

N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE

MINISTRE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES



FACULTE DES SCIENCES EXACTES

SIDI BEL ABBES

THESE

De Doctorat en Sciences

Présentée par: Aissa FELLAH

Spécialité: Informatique

Option: Interopérabilité et intégration des systèmes d'information
dans le web

Intitulée:

**Alignement des Ontologies dans un
Contexte Services Web Sémantiques**

Soutenue le: 05/10/2017

Devant le jury composé de:

<i>Président</i>	:	Sofiane BOUKLI HACENE	MCA	UDL-SBA
<i>Examineurs</i>	:	Reda ADJOUJ	MCA	UDL-SBA
		Sidi Mohamed BENSLIMANE	Professeur	ESI-SBA
		Djelloul BOUCHIHA	MCA	CU-NAAMA
		Adil TOUMOUH	MCA	UDL-SBA
<i>Directeur de thèse:</i>		Mimoun MALKI	Professeur	ESI - SBA

Année universitaire 2016/2017

Remerciements

Louange à ALLAH

C'est avec un réel plaisir que j'exprime mes très vifs remerciements à tous ceux qui, de près ou de loin, m'ont accordé leur soutien.

*Qu'il me soit permis tout d'abord d'exprimer ma gratitude et mes remerciements envers le Professeur **Mimoun MALKI**, mon directeur de thèse, qui est à l'origine de ce travail, aussi pour m'avoir fait confiance malgré les connaissances plutôt légères que j'avais au début sur la problématique de l'interopérabilité sémantique, puis pour sa patience et son encouragement tout au long de ce travail.*

C'est un honneur particulier pour moi d'avoir l'occasion de soutenir ma thèse devant un jury des plus respectés et avisés. Je ne peux que remercier chaleureusement :

- Monsieur **Sofiane BOUKLI HACENE** Maitres de conférence A, à l'université de SBA d'avoir accepté de présider mon jury de thèse,*
- Monsieur **Djelloul BOUCHIHA**, Maitre de conférence A, au Centre Universitaire de Naama, Monsieur **Sidi Mohamed BENSLIMANE** Professeur à l'école supérieure d'informatique de SBA et Messieurs **Adil TOUMOUH** et **Réda ADJOUJ** Maitres de conférence A, à l'université de SBA pour avoir accepté de faire partie de mon jury de thèse.*

Pour leurs encouragements et leur assistance aussi bien matérielle que morale qui m'ont permis de faire cette thèse dans de bonnes conditions, je remercie ma famille, ainsi que l'Université Moulay Tahar de SAIDA.

L'aboutissement de cette thèse a aussi été encouragé par de nombreuses discussions avec des collègues de disciplines variées. Je ne citerai pas de noms ici, pour ne pas en oublier certains

Je n'oublie pas les étudiants que j'ai eu le plaisir d'encadrer pour les richesses qu'ils m'ont apportés tant sur le plan scientifique que sur le plan humain.

Je n'oublie pas de penser aux enseignants que j'ai eus par le passé depuis l'école primaire.

A

- ❖ *Mes défunts **parents**. Que Dieu le Tout Puissant leur accorde sa Pleine Miséricorde et les accueille dans son vaste paradis.*
- ❖ *Ma femme pour son soutien sans faille*
- ❖ *Mes enfants qui excitent en moi l'envie de me battre. Que Dieu les bénisse.*

Résumé

L'augmentation du nombre de services web disponibles a généré un problème technique notable à résoudre, qui est la découverte automatique du service web pertinent. Pour traiter cette problématique, les chercheurs ont combiné les services web avec un modèle sémantique, en particulier l'ontologie qui est un choix du web sémantique, ce qui donnait naissance aux Services Web Sémantiques (SWS). Plusieurs travaux ont discuté la découverte sémantique, lorsque les deux services, requis et offert, utilisent la même ontologie. Nous, pensons qu'une approche basée sur l'utilisation d'une seule ontologie n'est pas pratique, car il est très peu probable que tous les fournisseurs de services et les demandeurs adhèrent à la même ontologie. Très peu d'approches considèrent le cas, lorsque les deux services, requis et offert appartiennent à différentes ontologies du même domaine. Les difficultés inhérentes à l'appariement (matchmaking) de deux services web sémantique, requis et offert dans un contexte multi-ontologie constituent les majeures problématiques étudiées dans cette thèse. Notre première contribution consiste à proposer un cadre général de matchmaking des services web sémantiques dans un contexte multi ontologies, dont le noyau est une nouvelle mesure de similarité sémantique entre concepts de deux ontologies. Dans une seconde proposition, nous avons développé un algorithme d'appariement des services web SAWSDL nommé SAWSDL-MOM, qui est une instantiation de notre cadre général et intégrait un algorithme d'alignement partiel des ontologies. Ce dernier utilise une version améliorée de notre mesure de similarité. Les résultats obtenus en expérimentations sont satisfaisants et encouragent les futurs travaux dans cette piste de recherche.

Mots clés : *Ontologie, Appariement des Services web Sémantiques, Alignement Partiel des Ontologies, Mesure de Similarité.*

ملخص

زيادة عدد خدمات الويب المتاحة على شبكة الإنترنت، ولدت مشكلة تقنية كبيرة هي الاكتشاف التلقائي لخدمة الويب المناسبة. لمعالجة هذه المشكلة قام الباحثون بالجمع بين خدمات الويب مع النموذج الدلالي (خصوصا الأنطولوجيا التي هي خيار الشبكة الدلالية)، الذي ولد خدمات الويب الدلالي (SWS). وهذا سعيًا لإظهار أنه إذا تم شرح أوصاف الخدمة ومتطلبات المستخدم بواسطة الأنطولوجيا، فإن النهج الدلالي للعثور على الخدمة المناسبة يصبح أسهل. العديد من الدراسات ناقشت الاكتشاف الدلالي، عندما تستخدم على حد سواء الخدمات المعروضة والمطلوبة نفس الأنطولوجيا. ونعتقد أن المقاربة التي تقوم على اساس استخدام الأنطولوجيا الوحيدة ليست بأمر عملي لأنه من غير المحتمل أن ينضم جميع مقدمي وطالبي الخدمات الى نفس الأنطولوجيا. عدد قليل جدا من المقاربات، نظرت في القضية عندما تنتمي الخدمتين المطلوبة والمقدمة لأنطولوجيات مختلفة في نفس المجال. الصعوبات الكامنة في التوفيق بين خدمتي ويب دلالي مطلوبة و معروضة في سياق أنطولوجيات متعددة هي القضايا الرئيسية التي نوقشت في هذه الأطروحة. لدينا أول مساهمة في توفير إطار عام للتوفيق بين خدمات الويب الدلالي في سياق الأنطولوجيات المتعددة، الذي أساسه هو مقياس جديد للتشابه الدلالي بين مفهومين من أنطولوجيات مختلفة. في اقتراح آخر قمنا بتطوير خوارزمية مطابقة (التوفيق بين) خاصة بخدمات الويب SAWSDL اسمها SAWSDL-MOM، التي هي مثيل للإطار العام المقترح واشتملت على خوارزمية لتوفيق الجزئي بين الأنطولوجيات. هذه الاخيرة تستخدم نسخة محسنة من مقياسنا للتشابه. يمكن القول ان النتائج التي تم الحصول عليها في التجارب مرضية وتشجع على مواصلة العمل في هذا النوع من البحوث.

كلمات البحث: أنطولوجيا، التوفيق بين خدمات الويب الدلالي، التوفيق الجزئي بين أنطولوجيات، قياس التشابه.

Abstract

The increase in the number of available web services, has generated a significant technical problem is automatic discovery of the most relevant web service. To address this problem, the researchers combined web services with semantic model (especially ontologies that are a semantic web choice), which gave birth to the Semantic Web Services (SWS). They sought to show that if the service descriptions and user requirements are annotated with ontology, then the semantic approach to find the appropriate service becomes easier. Several studies have discussed the semantic discovery, when both services offered and required use the same ontology. We believe that an approach based on the use of a single ontology is not practical because it is very unlikely that all the service providers and the applicant adhere to the same ontology. Very few approaches consider the case when the two services required and offered belong to different ontologies of the same domain. The difficulties inherent in the Matchmaking of two semantic web services required and offered in a multiple ontology contexts are the major issues discussed in this thesis. Our first contribution is to provide a general framework for semantic web services matchmaking in multiple ontology contexts, whose core is a new measure of semantic similarity between concepts of two ontologies. In another proposal, we have developed a matching algorithm (matchmaking) for SAWSDL web services named SAWSDL-MOM, that instantiated our general framework and included a partial ontology alignment algorithm. The latter uses an improved version of our similarity measure. The results obtained in experiments are satisfactory and encourage further work in this line of research.

Keywords: Ontology, Semantic Web Services Matchmaking, Partial Ontology Alignment, Similarity Measure.

Table des matières

REMERCIEMENTS	I
RESUME	III
ملخص	IV
ABSTRACT	V
TABLE DES MATIERES	VI
LISTE DES FIGURES	IX
LISTE DES TABLES	X
CHAPITRE 1 : INTRODUCTION	I
1.1 Préambule	1
1.2 Problématique	4
1.3 Objectifs	4
1.3.1 Sous-objectifs	4
1.4 Contributions	5
1.5 Organisation du manuscrit	5
PARTIE I : BACKGROUND & ETAT DE L'ART	8
Chapitre 2 : Background	9
2.1 Introduction	10
2.2 Service Web	10
2.2.1 Introduction	10
2.2.2 Définition	11
2.2.3 Standards et description de services Web	12
2.2.4 Synthèse sur les services web	17
2.3 Ontologies	17
2.3.1 Introduction	17
2.3.2 Synthèse de représentation des connaissances avant l'ontologie	18
2.3.3 La notion d'ontologie	19
2.3.4 L'ontologie techniquement	21
2.3.5 Quelques cas d'utilisation des ontologies	22
2.3.6 Les ontologies dans le Web sémantique	23
2.3.7 Langages d'ontologie	23
2.4 Service web sémantique	28
2.4.1 Introduction	28
2.4.2 Description des services web sémantiques	28
2.4.3 Langages de description sémantique	29
2.4.4 Annotation du langage de description	31
2.4.5 SAWSDL (Semantic Annotations for Web Services Description Language)	32
2.5 Conclusion	36

Chapitre 3 : Matchmaking des Services Web.....	37
3.1 Introduction	38
3.2 Concepts généraux.....	38
3.2.1 Découverte de services web	38
3.2.2 Matchmaking.....	39
3.3 Processus de Matchmaking	40
3.4. Dimensions d'un processus de matchmaking	41
3.4.1 Entrée du Matchmaker.....	41
3.4.2 Sortie du Matchmaker	42
3.4.3 Ressources & paramètres.....	42
3.5 Caractéristiques du processus de matchmaking	43
3.6 Les travaux de Matchmaking	45
3.6.1 Travaux de base.....	45
3.6.2 Matchmaking des services web dans un contexte ontologie unique	47
3.6.3 Matchmaking des services web dans un contexte multi-ontologie.....	48
3.6.4 Approches de Matchmaking pour SAWSDL	50
3.7 Synthèse et conclusion.....	52
3.7.1 Synthèse de l'état de l'art	52
3.7.2 Conclusion.....	55
Chapitre 4 : Appariement sémantique élémentaire dans le Matchmaking des SWS	56
4.1 Introduction	57
4.2 Similarité sémantique	57
4.2.1 Définition.....	58
4.2.2 Classification des mesures de similarité sémantique.....	58
4.3 Techniques d'appariement sémantique élémentaire	60
4.4 Besoin d'interopérabilité sémantique	61
4.4.1 Formes d'hétérogénéité.....	61
4.4.2 Surmonter l'hétérogénéité.....	64
4.5 Alignement des ontologies comme solution à l'interopérabilité sémantique	64
4.5.1 Terminologie	65
4.5.2 Les techniques de comparaison ou matchers.....	66
4.5.3 Quelques systèmes.....	68
4.6 Alignement partiel des ontologies	71
4.7 Synthèse sur les techniques d'appariement sémantique élémentaire.....	71
4.8 Conclusion.....	74
PARTIE II : CONTRIBUTION ET EVALUATION.....	75
Chapitre 5: Matchmaking des SWS dans un contexte multi-ontologie	77
5.1 Introduction	77
5.2 Problématique.....	77
5.3 Architecture de la découverte	78
5.4 Framework général	79
5.5 Mesure de Similarité inter Ontologies.....	82
5.5.1 Algorithme d'appariement des concepts	82
5.5.2 La similarité des URI (URI Matcher).....	86
5.5.3 La similarité des noms (Names Matcher).....	86
5.5.4 La similarité de contexte interne:	87
5.5.5 La similarité de contexte externe.....	89
5.5.6 Similarité globale (Global Matcher).....	95

5.6 Matchmaking des services web basé sur l’alignement partiel des ontologies	95
5.6.1 Spécification	96
5.6.2 Identification des éléments à apparier	97
5.6.3 Instanciation du Framework pour SAWSDL	97
5.6.4 Algorithme d’appariement des services web SAWSDL (SAWSDL Matchmaking)	99
5.6.5 Alignement partiel	102
5.6.6 Mesure de similarité	105
5.7 Vers une extension de SAWSDL-TC	106
5.8 Conclusion	108
Chapitre 6 : Evaluation & Expérimentation	109
6.1 Introduction	110
6.2 Principe et indicateurs de l’évaluation.....	110
6.2.1 Les indicateurs et mesures	110
6.3 Bases de tests standards	111
6.3.1 La base de test de l’alignement des ontologies OAEI(2014).....	111
6.3.2 La base de test SAWSDL –TC	111
6.4 Evaluation expérimentale	112
6.4.1 Evaluation de la mesure de similarité.....	112
6.4.2 Evaluation de l’algorithme d’alignement partiel	114
6.4.3 Evaluation de l’algorithme de Matchmaking SAWSDL-MOM.....	116
6.5 Discussion.....	118
6.6 Conclusion.....	120
CONCLUSION & PERSPECTIVES	121
Conclusion	122
Perspectives.....	123
BIBLIOGRAPHIE	125

