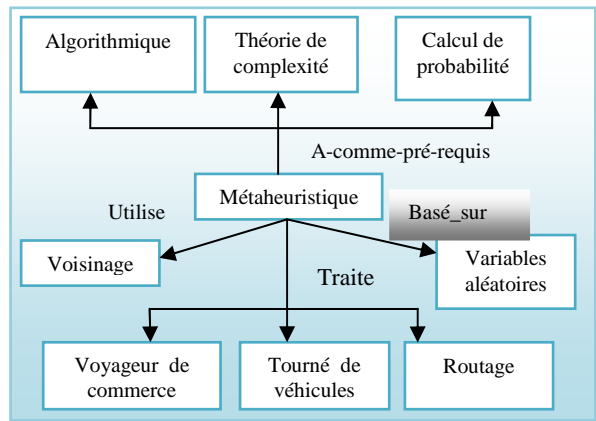
**Algorithme d'appariement ;**

*M* : Le module à affecter.  
*N* : Nombre des enseignants demandant le même module.  
*L* : Nombre des classes de compétences.  
*Nb\_R* : Nombre des compétences requises d'un module *M*.  
*Nb\_PR* : Nombre des Compétences Pré-requis d'un module *M*.  
 $C[i,j,k]$  : compétence *k* de l'enseignant *i* dans la classe des compétences *j*.  
 $Nb\_C[i,j]$  : Nombre des compétences de l'enseignant *i* dans la classe des compétences *j*.  
 $P_j$  : poids de la classe de compétence *j*.  
*Trouve* : booléen mis à vrai si la compétence est un requis.  
 $Sim_1$  : Nombre des compétences acquises de la classe *K* de l'enseignant *j* qui coïncident avec les compétences requises du module issues de la relation: traite, utilise, basé\_sur.  
 $Sim_2$  : Résultat de la fonction *Distance\_WPalmer*.  
 $Sim_3$  : Nombre des compétences acquises de la classe *K* de l'enseignant *j* qui coïncident avec les compétences requises du module issues de la relation a\_comme\_prérequis  
**Début**  
**Pour**  $i=1$  à *N* **faire**  
 $Sim_1 \leftarrow 0$ ;  $Sim_2 \leftarrow 0$ ;  $Sim_3 \leftarrow 0$ ;  
**Pour**  $j=1$  à *L* **faire**  
**Pour**  $k=1$  à  $Nb\_C[i,j]$  **faire**  
*Trouve*  $\leftarrow$  faux  
**Pour**  $l=1$  à  $Nb\_R$  **faire**  
**Si**  $C[i,j,k] = R[l]$  **alors**  
**Début**  
 $Sim_1 \leftarrow sim_1 + 1$  ;  
 $l = Nb\_R + 1$ ;  
*Trouve*  $\leftarrow$  vrai ;

**Fin**  
**Fin\_si**  
**Fin\_pour**  
**Si not ( trouve ) alors**  
**Si**  $Distance\_WPalmer (C[i,j,k] , M) < 0$  **alors**  
 $Sim_2 \leftarrow sim_2 + Distance\_Palmer (C[i,j,k] , M)$ ;  
**Sinon**  
**Pour**  $o=1$  à  $Nb\_PR$  **faire**  
**Si**  $C[i,j,k] = PR[o]$  **alors**  
**Début**  
 $Sim_3 \leftarrow sim_3 + 1$  ;  
 $o = Nb\_PR + 1$ ;  
**Fin**  
**Fin\_si**  
**Fin\_pour**  
**Fin\_si**  
**Fin\_si**  
**Fin\_pour**  
 $Sim [i,j] \leftarrow 0.5 * (Sim_1*2 / Nb\_R + Nb\_C[i,j])$   
 $+ 0.4 * Sim_2 + 0.3 * (Sim_3 / Nb\_PR + Nb\_C[i,j])$   
**Fin\_pour**  
**Fin\_pour**  
 $Enseignant\_choisi \leftarrow Argmax \{ \sum_{j=1}^L P_j * Sim [i,j] \}$   
**Fin**

**Figure 6-** L'algorithme d'appariement.

**Etude de cas-** Pour montrer l'efficacité du notre algorithme, nous présentons un exemple qui concrétise notre proposition. Supposons que le module à affecter est le module « Métaheuristique » (voir Fig. 7) et que deux enseignants possédants des compétences différentes veulent s'en charger (voir les Fig.8 et Fig.9).