Liste des figures

Figure 1.1 : Synthèse du plan de la thèse	7
Figure 2.1: Cycle de vie des interactions de services Web proposé par W3C(Klusch, 2014).	12
Figure 2.2: Exemple de description WSDL d'un service web (Dumas & Fauvet, 2008).....	15
Figure 2.3: Vue d'ensemble de l'ontologie service OWL-S.....	29
Figure 2.4 : Méta modèle SAWSDL (Klusch et al., 2009).	33
Figure 3.1 : Catégorisation des matchmakers des SWS.....	40
Figure 3.2 : Le processus général de matchmaking des SWS.	41
Figure 4.1: Classification des mesures de similarité intra ontologie.	59
Figure 4.2: Classification des mesures de similarité inter ontologies.....	60
Figure 4.3 : L'hétérogénéité conceptuelle (Euzenat et al., 2004)	62
Figure 4.4 : Techniques de base pour l'alignement des ontologies (inspirée de (Euzenat et al., 2004).	67
Figure 5.1: Découverte de service web sémantiques.....	79
Figure 5.2: Framework de Matchmaing des SWS multi ontologie.	81
Figure 5.3: Calcul de similarité entre deux concepts.	83
Figure 5.4: Le Processus de Matching des Concepts.....	84
Figure 5.5 : Calcul de similarité globale.....	84
Figure 5.6: Comparaison entre deux concepts: Network Node and Network Equipment.....	85
Figure 5.6: La similarité des contextes externes (pères) des deux concepts: CR & CS.....	91
Figure 5.7: La similarité des contextes externes des deux concepts: Entry & Reference.....	94
Figure 5.8: Le processus de Matchmaking SAWSDL-MOM.....	98
Figure 5.9 : Détail du processus de Matching.....	100
Figure 5.10 (Listing): Algorithme de Matchmaking.....	102
Figure 5.11(Listing) : Algorithme d'Alignement Partiel	104
Figure 5.12(Listing) : Calcul de similarité globale.....	106
Figure 6.1 : Résultats graphiquement (tests de notre mesure).....	114
Figure 6.2: Les tests de l'Alignement Partiel.....	116

Liste des tables

<i>Table 1.1: La Différence entre WSDL 1.1 et WSDL 2.0.</i>	13
<i>Table 2.1. Définitions choisies pour "L'ontologie".</i>	20
<i>Table 3.1: Synthèse des travaux de matchmaking</i>	54
<i>Table 4.1 : Synthèse sur les techniques d'appariement sémantique élémentaire</i>	73
<i>Table 6.1: Les résultats des tests de notre mesure.</i>	113
<i>Table 6.2: Résultats de comparaison des matchmakers.</i>	118

Chapitre 1 : Introduction

1.1 Préambule

Le WWW et l'avènement de l'Architecture Orientée Service (SOA : Service Oriented Architecture) nous ont fait entrer dans une nouvelle ère de la révolution de l'information. Ces technologies ont donné naissance à la «nouvelle économie» et ont changé la façon avec laquelle les entreprises fonctionnaient. Les avantages du couplage faible offert par les architectures orientées services (SOA) les ont fait un choix populaire pour les systèmes d'entreprise d'aujourd'hui. La popularité a conduit les efforts de normalisation dans ce domaine, de la publication à l'invocation de services. Les services spécifiés en utilisant ces normes (Booth et al., 2004), sont désignés comme des services Web. L'introduction de nouvelles normes telles que UDDI (Cerami et al., 2002 ; Bellwood al., 2002), WSDL (Christensen et al., 2001 ; Chinnici et al., 2006) et SOAP (Box et al., 2000) sont un pas dans la bonne direction. Les services web sont devenus un paradigme adéquat pour le développement des systèmes distribués à grand échelle. Néanmoins, comme toute nouvelle technologie, ils viennent avec une nouvelle série de défis, dont certains sont spécifiques et d'autres sont classiques liés à la technologie. Le développement des services Web soulève ainsi des difficultés similaires à celles qui ont accompagné la croissance du Web. En effet, ces services sont développés par différentes entités et il n'existe pas de consensus sur la manière de faire usage des descriptions de services. Les services sont créés, mis à jour ou supprimés à la volée (Cherifi, 2011) ; ces changements peuvent, de plus, ne pas être reportés dans les annuaires. Pour de multiples raisons, les services peuvent également présenter des problèmes de dysfonctionnement au moment de leur utilisation. Ces caractéristiques conduisent à un environnement extrêmement dynamique et volatile. Publier, découvrir et composer des services dans un tel environnement pose un certain nombre de problèmes. La recherche de services doit se faire dans un ensemble de grande taille non structuré et changeant. Un problème technique notable à résoudre est la découverte de service web le plus pertinent parmi un large éventail de services web offerts. La description des services est une étape essentielle pour la découverte et la composition. Les informations qu'elle contient permettent de guider la recherche et le choix des services et conditionnent leur interaction. Une description précise la fonction d'un service, ses contraintes de fonctionnement et la façon d'interagir avec lui. Schématiquement, on peut classer ces informations en quatre catégories : les informations générales relatives à la nature du service (nom, description), les informations techniques relatives à la façon de l'utiliser, les informations non fonctionnelles qui concernent la qualité de service et enfin les informations fonctionnelles qui décrivent les fonctionnalités du service. La description des services web repose actuellement sur des standards comme WSDL (Web Services Description Language)

Chapitre 1 : Introduction

(Christensen et al., 2001; Chinnici et al., 2006), qui fournit une description syntaxique des services. Il spécifie les fonctionnalités d'un service en définissant des messages et des opérations. Les messages fournissent une définition abstraite des données qui doivent être échangées. Les opérations sont fournies par les services pour transformer les messages. Chaque message contient un ou plusieurs paramètres. WSDL permet une description centrée sur la fonction du service qui est représentée par les paramètres en entrée et en sortie des opérations. La découverte de service Web est le mécanisme qui permet de trouver le service le plus approprié qui satisfait la demande de l'utilisateur. Le noyau de toute architecture de découverte est un processus d'appariement (Matchmaking : voir NB ci-dessous) entre la requête et le service offert pour déterminer leur degré de correspondance global (GDoM : Global Degree of Match ou simplement DoM). Un appariement élémentaire entre les composants de la requête et du service est nécessaire, ainsi le GDoM est le résultat des agrégations et des propagations des résultats de l'appariement élémentaire. Un certain nombre d'approches pour l'appariement des services web ont été proposées. Les mécanismes de découverte de services Web traditionnels impliquent une sélection manuelle des services web requis, par les utilisateurs, à partir d'une longue liste textuelle. Ce processus prend énormément de temps. Comme le nombre de Services Web augmente cette approche devient de plus en plus infaisable et trouver le meilleur service selon les besoins des utilisateurs devient un défi. Avec le web sémantique, les ontologies (GRUBER, 1993 ; GRUBER & OLSEN, 1994) émergent comme une solution prometteuse pour représenter un ensemble de connaissances et permettre leur partage et réutilisation. Les services Web, en tant qu'applications disponibles sur le Web, suivent cette même tendance, des premiers travaux de recherche ont identifié l'application des ontologies comme l'une des solutions à ce problème. Ils cherchaient à démontrer que si les descriptions de services et les besoins des utilisateurs sont annotés avec des ontologies, alors l'approche sémantique pour trouver le service approprié devient plus facile. Donc, les services web combinés avec un modèle sémantique, en particulier les ontologies qui sont un choix du web sémantique, donnaient naissance aux Services Web Sémantiques (SWS : Semantic Web Services). D'une façon générale, c'est un passage d'une description syntaxique à une description sémantique. Avec des descriptions sémantiques interprétables par la machine, les SWS visent l'automatisation des activités liées à l'exploitation comme la découverte et la composition. Au niveau fonctionnel, les services Web sémantiques sont décrits de la même façon que les services syntaxiques, c'est-à-dire avec des opérations et des paramètres d'entrée et de sortie. La différence majeure est que les paramètres sont dans ce cas associés à des concepts ontologiques. Citons les principaux langages qui sont OWL-S (Ontology Web Language for Services) (Martin et al., 2004) et WSMO (Web Services Modeling Ontology) (De Bruijn et al., 2006). Une approche médiane consiste à enrichir les descriptions

Chapitre 1 : Introduction

syntaxiques des services en utilisant des ontologies. Les deux principales propositions en ce sens sont WSDL-S (WSDL-Semantic) (Akkiraju et al., 2005) et SAWSDL (Semantic Annotation for WSDL) (Farrell & Lausen, 2007). Ces spécifications W3C établissent une correspondance entre certains éléments WSDL existants et des concepts ontologiques, avec le but d'aboutir à une annotation automatique des descriptions WSDL existantes. Pour atteindre une découverte automatique et pertinente, les efforts de recherche en appariement des services Web sémantiques essaient d'identifier les degrés de correspondance entre les concepts sémantiques qui décrivent les éléments des services et des requêtes (par exemple les entrées, les sorties). Des premiers travaux de recherche (Syeda-Mahmood et al., 2005; Plebani & Pernici, 2009 ; Klusch et al., 2009) ont identifié l'application des ontologies comme l'une des solutions à ce problème. Ils cherchaient à démontrer que si les descriptions de services et les besoins des utilisateurs sont annotés avec des ontologies, alors l'approche sémantique pour la découverte du service approprié devient plus facile, avec une découverte entraînant l'identification du service le plus pertinent avec un certain degré d'automatisation. Plusieurs travaux (Sycara et al., 1999; Paolucci et al., 2002; Plebani & Pernici, 2009; Klusch et al., 2009; Klusch, 2014 ; Schulte, 2010 ; Sycara et al., 2002 ; Li & Horrocks, 2004 ; Zinnikus, 2006) ont discuté l'appariement sémantique, lorsque les deux services, requis et offert utilisent la même ontologie. Une approche basée sur l'utilisation d'une seule ontologie n'est pas pratique car il est très peu probable que tous les fournisseurs de services et demandeurs adhèrent à la même ontologie. Très peu d'approches (Cardoso & Sheth, 2003 ; Oundhakar et al., 2005) considèrent le cas, lorsque les deux services, requis et offert appartiennent à différentes ontologies du même domaine.

NB : Le Matchmaking est un anglicisme désignant la mise en relation de deux individus par affinités, il est également utilisé dans les jeux vidéo, par certains jeux multi-joueurs, pour la sélection On-line de joueurs de même niveau ; aussi, il est utilisé dans le monde de travail pour aboutir à des rencontres professionnelles réelles. Les chercheurs du domaine des services web ont emprunté ce terme. La traduction en français de « web services Matchmaking » n'est pas simple. Jumelage des services web, entremise des services web ou mise en correspondance des services web, comme le proposent les dictionnaires de traduction, ne semblent pas utilisables par les chercheurs francophones du domaine des services web. Quelques auteurs de travaux en français comme Chabeb (2011), utilisent le terme « appariement des services web ». Dans cette thèse, on utilise le terme appariement des services web et on conserve le terme original « Matchmaking » d'une manière interchangeable. La même remarque est constatée pour le terme Matchmaker. Les dictionnaires proposent les termes (intermédiaire, marieur, entremetteur). Nous optons pour les termes Apparieur et Matchmaker, toujours d'une manière interchangeable.

Chapitre 1 : Introduction

1.2 Problématique

Étant donnés, deux services, requis et offert annotés avec différentes ontologies du même domaine. Les ontologies servent à enrichir les descriptions des services avec de la sémantique pour permettre un Matchmaking sémantique entre les services, l'appariement des services peut être réduit au problème d'interopérabilité sémantique d'ontologies d'annotation (Dietze et al., 2009). D'une part, les ontologies, elles-mêmes souffrent des hétérogénéités. D'autre part, l'alignement des ontologies est connu une technique puissante qui permet de surmonter l'hétérogénéité entre les ontologies. Un problème technique important à résoudre est celui de surmonter les difficultés inhérentes au Matchmaking des services web sémantiques dans un contexte multi ontologies, par l'intégration des techniques d'interopérabilité sémantique des ontologies, en particulier l'alignement des ontologies.

1.3 Objectifs

Notre objectif est de traiter la problématique d'interopérabilité sémantique des services web sémantique, en particulier le Matchmaking sémantique des services web dans un environnement multi ontologie.

Notre premier objectif consiste à proposer un algorithme d'appariement pertinent entre une requête de services et les descriptions des services publiés dans un contexte multi ontologie. Compte tenu du contexte de notre travail et des besoins décrits ci-dessus, nous avons identifié plusieurs points à considérer comme sous objectifs.

1.3.1 Sous-objectifs

- ✓ Proposer un cadre général d'appariement des services web sémantiques dans un contexte multi-ontologie.
- ✓ Identifier avec précision, les composants fondamentaux pour réussir l'appariement sémantique des services web dans un environnement multi-ontologie. En particulier, la nécessité de l'usage des techniques pour l'appariement des ontologies.
- ✓ Proposer un algorithme qui mesure de similarité entre concepts de différentes ontologies. La technique combine différents matchers élémentaires.
- ✓ Proposer une approche d'alignement partiel entre ontologies de domaine. La précision des techniques d'appariement est primordiale pour évaluer un certain aspect de la pertinence d'une approche : plus on réussit à apparier les éléments des deux services, requis et offert, plus le résultat de matchmaking est pertinent, et plus on identifie avec précision la similarité entre les éléments des deux services, requis et offert, plus le résultat de découverte est pertinent.

- ✓ Proposer une approche de matchmaking des services web (En particulier les services web SAWSDL), qui est basée sur l'alignement partiel des ontologies.
- ✓ Etudier, l'agrégation des résultats des appariements élémentaires (ou plus généralement partiels) pour calculer de manière pertinente le degré final/global d'appariement.

Les sous-objectifs visent une contribution guidée plutôt par l'aspect quantitatif (précision, rappel et de temps de réponse).

Ces travaux de thèse, présentent, une extension des travaux antérieurs, de matchmaking sémantique, lorsque les deux services, requis et offert utilisent la même ontologie. Nous utilisons une technique basée sur l'alignement partiel des ontologies. Nous introduisons de nouveaux matchers, en utilisant de nouvelles mesures de similarité. En particulier, pour apparier deux concepts, il faut prendre en compte leurs contextes internes et externes. Cette thèse présente également une évaluation de l'algorithme utilisant une implémentation d'un prototype. Les résultats préliminaires montrent que l'utilisation de mesures supplémentaires comme contexte peuvent conduire à une grande précision que les autres approches qui ne le considèrent pas.

1.4 Contributions

Les principales contributions de cette thèse sont:

- ❖ Un Framework général de Matchmaking de services web dans un contexte multi ontologie qui utilise l'alignement des ontologies,
- ❖ Deux variantes d'une mesure de similarité entre concepts de différentes ontologies,
- ❖ Un algorithme d'alignement partiel des ontologies,
- ❖ Une approche de Matchmaking des services web sémantique dans un contexte multi-ontologie pour les services web sémantiques SAWSDL.

Pratiquement tous efforts d'intégration et d'interopérabilité des données et des applications nécessitent une étape d'appariement. Alors que nous présentons nos idées dans le cadre de Matchmaking des services Web sémantiques, elles sont largement applicables à d'autres situations de même défi (besoin d'interopérabilité sémantique), si l'on envisage un environnement multi-ontologie.

1.5 Organisation du manuscrit

Ce mémoire, au plus de l'introduction générale, est organisé en deux grandes parties :

Partie 1 : Background & Etat de l'art et

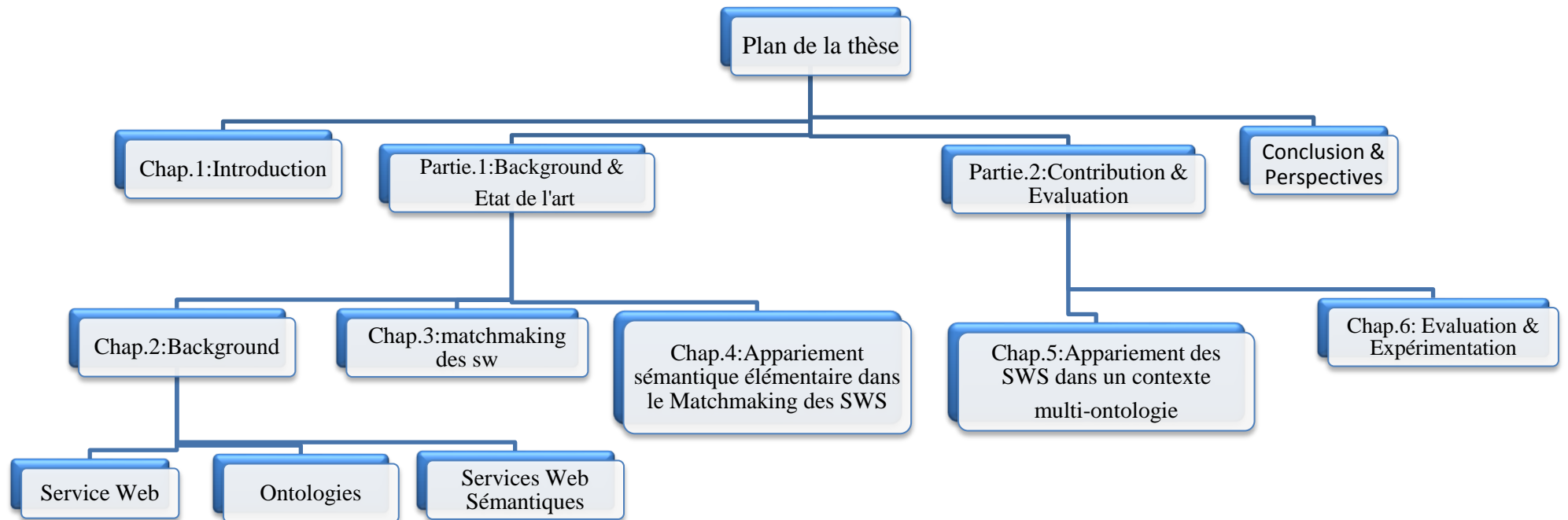
Partie 2 : Contribution & Evaluation.

Chapitre 1 : Introduction

La première partie compte trois chapitres : Le chapitre 2 (Background) traite la documentation pour cette thèse. Afin de mieux comprendre l'enchaînement des concepts fondamentaux de notre recherche, nous présentons des informations de base sur les services Web et les normes ou les technologies connexes. La section suivante du chapitre est consacrée à la notion d'ontologie, cette approche qui a émergé depuis les années 90 comme un enjeu stratégique dans la représentation et la modélisation des connaissances. Enfin, le chapitre présente le concept de service web sémantique avec les différentes approches de description. Dans la suite de la première partie, nous avons effectué un état de l'art sur le matchmaking des services web et l'appariement élémentaire entre composants de services. Le chapitre 3 est consacré à l'état de l'art du matchmaking des services web. On commence par quelques définitions. Ensuite, nous présentons la description détaillée du processus d'appariement des services web sémantiques. Dans le chapitre 4, nous présentons l'appariement élémentaire sémantique comme élément central du matchmaking des services web sémantiques. Un point important dans ce chapitre concerne l'étude de la similarité sémantique (intra et inter) ontologies et l'alignement des ontologies pour surmonter l'hétérogénéité.

Nous avons consacré la seconde partie (Contribution & Evaluation) à la présentation de nos propositions et de leurs évaluations. Le chapitre 5 présente notre approche d'appariement des services web sémantique dans un environnement multi-ontologie. En premier lieu, nous avons proposé un Framework qui détail : le matchmaking avec plusieurs ontologies tout en déterminant les composantes principales. Ensuite, nous avons proposé une approche de matchmaking des services web sémantiques SWSDL basée sur l'alignement partiel des ontologies. Le chapitre 6 décrit les principes de nos évaluations ainsi que les résultats et leur discussion. Enfin, les conclusions et les travaux futurs sont mentionnés.

Figure 1.1 : Synthèse du plan de la thèse



Partie I : Background & Etat de l'Art

- *Chapitre 2 : Background*
- *Chapitre 3 : Matchmaking des Services Web Sémantique*
- *Chapitre 4 : Appariement sémantique élémentaire dans le Matchmaking des SWS*

Chapitre 2 : Background

Chapitre 2 : Background

2.1 Introduction

Le présent chapitre est un background théorique qui contient trois sections. La première section traite des concepts de base sur les services Web et les normes ou les technologies connexes. La deuxième section présente la notion d'ontologie, nécessaire pour introduire les services web sémantiques dans une troisième section. Nous pensons que ces notions sont importantes pour suivre le reste du document.

2.2 Service Web

2.2.1 Introduction

Le World Wide Web (ou tout simplement Web) été étonnamment réussi. Le secret de ce succès qu'il est basé sur des standards de communication ouverts (par exemple, HTTP) et des documents de normes ouvertes (par exemple, HTML). D'autre part, XML est devenu le standard pour l'échange de données entre applications à la fois dans le Web et dans des environnements non-Web. Les services Web sont une technologie Internet qui permet aux entreprises de fournir leurs services et de découvrir des services d'autres entreprises. Ces services sont des applications modulaires autonomes et auto-descriptives que les entreprises peuvent publier, découvrir et invoquer sur le Web. S'appuyant sur le succès du World Wide Web et XML, pour la quête de l'interopérabilité, les services Web constituent un paradigme de l'informatique distribuée, basé sur des standards ouverts basés sur XML. Leur standardisation tourne autour des normes principales suivantes :

- WSDL : Web Service Description Language (Christensen et al., 2001) ; (Chinnici et al., 2006) ; pour décrire Les services Web,
- UDDI : Universal Description Discovery and Integration (Ceramiet al., 2002) ; (Bellwood et al., 2002) ; pour la publication et décrivant les services Web,
- SOAP : Simple Object Access Protocol (Box et al., 2000) ; pour la communication à travers des services Web.

En plus de ces normes, il y a d'autres normes et projets de normes qui traitent des aspects non fonctionnels de services Web. Les spécifications qui fournissent un langage indépendant du domaine pour décrire les aspects non fonctionnels de services :

- Web WS-Policy (Vedamuthu al., 2007),
- WS-Agreement (Andrieux al., 2004).

Chapitre 2 : Background

En outre, un certain nombre de spécifications sont développées pour décrire les différents aspects non fonctionnels. Pour Ainsi,

- WS-Security (Nadalin al., 2002) est un vocabulaire pour le domaine de la sécurité qui définit les jetons de sécurité, algorithmes de chiffrement et d'autres objets liés à la sécurité. D'autres vocabulaires spécifiques à un domaine inclure,
- WS-Trust (Anderson al., 2005),
- WS-Transaction (Cabrera al., 2002).

Deux importants efforts de standardisation dans l'aspect comportemental :

- le langage de description de processus métiers exécutables pour services Web BPEL (Web Services Business Process Execution Language) (Andrews et al., 2003) et
- WS-CDL (Web Services Choreography Definition Language) (Kavantzas, 2004)

2.2.2 Définition

L'objectif fondamental des services web est d'assurer une interopérabilité entre applications à travers le web, dans une optique de rendre le web plus dynamique. Le concept « Web service » désigne principalement un programme mis à disposition sur Internet par un fournisseur de services et accessible par les clients à travers des standards. Plusieurs définitions pour les services Web ont été mises en avant par différents auteurs. Ci-dessous, nous citons une définition largement acceptée et fournie par le consortium W3C (Booth al., 2004):

“A Web service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL¹). Other systems interact with the Web service in a manner prescribed by its description using SOAP² messages, typically conveyed using http with an XML serialization in conjunction with other Web-related standards.”

La figure 2.1 illustre l'interaction d'UDDI, WSDL et SOAP, comme suit :

- 1) Un fournisseur de services Web décrit son service via WSDL et publie la définition WSDL du service à l'annuaire UDDI.
- 2) Ensuite, un consommateur de service envoie une ou plusieurs requêtes au registre UDDI pour trouver un service et connaître le mode de communication avec ce service.

¹ Web Service Description Language

² Simple Object Access Protocol

Chapitre 2 : Background

- 3) Une partie du WSDL fourni par le fournisseur de services, serait transmise au client. Ce qui exprime le flux des échanges entre la requête du client et le service du fournisseur.
- 4) Ensuite, le consommateur communique en utilisant le SOAP pour envoyer une requête au fournisseur.
- 5) Le fournisseur retourne la réponse requise au client.

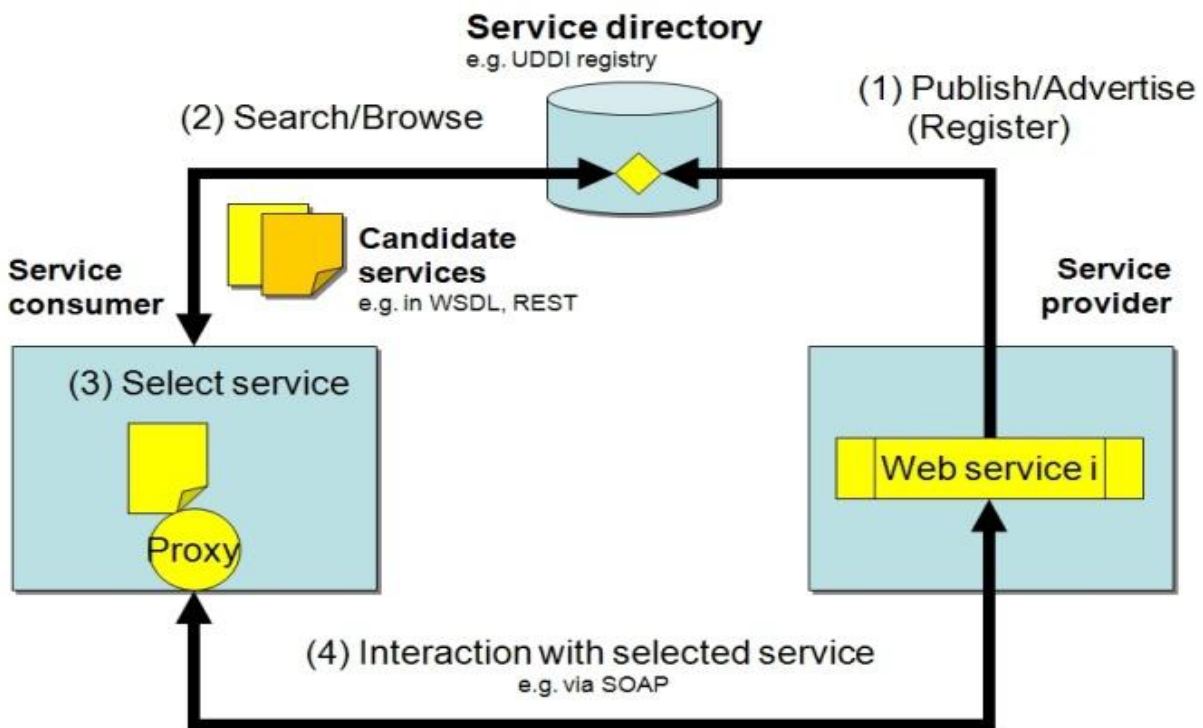


Figure 2.1: Cycle de vie des interactions de services Web proposé par W3C (Klusch, 2014).