**Figure 7-** Extrait de l'ontologie OntInformatique décrivant le module Métaheuristique.

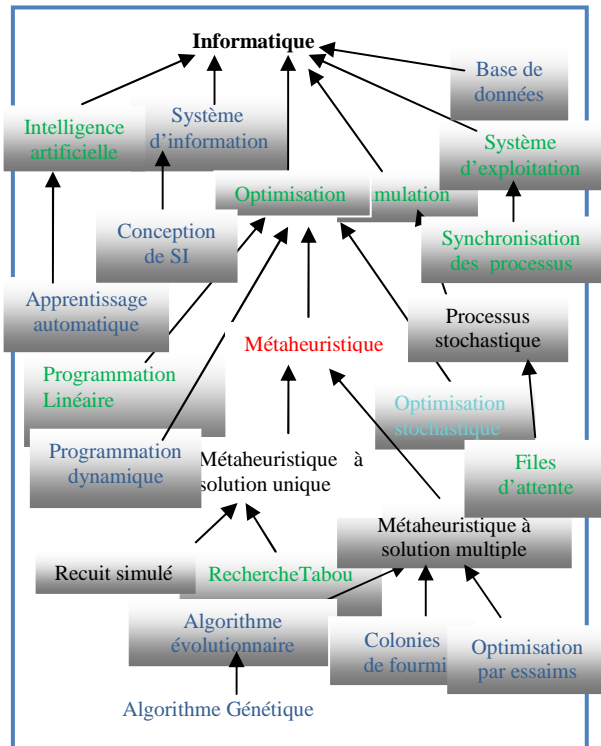
**Enseignant\_1 :**  
**Grade :** MCB ;  
**Modules enseignés :** Métaheuristique, Simulation, Optimisation, Programmation linéaire;  
**Thèse de doctorat :** Simulation, Synchronisation des processus, Files d'attente.  
**Mémoire magistère :** optimisation Stochastique, calcul de probabilité  
**Mémoire d'ingéniorat :** Problème de tournée de véhicule, recherche tabou  
**Travaux d'encadrement :** Optimisation, Simulation, Problème de voyageur de commerce.  
**Autres domaines d'intérêt :** Intelligence artificielle.

**Figure8-** Extrait du CV du premier enseignant décrivant ses compétences.

**Enseignant\_2 :**  
**Grade :** MAA ;  
**Modules enseigné :** Système d'information, Optimisation Stochastique, Programmation linéaire ;  
**Thèse de doctorat :** Colonie de fourmi, apprentissage automatique ;  
**Mémoire Magistère :** voyageur de commerce, Optimisation par essaim , voisinage  
**Mémoire d'ingénieur :** Problème combinatoire, algorithme génétique  
**Travaux d'encadrement :** Base de données, Conception de SI, algorithme évolutionnaire ;  
**Autres domaines d'intérêts :** Système d'exploitation;

**Figure 9-** Extrait du CV du deuxième enseignant décrivant ses compétences.

Pour offrir une image plus complète, nous illustrons dans la figure ci-dessous l'emplacement du module ainsi que les compétences des deux enseignants dans la hiérarchie des concepts de l'ontologie du domaine.