2.2.3 Standards et description de services Web

L'aspect description est fondamental pour la technologie des services web. Une description peut porter sur la dimension fonctionnelle structurelle et/ou dimension fonctionnelle comportementale, ainsi que la dimension non fonctionnelle. Les langages de description des dimensions différentes d'un service web sont basés sur le standard XML (eXtensible Markup Language). XML est un langage de balisage extensible, devenu une recommandation W3C et un standard pour la représentation de données semi-structurées. Dans un document XML, la structure de données est fournie par le biais de l'utilisation de balises comme en SGML (Standard Generalized Markup Language) (Wood, 1995), mais en s'affranchissant des aspects liés à la présentation des données). Dans un document XML, les données peuvent être spécifiées conformément à une grammaire ou un schéma, cependant, il est permis de définir des parties partiellement. Une problématique généralement associée à la description de services, est celle de leur publication et de leur

Chapitre 2 : Background

découverte. Un autre standard proposé pour assurer la publication et la découverte des services, c'est l'UDDI (Cerami et al., 2002; Bellwood et al., 2002). UDDI définit une interface de programmation pour publier des descriptions de services dans des répertoires dédiés, pour soumettre des requêtes à base de mots-clés sur ces répertoires et pour naviguer au travers des descriptions obtenues par le biais de ces requêtes. Dans la suite, on présente les langages de description des dimensions différentes des services.

2.2.3.1 Description fonctionnelle et structurelle (WSDL)

WSDL (**Web Services Description Language**) (Christensen et al., 2001; Chinnici et al., 2006), est un langage de définition d'interface basé sur XML qui est utilisé pour décrire la fonctionnalité offerte par un service Web. Il a été construit en combinant deux langages de description de service: NASSL (**Network Application Service Specification Language**) d'IBM et SDL (**Service Description Language**) de Microsoft. Il fournit une description interprétable par machine de la manière avec laquelle le service peut être invoqué, des paramètres qu'il attend et des structures de données qu'il renvoie. Donc, son objectif est semblable à celui d'une signature de méthode dans un langage de programmation ou aussi les interfaces IDL (**Interface Definition Language**) de CORBA. La version actuelle de WSDL est WSDL 2.0 (la *Table 1.1* présente les différences entre les principales versions WSDL).

WSDL 1.1	WSDL 2.0
Definition	Description
Types	Types
Message	
PortType	Interface
Binding	Binding
Service	Service
Port	PortType

Table 1.1: La Différence entre WSDL 1.1 et WSDL 2.0.

Dans la version WSDL1.1, un service comprend un ensemble de messages et un ensemble de points d'entrée (portType). Un point d'entrée est la description abstraite d'une interface et de son implantation. La description abstraite inclue : - la définition des messages qui sont consommés et générés par le service (les entrées et les sorties), et - la signature des opérations offertes par le service. La liaison entre l'interface et son implantation enveloppe l'information sur le protocole

Chapitre 2 : Background

employé pour l'échange des messages avec le service (par exemple SOAP au-dessus de HTTP) et les liens entre la description de l'interface abstraite du service et les types de messages supportés par le protocole bas de communication (par exemple SOAP).

Exemple WSDL

```
<?xml version="1.0"?>
<definitions name="RecevoirBdC"
targetNamespace="http://www.yothuyindi.fr:8080/exemple/fournisseur.wsdl"
xmlns:xsd="http://www.w3.org/2001/XMLSchema/"
xmlns:wns="http://www.yothuyindi.fr:8080/exemple/fournisseur.wsdl"
xmlns:xsdl="http://www.yothuyindi.fr:8080/exemple/fournisseur.xsd"
xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
xmlns="http://schemas.xmlsoap.org/wsdl/">
<types>
<schema
targetNamespace="http://www.yothuyindi.fr:8080/exemple/fournisseur.xsd">
<xsd:complexType name="Commande"><xsd:sequence>
<xsd:element name="dateCommande" type="xsd:date">
<xsd:element name="LigneDeCommande" minOccurs="1"
maxOccurs="unbounded">
<xsd:complexType><xsd:sequence>
<xsd:element name="RéférenceProduit" type="xsd:string"/>
<xsd:element name="Quantité" type="xsd:positiveInteger"/>
</xsd:sequence></xsd:complexType></xsd:element>
</xsd:sequence></xsd:complexType>
...</schema>
</types>
<message name="SoumissionBdC">
<part name="body" element="xsdl:Commande"/>
</message>
<message name="RécépisséBdC">
<part name="body" element="xsd:string"/>
</message>
<portType name=pt_RecevoirBdC>
<operation name="op_Commande">
<input message="wsn:SoumissionBdC">
<output message="wsn:RécépisséBdC">
</operation>
<operation name=...
...
</portType>
<binding name="bd_opCommande" type=pt_RecevoirBdC>
<soap:binding style="document"
transport="http://schemas.xmlsoap.org/soap/http"/>
<operation name="op_Commande">
<soap:operation
```

Chapitre 2 : Background

```
soapAction="http://www.yothuyindi.fr/exemple/op_Commande"/>
<input>
<soap:body
use="literal"
namespace="http://www.yothuyindi.fr/exemple/fournisseur.xsd"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding"/>
</input>
<output>
<soap:body
use="literal"
namespace="http://www.yothuyindi.fr/exemple/fournisseur.xsd"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding"/>
</output>
</operation>
</binding>
<service name=ServiceFournisseur>
<documentation> Service de réception des commandes </documentation>
<port name="PortSoumissionCommande" binding="bd_opCommande">
<soap:address location="//www.yothuyindi.fr/exemple/fournisseur"/>
</port>
</service>
<!-- fermeture de la balise racine -->
</wsdl:definitions>
```

Figure 2.2: Exemple de description WSDL d'un service web (Dumas & Fauvet, 2008)

La Figure 2.2 donne la description WSDL de l'interface du service de réception d'une commande offert par le fournisseur. L'élément *definitions* constitue la racine du document et fournit les espaces de noms. Les types de données, paramètres d'entrée et/ou de sortie sont éventuellement décrits ensuite en XMLSchema (Gao et al., 2009). La description d'une commande est décrite comme une date et une liste non vide de produits a commandé, avec une quantité non nulle pour chaque produit. Ensuite, la description de la liste des définitions des messages échangés indépendamment de l'implantation du service et du protocole utilisé. Le document décrit deux messages pour l'interaction : le premier message est la requête reçue par le service, l'autre est l'accusé de réception renvoyé. Les valeurs fournies pour les attributs element sont des noms de types XML Schema ou définis dans la partie types ci-dessus. Les opérations offertes par le service sont exposées par le biais de points d'entrée. Un point d'entrée (élément portType) fournit la signature de chaque opération et doit par la suite être associé à une implantation particulière (voir plus loin la description de la partie binding). WSDL permet l'existence de plusieurs points d'entrée dans un même document. Ainsi, la même opération peut être rendue disponible à travers des implantations différentes.

Chapitre 2 : Background

Dans la description ci-dessus les valeurs des attributs message font référence à un message défini plus haut (élément message du document WSDL). La définition de la signature d'une opération s'appuie sur trois sous-éléments possibles : input pour le message en entrée, output pour celui en sortie et fault pour un message d'erreur émis par le service. Les combinaisons possibles de ces sous-éléments permettent de décrire divers modèles d'opérations :

- Réception de message (sous-élément input) : le service reçoit un message envoyé par un client ;
- Réception - émission (sous-élément input suivi de output et éventuellement de fault) : le service reçoit une requête sous forme d'un message puis émet une réponse sous forme d'un retour de message. C'est le modèle classique RPC (Remote Procedure Call) ;

-Emission - retour (sous-élément output suivi de input et éventuellement de fault) : le service envoie une requête à un client (suite aucune inscription préalable) sous forme d'un message ; le client répond sous forme d'un message de retour.

-Emission (sous-élément output) : le service envoie un message (d'alerte par exemple) à un client. La dernière partie du document fixe la mise en correspondance entre chaque point d'entrée (un ensemble d'opérations) et son implantation, et permet de définir quels services sont associés à quelle mise en correspondance.

2.2.3.2 Description comportementale

La portée de WSDL est limitée à la description des types des données incluses dans les messages qu'un service est capable de recevoir ou d'émettre. Dans de nombreuses applications, ces descriptions uniquement structurelles n'offrent pas assez d'information sur les contraintes régissant les interactions dans lesquelles un service peut ou est censé s'engager. Dans certains cas, ces contraintes sont assez simples, comme, par exemple : « le fournisseur n'envoie le bordereau de livraison qu'après avoir reçu la confirmation du paiement ». D'autres fois, ces contraintes peuvent être relativement complexes. Quelques travaux existent dans ce domaine, par exemple :

- le langage de description de processus métiers exécutables des services Web (Web Services Business Process Execution Language, BPEL) (Andrews et al., 2003),
- le langage de description de chorégraphies de services Web (Web Services Choreography Definition Language, WS-CDL) (Kavantzias, 2004).

BPEL est un langage basé sur XML, son but est de décrire les interfaces comportementales, ainsi que des orchestrations complètement exécutables. WS-CDL tente d'aller plus loin que BPEL, en

Chapitre 2 : Background

s'attaquant non seulement à la description d'interfaces comportementales, mais aussi à la description de chorégraphies à différents niveaux de détails, allant de la modélisation conceptuelle jusqu'à l'implantation.

2.2.3.3 Description des aspects non-fonctionnels

Les standards WSDL et BPEL ne permettent pas de décrire des aspects non-fonctionnels des services tels que la qualité de services, la sécurité, la fiabilité, le coût, etc. Cette lacune de description est en partie comblée par d'autres standards proposés comme, WS-Policy, WS-Security-Policy, WSReliable-Messaging, WS-Addressing, etc. ainsi que d'autres approches spécifiques. Par exemple le WS-Policy est un langage extensible permettant d'exprimer des politiques (ou règles) d'usage sous forme de conjonctions et de disjonctions (au sens logique) d'assertions. Dans ce contexte, une assertion est une donnée par laquelle un service exprime qu'il permet aux clients (ou qu'il requiert des clients) de procéder d'une certaine manière lorsqu'ils accèdent aux opérations du service. WS-Security-Policy, WSReliable-Messaging et WS-Addressing, définissent respectivement des types d'assertion liés à la sécurité, la fiabilité, et l'adressage.

2.2.4 Synthèse sur les services web

Dans une entreprise des activités d'intégration d'applications par l'utilisation des services web, exige un effort humain prédominant par exemple pour la découverte de services web. Cela est dû à l'absence de connaissances, indispensables pour l'automatisation des services, ou à leurs descriptions adéquates seulement pour l'usage humain. Il semble donc nécessaire de passer à des services compréhensibles par des machines : c'est la technologie des services web sémantiques, qui s'appuie principalement sur l'enrichissement de service web avec un modèle sémantique en particulier l'ontologie. La section suivante présente la notion d'ontologie, ce qui va nous permettre d'aborder le concept de service web sémantique(SWS) dans la section (2.4).

2.3 Ontologies

2.3.1 Introduction

La recherche sur les machines intelligentes a pour objet la production de systèmes qui peuvent raisonner avec les connaissances disponibles et ainsi se comporter intelligemment (c.-à-d. des systèmes capables de détecter, de comprendre, d'agir et d'apprendre), afin d'automatiser certaines tâches (basées sur le savoir et le savoir-faire) qui occupent une part importante du temps des

Chapitre 2 : Background

humains. L'un des principaux problèmes est alors comment incorporer les connaissances dans ces systèmes ? Comment réduire cette notion abstraite (connaissance) dans des formes qui peuvent être décrites dans l'ordinateur ? On appelle cela le problème de la représentation des connaissances. En effet, de nombreux chercheurs dans le domaine de l'intelligence artificielle sont convaincus que la représentation de la connaissance est une étape clé dans la tentative de compréhension de l'intelligence. Il en est ainsi parce qu'une mauvaise représentation peut créer de nombreux problèmes dans la conception et l'exécution du système intelligent. Une nouvelle approche a émergé au début des années 90 est basée sur la notion d'ontologie. L'ontologie est une technique récente pour la représentation des connaissances. Elle est connue comme la technologie de base pour le web sémantique, qui est une vision future de développement de WWW (Berners-Lee et al., 2002) et elle est utilisée comme la spécification formelle d'un domaine de connaissance pour les descriptions SWS et la programmation sémantique en général.

2.3.2 Synthèse de représentation des connaissances avant l'ontologie

La notion de connaissance est essentielle pour un nombre important de domaines, par exemple, la philosophie, la psychologie, la logique et l'éducation. Même dans les mathématiques et la physique, depuis la formalisation mathématique des fondements de la physique, la communauté scientifique exige, que les lois de toutes natures doivent être décrites de manière similaire. La représentation des connaissances pose le problème de choix de formalisme de représentation qui nous permet de capturer la sémantique. À l'exception de la logique, l'impact révolutionnaire de l'avènement des ordinateurs a donné une nouvelle force de motivation, ainsi de nombreux chercheurs abordent maintenant le problème de représentation des connaissances. Brièvement voici quelques-uns des principaux formalismes de représentation qui ont résulté de leurs travaux:

La logique est un moyen naturel d'exprimer certaines notions. L'expression d'un problème dans la logique correspond à notre compréhension intuitive d'un domaine. Cela donne une dimension de clarté à la représentation. Un autre avantage de la logique est qu'elle est précise. La modularité et la flexibilité de la logique représentent un avantage significatif. L'inconvénient majeur de la logique est que les preuves pour des problèmes réels ont tendance à être informatiquement infaisable. Donc, sa puissance de raisonnement est limitée par des contraintes pratiques. Son grand échec comme un système de représentation est son incapacité à représenter adéquatement le temps ou les concepts d'ordre supérieur comme la généralisation.

Les opposants de l'approche déclarative adhèrent à la représentation procédurale où la connaissance est intrinsèquement liée dans les routines et les procédures qui l'utilisent. Ces procédures et routines savent comment faire une tâche particulière qui serait considérée comme intelligente. Un système

Chapitre 2 : Background

célèbre d'intelligence artificielle qui utilise la connaissance procédurale est le fameux SHRDLU de Winograd. Deux problèmes majeurs qui se posent dans une approche procédurale concernent l'exhaustivité et la cohérence.

Les idées de l'association des relations comme l'héritage et d'autres relations sont au cœur des concepts de la connaissance. Les réseaux sémantiques sont un paradigme populaire qui reflète ces idées. Un réseau est constitué de nœuds représentant des objets, des concepts et des événements et des liens entre les nœuds représentant leurs relations. L'un des principaux problèmes avec les réseaux sémantiques c'est qu'au début, il n'y avait pas de distinction entre les individus et les classes d'objets. Plus tard plusieurs solutions ont été proposées. Le grand avantage des réseaux sémantiques est bien sûr que les associations importantes peuvent être faites explicitement et succinctement. L'un des grands inconvénients de ce système de représentation est qu'il n'y a pas de conventions standards sur l'interprétation formelle des réseaux sémantiques.

La représentation des connaissances sur les objets et les événements typiques à des situations spécifiques a fait l'objet de la recherche sur un concept appelé Frames (cadres). Les Frames ont été initialement proposés, en 1975, par Minsky, comme base pour la compréhension de la perception visuelle, des dialogues en langage naturel et d'autres comportements complexes. Une trame est une structure de données composite qui se compose d'un certain nombre de créneaux qui correspondent à différents aspects de l'objet représenté.

Enfin, un autre formalisme qui sont les Scripts. Un script est une représentation structurée décrivant une séquence stéréotypée des événements dans un contexte particulier. Initialement proposé par Schank et Ableson en 1977. Les scripts peuvent être utilisés pour répondre à des questions sur des situations particulières. Ils peuvent être utilisés pour résoudre les ambiguïtés référentielles. Parmi les inconvénients des scripts est le problème lorsque deux scripts ou plusieurs sont applicables à une situation donnée. Dans la suite, nous centrons notre présentation sur les ontologies, qui constituent un enjeu stratégique dans la représentation et la modélisation des connaissances. Dans la suite nous, centrons notre présentation sur la notion d'ontologie.

2.3.3 La notion d'ontologie

À l'origine, l'ontologie est une notion philosophique, qui représente une branche de la métaphysique qui désigne la science de l'être en général. Dans le domaine de l'IA, de façon moins ambitieuse, on ne considère que des ontologies, relatives aux différents domaines de connaissances. Dans différents contextes d'application, et au sein de différentes communautés, les ontologies ont été explorées à partir de différents points de vue, et il existe plusieurs définitions d'une ontologie. Des exemples de définitions les plus essentielles de l'ontologie sont présentés dans la Table 2.1.

Chapitre 2 : Background

T. R. Gruber: «Une ontologie est une spécification explicite d'une conceptualisation » (GRUBER, 1993)
« une ontologie est une spécification formelle d'une conceptualisation partagée » (Borstet, 1997)
Studer : « Une ontologie est une spécification formelle et explicite d'une conceptualisation partagée » (Studer et al., 1998)
M. Uschold, M. Gruninger: «Une ontologie est une compréhension commune de certains domaines d'intérêt. » (Uschold & Gruninger, 1996)
« Une ontologie peut prendre diverses formes, mais nécessairement elle comprendra un vocabulaire de termes, et certaines spécifications de leurs significations » (USCHOLD & KING, 1995)
Une d'ontologie peut être définie formellement comme une paire $O = (S, A)$, où S est La signature d'ontologie qui décrit le vocabulaire, et A est un ensemble axiomes, qui spécifient l'interprétation prévue d'un domaine particulier. (Kalfoglou & Schorlemmer, 2003)
Enfin, voici une "grande" définition formelle: Niepert & al.: «Une ontologie est un six-uplet, $O = (L, C, HC, RH, FC, FR)$. Avec: C : est l'ensemble des concepts, HC : une taxonomie induite sur les concepts, RH : ensemble de relations non-taxonomiques, L : ensemble de termes (lexicales) qui se réfèrent à des concepts et des relations, FC : une fonction de correspondance qui associée des termes de L à des concepts FR : une fonction de correspondance qui associée des termes de L à des relations Auquel on ajoute un élément A : un ensemble de d'axiomes ou de contraintes (déclarations avec des éléments de C et RH). (Niepert et al., 2008)

Table 2.1. Définitions choisies pour "L'ontologie".

Au sein de la communauté du Web sémantique, la définition la plus dominante d'une ontologie est celle de Tom Gruber qui est complétée par (Studer et al., 1998). On remarque que toutes ces définitions partagent des notions communes:

Chapitre 2 : Background

- Formelle : indique que l'ontologie est adéquate à une appréhension et une interprétation par la machine sans ambiguïté ;
- Explicite : indique que l'information sémantique est décrite d'une manière claire et explicite via la terminologie et les axiomes de l'ontologie;
- Conceptualisation : indique qu'une modélisation abstraite d'une partie du monde est réalisée avec des concepts adéquats;
- Partagée : indique qu'ontologie véhiculait une connaissance consensuelle, et elle n'est pas limitée à un usage individuel, par contre, elle est partagée par un groupe.

2.3.4 L'ontologie techniquement

Techniquement, les principaux constituants d'une ontologie sont des concepts, des relations et des instances. Les concepts correspondent aux nœuds génériques dans les réseaux sémantiques, ou aux prédicats unaires dans la logique, ou à des concepts comme dans les logiques de description. Ils représentent les catégories ontologiques qui sont pertinents dans le domaine d'intérêt. Les relations sont l'équivalent des arcs dans les réseaux sémantiques, ou aux prédicats binaires dans la logique, ou à des rôles dans les logiques de description. Ils connectent sémantiquement les concepts, ainsi que les instances, en précisant leurs interrelations. Les instances correspondent aux nœuds individus dans les réseaux sémantiques, ou à des constantes dans la logique. Ils représentent des objets concrets identifiables dans le domaine d'intérêt, à savoir les individus particuliers qui sont classés par concepts. Ces éléments constituent un vocabulaire ontologique. Une ontologie peut être considérée comme un ensemble de déclarations, exprimé en termes de ce vocabulaire, qui sont également appelés comme des axiomes. Un simple axiome serait, par exemple, affirmé que «monsieur X est un employé », impliquant une instance et un concept. Un axiome plus complexe pourrait indiquer que «seuls les employés d'une entreprise particulière peuvent être, sur les voyages réservés par cette société», imposer une restriction à une relation entre deux concepts. La modélisation conceptuelle avec les ontologies semble être très similaire à la modélisation en orientée objet ou à la conception de diagrammes entité-relation des schémas de base de données. Cependant, il y a une double différence subtile. Premièrement, les langages d'ontologie fournissent une sémantique formelle plus riche que les formalismes orientés objets ou ceux des bases de données. Ils prennent en charge le codage de l'information axiomatique complexe en raison de leurs notations basées sur la logique. Par conséquent, une ontologie spécifie une riche axiomatisation sémantique des connaissances de domaine plutôt qu'une simple donnée ou modèle objet. Deuxièmement, les ontologies sont généralement développées pour un but différent de celui des modèles orientés objet ou des diagrammes entité-relation. Tandis que les modèles orientés objet

Chapitre 2 : Background

décrivent la plupart des composants d'un système informatique à exécuter sur une machine et les diagrammes entité-relation décrivent un schéma pour le stockage de données, une ontologie capture des connaissances de domaine et permet un raisonnement sur ces dernières.

Au sein des systèmes d'information, les ontologies apparaissent dans différentes formes liées aux formes de représentation des connaissances déjà citées. L'ingénieur de connaissance considère une ontologie via une certaine visualisation graphique ou formelle, alors que pour le stockage ou le transfert, elle est codée dans un langage d'ontologie avec un certain format de sérialisation traitable par la machine. Un raisonneur, à son tour, interprète une ontologie comme un ensemble d'axiomes constituant une théorie logique.

2.3.5 Quelques cas d'utilisation des ontologies

Les ontologies sont employées d'une manière intensive dans divers champs comme l'ingénierie des connaissances, l'intelligence artificielle et d'autres applications relatives à la gestion des connaissances, la recherche d'informations, les données liées et le web sémantique. Les ontologies peuvent également être utilisées à des fins de documentation et de référence, ciblant des agents humains plutôt que des machines, pour l'échange de leurs spécifications. De cette façon, la documentation d'un domaine bénéficie de modèle de spécification précis à travers la sémantique formelle des langages d'ontologie. Plusieurs applications ont évolué et font usage des ontologies de différentes manières. Nous énumérons certains d'entre eux comme des exemples de la façon dont les applications peuvent tirer profit des spécifications formelles que les ontologies fournissent.

- **L'intégration de l'information**

Un domaine prometteur qui requiert les ontologies est leur utilisation pour l'intégration de sources d'informations hétérogènes. Souvent, différentes sources d'information stockent le même genre d'information, mais adhèrent à différents modèles de données. Une ontologie peut être utilisée pour la médiation entre les schémas des bases de données par exemple, permettant d'intégrer des informations provenant de sources hétérogènes et d'interpréter les données d'une source sous le schéma d'une autre. Permettant l'interrogation unifiée et le raisonnement à la fois sur les différentes sources d'informations.

- **Recherche d'informations**

Motivé par le succès et le rôle-clé de Google dans le Web. La recherche d'informations sur les documents Web est un domaine d'application par excellence des ontologies. L'idée derrière la recherche d'informations basée sur l'ontologie est d'augmenter la précision des résultats de la

Chapitre 2 : Background

recherche en tenant compte de l'information sémantique contenue dans les requêtes et les documents, remplaçant les mots-clés par des concepts ontologiques.

- **L'enrichissement sémantique pour la gestion du contenu**

Dans de nombreux domaines de calcul des données ont besoin d'être annotés avec des métadonnées à des fins diverses. Les ontologies fournissent un vocabulaire spécifique au domaine pour annoter les données avec des métadonnées. L'aspect formel des langages d'ontologie permet un traitement automatique de ces métadonnées et leur représentation dans une forme interprétable par machine.

- **Systemes experts et les systemes à base de connaissances**

Dans divers domaines, tels que le diagnostic médical ou des conseils juridiques dans la jurisprudence, il est souhaitable de simuler un expert de domaine qui peut répondre à des questions sophistiquées. Dans un système expert, cela est réalisé en incorporant une ontologie de domaine pour formaliser les connaissances d'experts. Des requêtes spécifiques au domaine peuvent être satisfaites par le raisonnement sur des connaissances hautement spécialisées.

2.3.6 Les ontologies dans le Web sémantique

Dans le contexte du Web sémantique, les ontologies jouent un rôle crucial. L'idée du Web sémantique est d'annoter du contenu Web par des métadonnées interprétables par la machine pour que les ordinateurs soient en mesure de traiter ce contenu au niveau sémantique. Les ontologies fournissent le vocabulaire de domaine en matière d'annotation sémantique. Les métas déclarations sur le contenu Web dans ces annotations se réfèrent à un modèle de domaine couramment utilisé en incluant les concepts, les relations et les instances d'une ontologie de domaine. L'aspect formel des langages d'ontologie permet de raisonner sur les annotations sémantiques de différentes sources, connecté à des connaissances de base dans le domaine d'intérêt. Les ontologies ne sont pas disponibles localement à un seul nœud mais réparti sur différents sites. Cela pose des contraintes supplémentaires sur l'utilisation des ontologies dans le Web sémantique, en tenant compte des connaissances distribuées.

2.3.7 Langages d'ontologie

Pour rendre les ontologies disponibles pour les systèmes d'information, différents langages de représentations de l'ontologie ont été conçus et proposés pour la normalisation. Les influences des différents milieux de recherche et des utilisateurs ayant des besoins multiples ont donné lieu à un

Chapitre 2 : Background

paysage complexe d'une multitude de langages. Il est encore un sujet ouvert stimulant des discussions animées dans la recherche actuelle ; quels langages les mieux adaptés, par exemple dans la façon dont ils peuvent être efficacement mis en œuvre, ou leurs aptitudes techniques à garantir l'interopérabilité sémantique. Dans cette section, nous donnons un aperçu des langages d'ontologie les plus connus et les plus utilisés.

2.3.7.1 KIF (Genesereth, 1995)

Il a été créé à l'origine par Michael Genesereth et d'autres participants dans le projet DARPA de partage des connaissances. KIF (Knowledge Interchange Format) est un langage conçu pour une utilisation dans l'échange de connaissances entre les systèmes informatiques hétérogènes. Est un langage basé sur les prédicats du premier ordre avec des extensions pour représenter des définitions et des méta-connaissances, la logique du premier ordre étant un langage de bas niveau pour l'expression d'ontologies. Une extension du langage KIF, ONTOLINGUA, est utilisée dans le serveur d'édition d'ontologies, Ontolingua du même nom. KIF dispose d'expressivité sémantique complète. Un inconvénient de ce langage est sa complexité de calcul trop élevée, et l'absence d'une normalisation.