**Figure 10-** Extrait de l'ontologie OntInformatique.

Dans le tableau ci-dessous nous récapitulons les résultats du calcul de similarité par classe de compétence pour chaque enseignant :

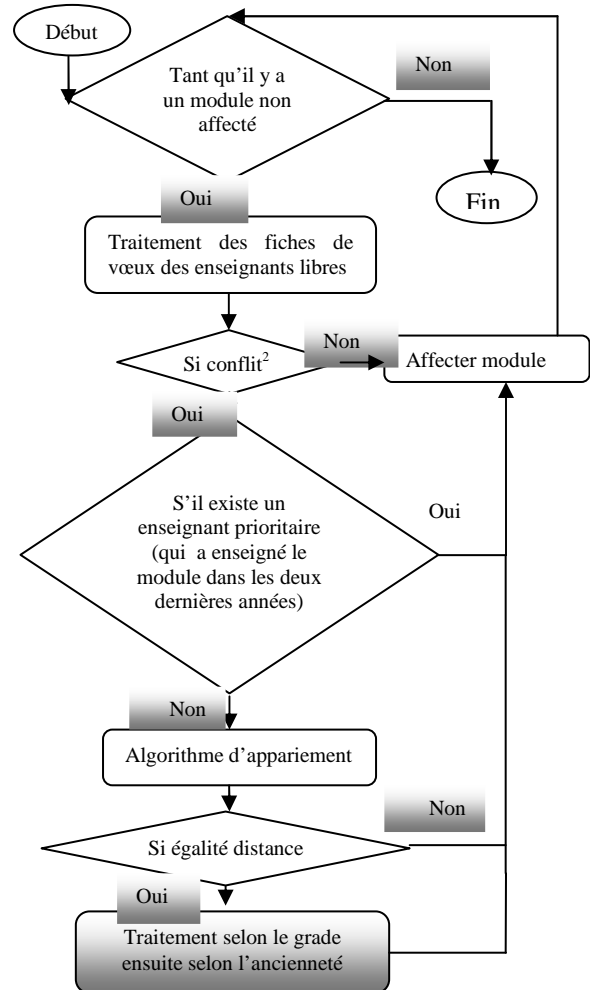
**Table 3-** Résultats de l'étude de cas

	Enseignant 1	Enseignant 2
Distance_Modules enseignés	0,43	0,2
Distance_Thèse de doctorat	0	0,13
Distance_Mémoire magistère	0,173	0,17
Distance_Mémoire d'ingénieur	0,2	0,23
Distance_Travaux d'encadrement	0,14	0,1
Distance_Domains d'intérêts	0	0
Distance Total	0,94	0,83

D'après les calculs, l'enseignant\_1 est le plus adéquat pour l'enseignement du module Métaheuristique. Ce résultat était bien attendu du fait que les compétences de l'enseignant\_1 sont clairement plus proches aux compétences requises pour l'enseignement du module en question. De ce fait, nous considérons que l'algorithme proposé répond très bien à la problématique d'affectation des modules aux enseignants à base de leurs compétences.

### 5. Le système de gestion d'affectation des modules

Le système « SyGAM » (Fig. 12) que nous avons développé est mis à la disposition des enseignants et d'un administrateur. Concernant l'affectation des modules aux enseignants, plusieurs critères de décisions sont mis en jeu : l'ancienneté, le nombre d'années d'enseignement d'un même module<sup>1</sup>, préférences de l'enseignant lui-même (sa fiche de vœux) en plus bien sûr de ses compétences. Nous récapitulons la procédure d'affectation dans le diagramme suivant :



<sup>1</sup>On a appliqué la règle suivante: un enseignant a le droit d'enseigner un module pour 3ans consécutives.

<sup>2</sup> le conflit survient quand plusieurs enseignants veulent se charger du même module.

**Figure 11-** Diagramme décrivant l'enchaînement de la tâche d'affectation des modules d'un semestre



**Figure 12.** Page d'accueil du système SyGAM

A part l'affectation de tous les modules de tous les semestres à tous les enseignants avec utilisation des fiches de vœux, SyGAM offre Différents scénarios d'utilisation du composant d'appariement:

1) L'affectation par semestre - Dans ce cas, l'administrateur choisit le cycle et le semestre dont il veut affecter les modules. L'avantage de ce type d'affectation est de laisser le choix à l'administrateur pour préciser le semestre qu'il trouve le plus prioritaire dans l'affectation.