2.3.7.2 RDF et RDF Schéma

Resource Description Framework (RDF) (Klyne et al. 2014) est un langage recommandé par le W3C pour représenter des informations sur les ressources du World Wide Web. Basé sur une syntaxe XML (Extended Markup Language) qui constitue déjà un standard, RDF permet de décrire des ressources Web en matière de ressources, propriétés et valeurs. Il est particulièrement destiné à la représentation des métadonnées sur les ressources web identifiables, tel que le titre et l'auteur d'une page Web, le sujet et le copyright d'un document électronique accessible à partir du Web ou la fonctionnalité et les conditions d'accès d'un service web. Abstraction faite des ressources Web récupérables ou traitables par voie électronique, une identification précise par un URI de référence est la seule exigence, RDF peut être utilisé pour représenter des informations quelconques. En ce sens, RDF peut servir comme un langage pour représenter la connaissance comme métadonnées sur les entités par exemple. Le langage de description RDF Schema (RDFS) (Brickley & Guha, 2000) est une extension RDF qui facilite la formulation de vocabulaire pour les métadonnées RDF. Alors que RDF est utilisé pour relier les ressources au moyen de propriétés, RDFS introduit les notions de ressources classes et leurs hiérarchies. L'utilisation combinée des deux (RDF et RDFS) est souvent

Chapitre 2 : Background

appelée comme RDF (S) et fournit un langage d'ontologie simple pour la modélisation conceptuelle avec une certaine capacité d'inférence.

Le W3C a adopté le langage RDF (Resource Description Framework) comme un des formalismes standards de représentation de connaissances sur le Web.

▪ **Outils logiciels pour RDF (S)**

Le langage RDF (S) est utilisé pour décrire les métadonnées. Un certain nombre d'outils sont disponibles pour l'édition visuelle et la manipulation par programmes des descriptions RDF (S). L'un des éditeurs visuels le plus utilisé pour RDF (S) est Protégé (NOY et al., 2000; PRO, 2002), bien qu'actuellement il orientait vers OWL. Protégé permet de naviguer et de modifier une hiérarchie de classes RDF (S), ainsi que la possibilité de remplissage d'un schéma RDF avec des instances en utilisant les formulaires de saisie personnalisés. D'autres éditeurs d'ontologies qui soutiennent RDF (S) comme WebODE (Corcho et al., 2002), OntoEdit (ONTOEDIT, 2004). Pour le traitement par programmes des descriptions RDF (S), une suite d'outils : Sesame et Jena, qui fournissent des bibliothèques de logiciels qui permettent aux développeurs de logiciels de traiter les descriptions RDF (S) dans leurs applications. Ils comprennent l'analyse et la sérialisation pour le format XML RDF, une représentation de l'objet en mémoire pour les descriptions RDF (S) ainsi que la persistance de la base de données et de fonctionnalités, y compris la possibilité d'interrogation et de raisonnement.

2.3.7.3 DAML + OIL (Horrocks, 2002)

DAML + OIL est un langage d'ontologie spécifiquement conçu pour une utilisation sur le Web; il exploite des standards Web (XML et RDF) existants, en ajoutant la rigueur formelle de la logique de description et les primitives ontologiques de l'orienté objet et les frames. Le langage OIL (Ontology Inference Layer or Ontology Interchange Language) a été fusionné avec le langage DAML (Darpa Agent Markup Language) (HENDLER & MCGUINNESS, 2001) pour former le DAML+OIL. DAML est conçu pour permettre l'expression d'ontologies dans une extension du langage RDF. Il offre les primitives usuelles d'une représentation à base de frames et utilise la syntaxe RDF. L'intégration de OIL rend possible les inférences compatibles avec les logiques de description, essentiellement les calculs de liens de subsomption.

Chapitre 2 : Background

2.3.7.4 OWL (McGuinness & Van Harmelen, 2004)

OWL (Ontology Web Language) a été normalisé par le consortium W3C comme un langage pour l'annotation sémantique du contenu web, dans une optique de rendre les ressources sur le Web aisément accessibles aux processus automatisés. Une combinaison de RDF/RDF-S et de DAML+OIL a permis l'émergence d'OWL comme, un langage standard de représentation des connaissances qui est largement accepté dans la communauté du Web sémantique. Une vision importante pour la conception d'OWL est le compromis entre expressivité de langage d'une part et l'évolutivité du raisonnement d'autre part. À cet effet, OWL possède trois variantes différentes, à savoir OWL-Lite, OWL-DL et OWL-Full, reflétant différents degrés d'expressivité. La conception d'OWL-Lite et d'OWL-DL a été considérablement influencée par les logiques de description, donc ces deux variantes correspondent respectivement aux dialectes de la logique de description SHIF(D) et SHOIN (D). Contrairement, OWL-Full, écarte la sémantique de logiques de description pour assurer la compatibilité avec RDF (S). Les variantes OWL basées sur la logique de description bénéficient des propriétés de calcul bien comprises et de la décidabilité de la logique de description, tandis qu'OWL-Full a montré son indécidabilité.

▪ Support logiciel pour OWL

Le fait qu'OWL soit techniquement construit sur RDF (S), certains outils spécifiques à RDF (S) peuvent facilement être utilisés, par exemple, pour l'analyse et la sérialisation dans le format OWL RDF / XML, tandis que d'autres outils ont également été mis à niveau pour supporter les versions OWL. L'éditeur d'ontologie Protégé (NOY et al., 2000 ; PRO, 2002). Supporte également OWL et il est livré avec une variété de plugins qui permettent la visualisation et la gestion des ontologies OWL. En plus des différentes vues graphiques des hiérarchies de classes et des propriétés explicites, il facilite l'édition visuelle des axiomes OWL et permet l'intégration d'outils de raisonnement. D'autres éditeurs visuels pour les ontologies OWL qui offrent la même fonctionnalité sont SWOOP (Kalyanpur et al., 2004), Altova Semantic Works (Altova, 2009) et TopBraid(COMPOSER, 2011). Pour la manipulation des ontologies OWL par des programmes, l'API OWL, ainsi que Jena peuvent être utilisées par les développeurs de logiciels pour traiter les descriptions OWL dans leurs applications. Ils fournissent des moyens pour l'analyse et la sérialisation des différents formats de syntaxe OWL et pour la manipulation en mémoire des ontologies. Comme, OWL est un langage de représentation de connaissance, le raisonnement est un facteur important, il y a un certain nombre de raisonneurs disponibles qui peuvent être utilisés pour interroger les ontologies OWL par rapport à la connaissance déduite ou pour vérifier leurs

Chapitre 2 : Background

cohérences. Les raisonneurs de la logique de description les plus courants dans le contexte du Web sémantique sont basés sur la méthode de tableau, et les systèmes qui prennent en charge le langage OWL sont Racer, FaCT et Pellet.

2.3.7.5 F-Logic

Frame logic (F-Logic) (Kifer et al., 1995) est un langage de base de données orientée objet déductive qui combine la sémantique déclarative et l'expressivité de la programmation logique avec les riches capacités de modélisation conceptuelle, des systèmes à base de frames. Les caractéristiques linguistiques les plus importantes de F-Logic comprennent l'identité des objets, des objets complexes, classes, l'héritage, les types polymorphes, les règles et les requêtes. Dans la suite, un exemple de description d'ontologie avec F-Logic :

$C1 :: C2$ signifie que le concept $C1$ est un sous-concept (direct ou indirect) de $C2$;

$I : C1$ signifie que I est une instance du concept $C2$;

$C1 [P \Rightarrow C2]$ signifie que les deux concepts $C1$ et $C2$ sont reliés par la propriété P ;

$C[A \Rightarrow V]$ signifie que le concept C dispose de l'attribut A ayant comme valeur V ;

$I1 < : I2$ signifie que, l'instance $I1$ fait partie de l'instance $I2$;

etc.

2.3.7.6 WSML

L'initiative WSMO (Web Services Modeling Ontology) (De Bruijn et al., 2006) vise à fournir un cadre global pour la manipulation des Services Web Sémantiques. WSMO est décrite en détail dans la section 2.3, alors qu'ici nous sommes préoccupés par un composant essentiel, le langage WSML. WSML (Web Services Modeling Language) est un langage pour décrire formellement les éléments définis dans le modèle conceptuel WSMO, fournissant la syntaxe et la sémantique formelle. WSML est particulièrement conçu pour décrire les services Web sémantiques et il n'est pas un simple langage d'ontologie. Outre les notions relatives aux ontologies, il fournit également des constructions spécifiques au SWS, comme "objectif", "service web", "interface", "chorégraphie" ou «capacité», ainsi que les différents aspects d'un service Web sémantique. WSML fournit également des moyens pour décrire les ontologies qui sont des pièces maîtresses dans WSMO pour annoter sémantiquement les services web.

▪ Support logiciel pour WSML

Chapitre 2 : Background

Le groupe de travail « WSMO working group », inclut le groupe WSML working group, dont les missions sont la recherche et le développement d'outils pour WSML et en général pour WSMO. Le site WSMO (www.wsmo.org) présente des liens vers plusieurs outils logiciels dans le cadre de WSMO. Par exemple WSML2Reasoner Framework, WSMO Studio, WSML Rule Reasoner, WSMO4J API.

2.4 Service web sémantique

2.4.1 Introduction

le but des services web sémantiques est de créer un web sémantique de services dont les propriétés, les capacités, les interfaces et les effets, sont décrits et annotés de manière non ambiguë et exploitable par des machines en utilisant des ontologies. Le web sémantique et les services web représentent une évolution naturelle du web dans deux directions, La première direction est sémantique et s'intéresse principalement aux moyens et technologies pour décrire de manière intelligible par la machine des informations statiques disponibles sur le Web. La préoccupation principale de la deuxième direction est l'interopérabilité entre applications via le Web en vue de rendre le Web plus dynamique. La convergence des deux évolutions du web a donné naissance au concept du Service Web Sémantique (en abrégé : SWS) (Weise et al., 2014 ; Plebani & Pernici, 2009 ; Klusch, 2008). Donc, les SWS sont un domaine de recherche qui se concentre sur la création des descriptions sémantiques pour les services Web, afin de faciliter l'usage automatique, telle que la découverte et la composition. À partir du Web sémantique, il emprunte le concept de l'utilisation des ontologies pour créer des modèles explicites et les introduire à la machine.

2.4.2 Description des services web sémantiques

WSDL décrit seulement les aspects fonctionnels et syntaxiques d'un service. Il ne fournit pas l'information comportementale ou non fonctionnelle des services. L'absence d'une description interprétable par la machine représente un obstacle pour l'automatisation des tâches, telle que la découverte, l'invocation, la composition et l'interopérabilité de services web. Afin de répondre à ces exigences, la communauté du Web sémantique propose des solutions reposant sur des ontologies pour fournir une description explicite de la sémantique des services web, lesquels sont ensuite appelés services Web sémantiques. L'objectif des services Web sémantiques est de faciliter les tâches liées à leur utilisation, telle que la découverte, la sélection, l'orchestration et l'invocation, par le biais de leurs descriptions qui rendent la sémantique explicite et compréhensible par les

Chapitre 2 : Background

machines. Deux tendances possibles pour aboutir à une description services Web sémantiques (Mrissa, 2007) :

- (i) La première tendance consiste à développer un langage complet pour décrire dans un seul bloc le service web et sa sémantique.
- (ii) La deuxième manière consiste à enrichir les descriptions existantes des services web avec de l'information sémantique.

De nombreux langages et architectures sont proposés afin de décrire les services web sémantiques. Les sections suivantes présentent les approches les plus populaires.

2.4.3 Langages de description sémantique

▪ Web Ontology Language for services Web (OWL-S)

OWL-S (Martin et al., 2004) (précédemment appelé DAML-S), signifie Ontology Web Language pour Services, c'est une ontologie OWL et un langage pour décrire formellement les services Web. Une description OWL-S se compose de trois éléments, comme il est illustré dans la *Figure 2.3*: le *service profile*, le *grounding* et le *process model*, lesquels décrivent respectivement *< que fait le service >*, *< comment le service fonctionne >* et *< comment accéder au service >* :

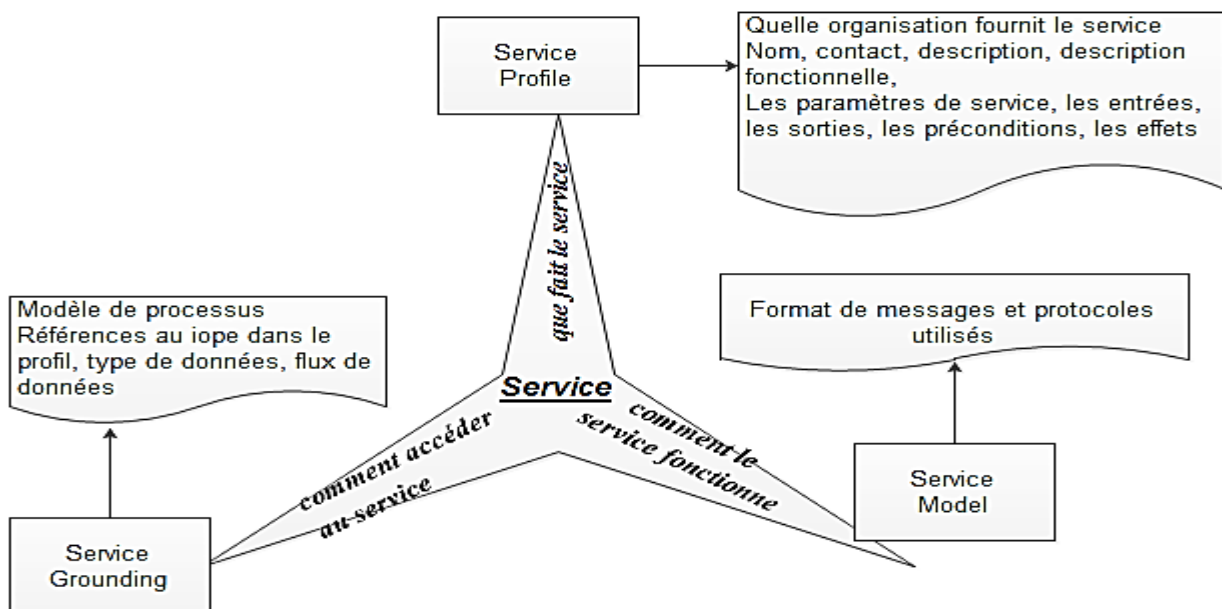


Figure 2.3: Vue d'ensemble de l'ontologie service OWL-S

Chapitre 2 : Background

- Le *service profile* décrit les fonctionnalités des services Web, il est utile pour leur découverte et leur sélection.
- Le *process model* détaille la sémantique des données échangées, au niveau des messages échangés entre services Web.
- Le *grounding* spécifie l'encodage des données échangées, les protocoles de communication, ainsi que tous les détails concrets nécessaires à l'invocation du service.

Suivant les idées développées par les auteurs d'OWL-S, de nombreux travaux de recherche ont été proposés. Notamment, ODESW (Gómez-Pérez et al., 2004) est une architecture de description et de composition des services Web reposant sur les méthodes de résolution de problèmes, ou < Problem-Solving Methods > (PSM). Dans cette architecture, une fonctionnalité est représentée comme un problème et la description des fonctionnalités offertes par le service permet d'évaluer les possibilités de résolution du problème.

▪ **Web Service Modeling Ontology (WSMO)**

L'architecture Web Service Modeling Ontology (WSMO) (De Bruijn et al., 2006) définit les éléments d'une modélisation pour décrire des services Web sémantiques, en se basant sur une conceptualisation mise en place dans le cadre Web Service Modeling Framework (WSMF) (Fensel & Bussler, 2002), dans lequel quatre éléments principaux ont été définis:

– **Les services Web**

Chaque service Web représente un bloc computationnel atomique de fonctionnalité, qui peut être réutilisé pour construire des processus plus complexes. Les services Web sont décrits dans WSMO sous trois angles différents: les propriétés non fonctionnelles, la fonctionnalité et le comportement.

– **Les Objectifs**

Les objectifs précisent les objectifs de fonctionnalités requises par le client lors de la consultation d'un service Web. Un objectif décrit la fonctionnalité, les entrées/sorties, les prés-conditions et les post conditions d'un service Web. La coexistence des objectifs et des services Web comme des entités différentes assure le découplage entre la requête et le service Web. Ce genre de situation, dans lequel le demandeur formule des objectifs sans tenir compte des services Web pour une résolution connue sous le nom d'approche axée sur les buts, dérivé de l'approche de l'agent rationnel AI.

– **Les Ontologies**

Chapitre 2 : Background

Les ontologies représentent des éléments-clés dans WSMO, car elles fournissent les terminologies (domaine spécifique connaissance), pour décrire les autres éléments. Elles servent un double objectif: définir la sémantique formelle de l'information et faire la liaison entre les terminologies humaines et celle de la machine.

– Les Médiateurs

Les médiateurs décrivent des éléments qui visent à surmonter les disparités qui apparaissent entre les différents composants qui accumulent une description de WSMO. L'existence de médiateurs permet de lier les ressources éventuellement hétérogènes. Ils résolvent les incompatibilités qui se posent à différents niveaux:

- Niveau de données : la médiation entre les différentes terminologies utilisées, plus spécifiquement la résolution du problème d'intégration d'ontologies.
- Niveau du processus : la médiation entre les modes de communication hétérogènes. Ce type d'hétérogénéité apparaît lors de la communication entre les services Web.

2.4.4 Annotation du langage de description

Les entreprises font des investissements dans des projets d'intégration basés sur les services Web. Par conséquent, une bonne approche pour l'ajout de la sémantique aux services Web devrait être spécifiée d'une manière compatible vers le haut afin de ne pas perturber les anciennes installations de services Web et les investissements associés à la formation humaine et aux solutions techniques. Également un mécanisme pour annoter les services web sémantique doit être indépendant du langage de représentation sémantique, du fait qu'il existe plusieurs langages pour représenter la sémantique comme OWL, WSML et UML d'une part et, d'autre part, chaque langage offre de différents niveaux d'expressivité sémantique et des différents supports développeur. Conclusion faite est de construire sur des normes de services web existantes. La section suivante est brève présentation de la première approche WSDL-S. Comme notre travail se focalise sur les services SAWDSL, l'approche sera présentée en détail dans la section (2.3.6).

➤ WSDL-S (WSDL-Semantic) (Akkiraju et al., 2005)

IBM et l'Université de Géorgie ont conjointement développé une approche nommée WSDL-S qui vise à fournir une approche < allégée > d'annotation sémantique de services Web. Les auteurs proposent l'ajout des annotations sémantiques aux éléments d'un document WSDL, en fournissant un mécanisme capable de faire référence aux URI des concepts d'un modèle sémantique externe aux descriptions. En bref, les extensions fournies dans cette approche sont:

Chapitre 2 : Background

-L'attribut *wssem: modelReference*, pour permettre des associations un-à-un, entre les éléments de schéma WSDL et des concepts dans un modèle sémantique.

-l'attribut *wssem: schemaMapping*, pour permettre des associations d'éléments de schéma d'entrée WSDL et aux concepts dans un modèle sémantique.

-l'attribut *wssem: precondition* et *wssem: effect*, qui sont des éléments qui spécifient le fonctionnement et la description sémantique de l'opération

-l'attribut *wssem: serviceCategorization*. Il se compose d'informations de catégorisation de service qui pourrait être utilisé lors de la publication d'un service dans un registre des services Web tels qu'UDDI. Il correspond à la notion de catégorisation proposée dans OWL-S.

2.4.5 SAWSDL (Semantic Annotations for Web Services Description Language)

SAWSDL (Semantic Annotations for WSDL) (Farrell & Lausen, 2007) est le plus populaire Framework pour l'annotation sémantique de WSDL, devenu une recommandation W3C. *SAWSDL* est conçu comme une extension du langage WSDL. Cette proposition étend WSDL-S, elle peut être considérée comme une continuité de ce langage et un moyen d'annoter les descriptions (WSDL 2.0) tout en supportant (WSDL 1.1.). Permettant aux fournisseurs de services d'enrichir les descriptions de service avec des informations sémantiques supplémentaires. À cet effet, les notions de *modelreference* et (*LiftingSchemaMapping* et *LoweringSchemaMapping*) ont été mises en place en matière d'attributs XML (tags) qui peuvent être ajoutés aux éléments WSDL existant dans une description de service, y compris des définitions de schéma XML pour les paramètres des messages comme, le montrent les figures *Figure 2.4* et *Figure 2.5*.

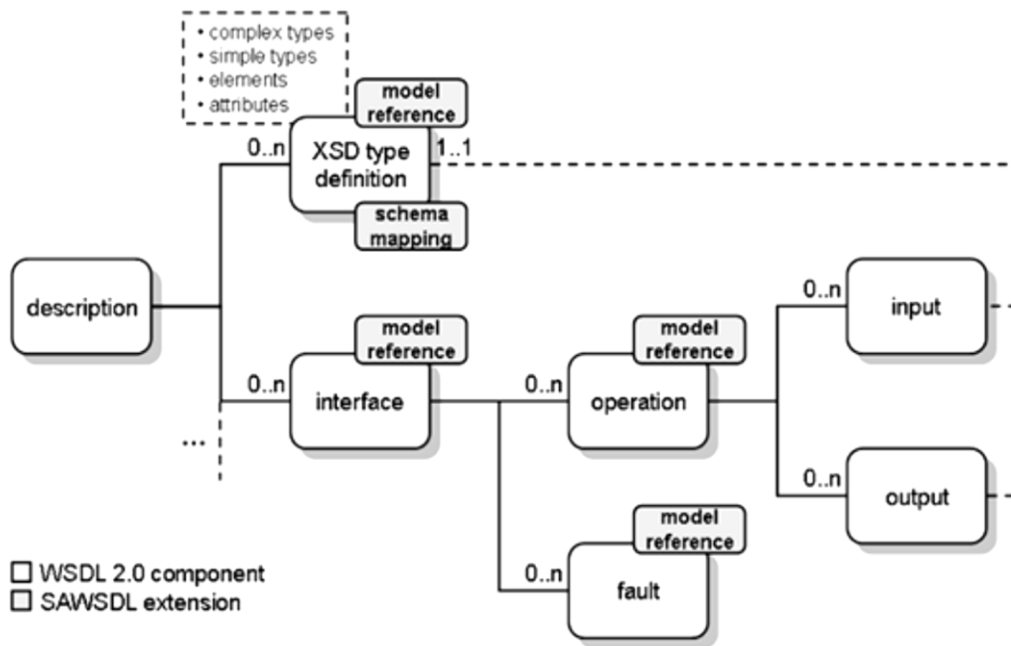


Figure 2.4 : Méta modèle SAWSDL (Klusch et al., 2009).

Les annotations sémantiques des services WSDL. Plus précisément, les extensions suivantes sont utilisées pour les annotations sémantiques des services WSDL:

- *ModelReference*: L'attribut *ModelReference* pointe sur un ou plusieurs concepts avec une intention exprimée dans n'importe quel langage de représentation de sémantique. La norme SAWSDL spécifiait que l'on peut annoter par les attributs *ModelReference*, les interfaces (ou les *portTypes*) WSDL, des opérations, des défauts ainsi que des éléments de schéma XML, les types complexes, les types simples et des attributs. La proposition de l'attribut *ModelReference* est essentiellement pour le support de l'appariement automatique des services. La spécification SAWSDL ne présente aucune restriction sur le type du concept sémantique. La seule exigence est que les concepts sont identifiables par un URI de référence. Dans le contexte de notre travail, nous supposons que les annotations sont des concepts des ontologies OWL. Comme, deuxième contrainte, notre outil de matchmaking (SAWSDL-MOM) prend en considération seulement le premier *ModelReference* d'un élément.

- (*LiftingSchemaMapping*, *LoweringSchemaMapping*): les mappings de schémas sont destinés à soutenir l'exécution automatique du service en fournissant des règles précisant les correspondances entre les concepts d'annotation sémantique définis dans une ontologie donnée (le niveau "haut") et

Chapitre 2 : Background

la représentation XML Schema de données réellement nécessaires pour invoquer le service Web (le niveau "inférieur"), et vice versa.

- ✓ l'attribut *liftingSchemaMapping* décrit la transformation du niveau "inférieur" du schéma XML vers l'ontologie utilisée pour l'annotation sémantique.
- ✓ l'attribut *loweringSchemaMapping* décrit la transformation du niveau "haut" d'une ontologie donnée au niveau «inférieur» dans le schéma XML.

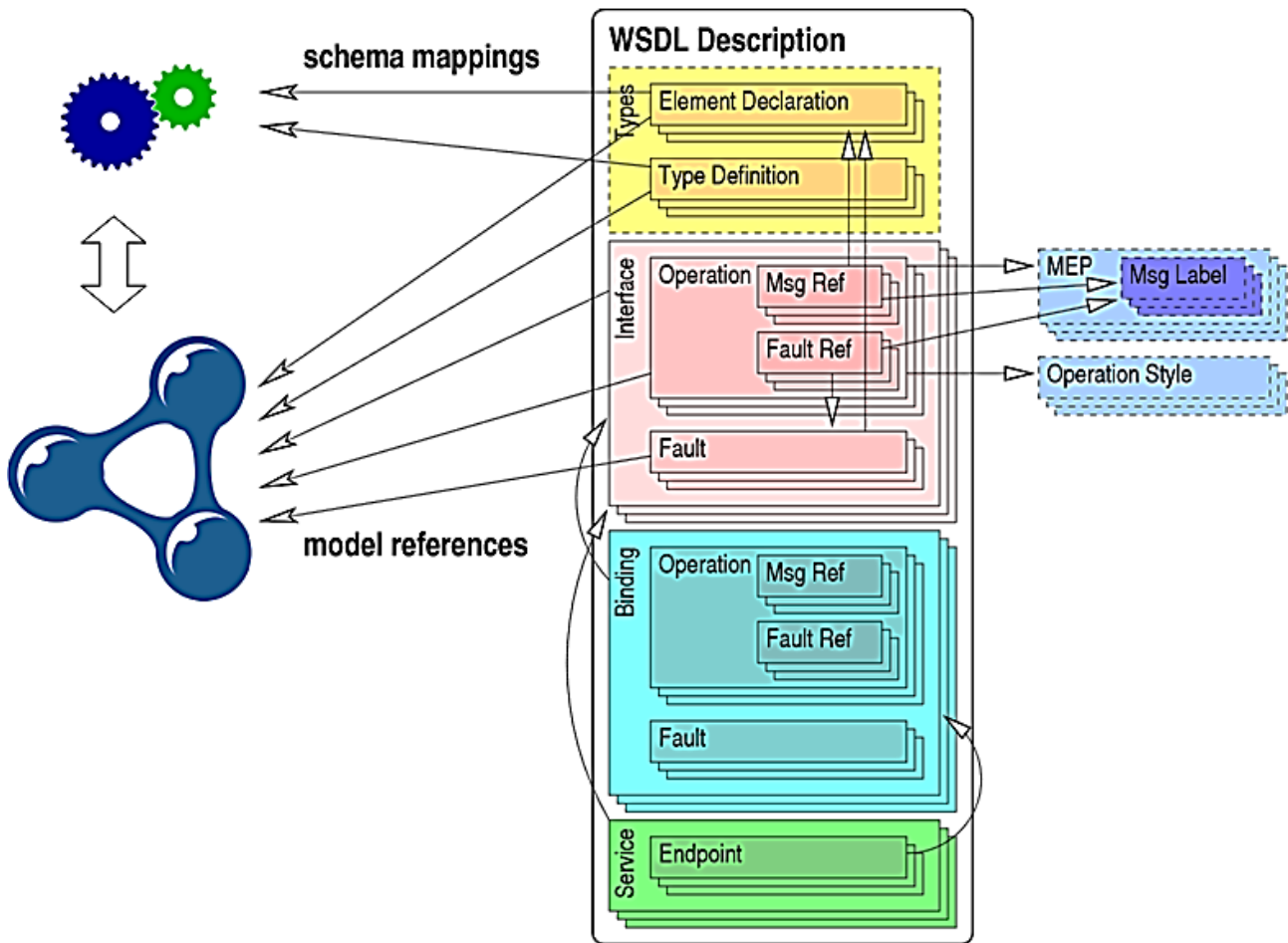


Figure 2.5 : Annotation sémantique de WSDL (Farrell & Lausen, 2007).