2) Affectation par module- Cette affectation est entièrement basée compétences où l'administrateur doit choisir le module à affecter avec la possibilité de considérer tous les enseignants du département ou seulement un sous-ensemble. Ce type d'affectation est adapté à la situation où de nouveaux enseignants sont recrutés avec des modules non encore affectés. Cette affectation peut être également exploitée dans une application de recrutement d'enseignants.

## 6. Conclusion

L'algorithme d'appariement décrit dans cet article exploite une combinaison de la distance de Wu-Palmer avec celle de Dice. L'algorithme proposé prend en entrée un module et un ensemble d'enseignants et donne en sortie l'enseignant le plus adéquat en termes des compétences pour enseigner ce module. Pour atteindre ce but, un rapprochement entre les compétences acquises des enseignants et de celle requises par les modules est effectué en se référant à l'ontologie du domaine Informatique. Notons que nous avons associé des poids aux classes de compétences suivant leur importance. Le même principe de pondération était appliqué sur les distances calculées par rapport aux compétences requises d'un module et qui sont issues de relations

différentes dans l'ontologie du domaine informatique. L'algorithme proposé a été implémenté au sein du Système de Gestion d'Affectation des Modules (SyGAM) comme un composant d'aide à la décision.

Concernant les voies de recherche futures, nous envisageons l'extension de l'ontologie OntInformatique pour aboutir à une granularité plus fine permettant une meilleure affectation. De même, nous planifions l'enrichissement de l'ontologie Compétences\_Enseignant pour prendre en compte d'autres critères servant à l'évaluation d'un enseignant.

## References

- [1] W. N. Borst, «Construction of Engineering Ontologies », Centre for Telematica and Information Technology, University of Twente, Enschede, The Netherlands, 1997.
- [2] N. Chergui, « Une approche de mapping pour l'intégration des ontologies », thèse de doctorat, Ecole Doctorale de l'Est en Informatique, Pole de Constantine, 2008.
- [3] J. Euzenat, P. Shvaiko, «Ontology Matching», ISBN 978-3-540-49611-3, Springer-Verlag Berlin Heidelberg, 2007.
- [4] A. Elbyed, «ROMIE, une approche d'alignement d'ontologies à base d'instances », Thèse de doctorat, l'INSTITUT NATIONAL DES TELECOMMUNICATIONS dans le cadre de l'école doctorale S&I en co-accréditation avec l'UNIVERSITE D'EVRY-VAL D'ESSONNE, 2009.
- [5] Z. WU, M. PALMER, «Verb Semantics and Lexical Selection », Proceedings of the 32nd Annual meeting of the Association for Computational Linguistics, 1994.
- [6] H. Zargayouna, V. Salotti, « Mesure de similarité dans une ontologie pour l'indexation sémantique de documents XML », LIMSI/CNRS, Université Paris 11, LIPN - CNRS UMR 7030, Université Paris 13.
- [7] R. Harrathi, S. Calabretto, «Une approche de recherche sémantique dans les documents semi-structurés », LIRIS UMR 5205 - INSA de Lyon, 7 avenue Jean Capelle, 69621 Villeurbanne cedex, France, 2010.
- [8] T. Slimani, B. Boutheina, K. Mellouli, «Une extension de mesure de similarité entre les concepts d'une ontologie », 4th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications – TUNISIA, 2007.
- [9] D. Lin. An Information-Theoretic Definition of similarity. In Proceedings of the Fifteenth International Conference on Machine Learning (ICML'98). Morgan-Kaufmann: Madison, WI, 1998.
- [10] M. Fernandez, A. Gomez-Pérez, N. Juristo «METHONTOLOGY: From Ontological Art Towards Ontological Engineering.», Proceedings of the AAAI-97 Spring Symposium Series on Ontological Engineering, Stanford, CA, USA, 1997.