La description suivante est un fragment SAWSDL qui utilise l'ontologie de Rosetta (Farrell & Lausen, 2007).

```
<wsdl:operation name="Order" pattern="http://www.w3.org/2006/01/wsdl/in-out"
sawSDL:modelReference="http://www.w3.org/2002/ws/sawSDL/spec/ontology/
purchaseOrder#RequestPurchaseOrder">
<wsdl:input element="OrderRequest" />
<wsdl:output element="OrderResponse" />
```

Chapitre 2 : Background

```
</wsdl:operation>
```

L'annotation de l'opération nommée **order** est un `modelReference` qui pointe vers le concept **RequestPurchaseOrder** dans l'ontologie **PurchaseOrder**. Cette ontologie est librement basée sur la norme RosettaNet pour le commerce électronique, qui comprend des opérations bien définies. Les aspects d'une opération incluent les entrées et les sorties et même les pré conditions et les effets (Les conditions préalables et les effets font partie du WSDL-S, mais pas de SAWSDL pour plus de simplicité).

Ensuite, nous examinons les annotations relatives à l'entrée de l'opération **order**.

```
<xs:element name="OrderRequest"
sawSDL:modelReference="http://www.w3.org/2002/ws/sawSDL/spec/ontology/
PurchaseOrder#OrderRequest"
sawSDL:loweringSchemaMapping="http://www.w3.org/2002/ws/sawSDL/spec/mapping
/RDFOnt2Request.xml">
<xs:complexType>
<xs:sequence>
<xs:element name="customerNo" type="xs:integer" />
<xs:element name="orderItem" type="item" minOccurs="1" maxOccurs="unbounded" />
</xs:sequence>
</xs:complexType>
</xs:element>
```

L'élément **OrderRequest** est annoté avec le concept **OrderRequest** de l'ontologie.

Finalement, nous examinons les annotations de la sortie de l'opération:

```
<xs:element name="OrderResponse" type="confirmation" />
<xs:simpleType name="confirmation"
sawSDL:modelReference="http://www.w3.org/2002/ws/sawSDL/spec/ontology/
PurchaseOrder#OrderConfirmation">
<xs:restriction base="xs:string">
<xs:enumeration value="Confirmed" />
<xs:enumeration value="Pending" />
<xs:enumeration value="Rejected" />
</xs:restriction>
</xs:simpleType>
```

OrderRéponse est un élément qui est annoté avec **OrderConfirmation** de la même ontologie.

Toutefois, la spécification actuelle de références du modèle SAWSDL pose pas mal de problèmes pour le matchmaking sémantique des services comme suit :

Chapitre 2 : Background

- Pas de langage uniforme de l'ontologie formelle. Contrairement à OWL-S ou WSML, la spécification de SAWSDL ne limite pas le développeur à aucun langage formel d'ontologie, comme OWL ou tel que défini dans le cadre de WSML. En conséquence, tous les outils d'appariement sémantique de services web doivent faire face à des problèmes d'interopérabilité sémantique des ontologies hétérogènes et des langages d'ontologies. Bien que ce problème puisse être résolu dans certains cas, par le biais de transformations syntaxiques et sémantiques, comme pour OWL-DL et WSML-DL, il reste difficile en général.
- De multiples références aux différentes ontologies. En fait, SAWSDL permet de multiples références à différents types d'ontologies pour annoter le même élément d'une description du service. Plusieurs questions sur le comportement d'un agent de matching de service sémantique devant cette situation. Comment, faire pour comprendre la sémantique de cet élément unique ? Les annotations sont censées être complémentaires ou équivalentes ? Si complémentaires, de quelle façon on peut les agréger ? si elles sont équivalentes, laquelle est la meilleure à sélectionner pour un traitement ultérieur ? Cela ouvre un large éventail de différentes solutions pragmatiques pour l'appariement des services SAWSDL.
- En outre, il n'existe pas de collection de tests native pour SAWSDL disponible publiquement, mais une conversion semi-automatique en SAWSDL de la base de tests OWLS-TC d'OWL-S, pour évaluer les Matchmakers SAWSDL.

2.5 Conclusion

Afin de mieux cerner les concepts essentiels pour suivre le reste du document ; nous avons présenté dans ce chapitre un background théorique qui traite en premier lieu les concepts de base sur les services Web et les normes ou les technologies connexes. L'absence de connaissances, indispensables pour l'automatisation des services, exige un effort humain important. Donc, il est nécessaire de passer à des services compréhensibles par des machines. Une seconde partie présente les ontologies comme choix emprunté à partir du Web sémantique, pour créer des modèles explicites et les introduire à la machine. L'enrichissement de service web avec l'ontologie permet de passer à la technologie des services web sémantiques, la troisième partie de ce chapitre, présente le concept SWS, ainsi que les approches et langages de description comme OWL-S et OWSM. L'approche SAWSDL est présentée en détail car notre approche de matchmaking du chapitre 5, est spécifique à ce langage.

Chapitre 3 : Matchmaking des Services Web

Chapitre 3 : Matchmaking des Services Web - Etat de l'art

3.1 Introduction

La vision ultime de l'architecture orientée services SOA (Brown et al., 2002), est de permettre à un client de sélectionner automatiquement un service approprié et de l'invoquer sans n'avoir aucune connaissance a priori sur le fournisseur de services et les spécificités du service lui-même. Un défi majeur de cette vision est la découverte automatique de service. Pour surmonter ce challenge, un premier pas est l'enrichissement des descriptions syntaxiques de services avec la sémantique. Une fois les descriptions sémantiques (Services Web Sémantique : SWS) sont disponibles, il faut faire face à une autre gageure qui est l'appariement (matchmaking) de ces descriptions avec une requête. Dans ce contexte, diverses approches pour le matchmaking des SWS ont été proposées. Comme ces approches sont différentes, une solution peut être meilleure qu'une autre en fonction des besoins. Dans la suite, nous présentons les concepts inhérents aux matchmaking sémantiques des SWS et nous donnons un aperçu des différentes approches d'appariement décrites dans la littérature. Nous présentons, une étude comparative, dont le résultat nous aidera à identifier les mécanismes qui constituent les points forts des approches existantes, et par la suite servira de ligne directrice pour proposer de nouvelles approches d'appariement des SWS plus fiables et adéquates au contexte actuel des SWS.

3.2 Concepts généraux

Dans cette section, nous présentons les concepts de base essentiels et nécessaires pour comprendre le matchmaking des services web en particulier :

- (i) La découverte de services web
- (ii) Les dimensions de matchmaking
- (iii) Les caractéristiques principales d'un processus de matchmaking

3.2.1 Découverte de services web

Dans (Klusch, 2008), Klusch définit la découverte de service comme " le processus de localisation de services Web existant sur la base de la description de leur sémantique fonctionnelle et non fonctionnelle ". Il identifie trois aspects liés à cette tâche:

- Langage de description du service: le(s) langage(s) de description du service (par exemple, OWL-S et WSML, SAWSDL, OWL-DL ou WSML-DL) utilisés pour représenter la sémantique fonctionnelle et non fonctionnelle du service Web.

- L'appariement sémantique pour la sélection du service: la comparaison d'une requête avec un service pour déterminer le degré de leur correspondance sémantique.
- Architecture de découverte: concerne l'environnement dans lequel la découverte est supposée effectuer.

Le Scénarios de découverte se produit généralement quand on essaie de réutiliser des fonctionnalités existantes (représentées comme un service Web) dans la construction de nouvelles ou l'amélioration des processus métiers.

Nos travaux se concentrent principalement sur le second aspect, l'appariement sémantique (Semantic Matchmaking).

Un Matchmaker du service Web est un élément central de tout modèle de découverte de service Web. Il réalise la correspondance entre la requête utilisateur avec les services web disponibles et trouve l'ensemble de services web candidats qui peuvent satisfaire la requête.

3.2.2 Matchmaking

Le matchmaking des services Web (sémantiques), est le processus de recherche des services offerts appropriés pour répondre à une requête donnée (Klusch et al., 2009 ; Klusch, 2008). En conséquence, un moteur d'appariement est un outil qui met en œuvre un tel processus.

Appariement sémantique de service détermine si la sémantique d'un service requis (ou but) est conforme à celui d'un service offert. Cette question est le noyau de tout cadre de découverte de service sémantique. Selon Klush(2014), les approches actuelles d'appariement sémantique de services peuvent être classées en fonction de :

- ✓ Quels types et quelles parties de la sémantique de services sont considérés pour l'appariement :
La plupart des matchmakers de service Web sémantique réalisent un appariement de profil de service plutôt que de modèle de processus. Le Matching de service correspondant à "une boîte noire" qui détermine la correspondance sémantique entre les services, en se basant sur les descriptions de leurs profils. Le profil d'un service décrit ce qu'il fait en termes de sa signature, ses entrées et ses sorties (IO), ainsi que les conditions préalables (P) et les effets ou les post (E), et les aspects non fonctionnels, tels que la catégorie d'affaires, la qualité et les règles de tarification du service. L'appariement du modèle du processus est axé sur la détermination de la correspondance entre le comportement fonctionnel désiré d'un service donné en termes de flux de contrôle de processus et de données avec un autre service (voir *Figure 3.1*).

- ✓ Comment, l'appariement, est effectivement effectué (repose sur des techniques non logiques ou basées sur le raisonnement logique sur la sémantique du service ou une combinaison hybride des deux).
- ✓ Quelles sont les ressources utilisées pour atteindre les objectifs de l'appariement

Un aperçu schématique des dimensions dans lesquelles évoluent les appariements selon une exploration de Klusch est représenté par *Figure 3.1*.

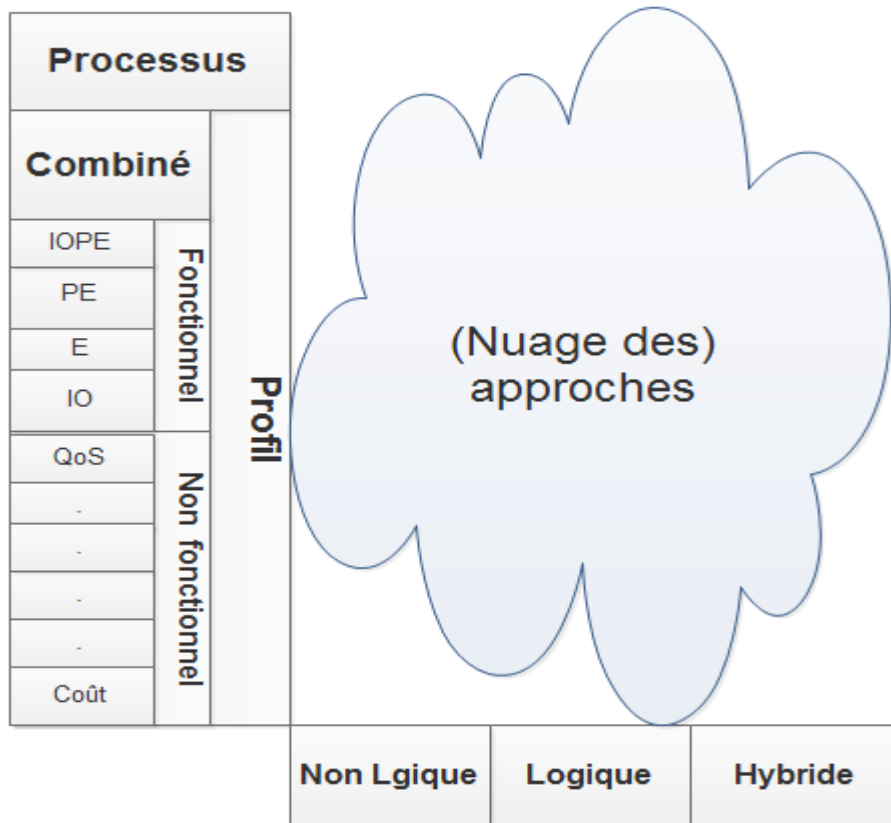


Figure 3.1 : Catégorisation des matchmakers des SWS

3.3 Processus de Matchmaking

Le but de cette section est de donner une définition précise du processus de matchmaking en général, et ses principales dimensions. La détermination de ces dimensions est très importante pour la caractérisation de ce que l'on sait ou encore pour les futurs algorithmes de matchmaking. Il devrait également être très utile dans la conception de tests d'évaluation et de comparaison des algorithmes similaires.

Définition (*processus de matchmaking*) Le processus de matchmaking peut être considéré comme une fonction f qui, à partir d'une requête R et un service S , un ensemble de paramètres p et un ensemble de ressources r , retourne un degré d'appariement DoM (DoM : Degree of Match):

$$\text{DoM} = f(\text{R}, \text{S}, \text{p}, \text{r})$$

Le processus de matchmaking consiste simplement à générer un degré de match à partir d'une description d'une requête et une description d'un service (R et S). Cependant, il existe divers autres paramètres qui peuvent étendre la définition du processus de matchmaking. Notamment, l'utilisation, de paramètres des méthodes d'appariement (qui peut être par exemple le poids) et quelques ressources externes utilisées par le processus de matchmaking (qui peuvent être des ressources à usage général, par exemple, lexiques, bases de données). Ce processus peut être défini par la *Figure 3.2* :

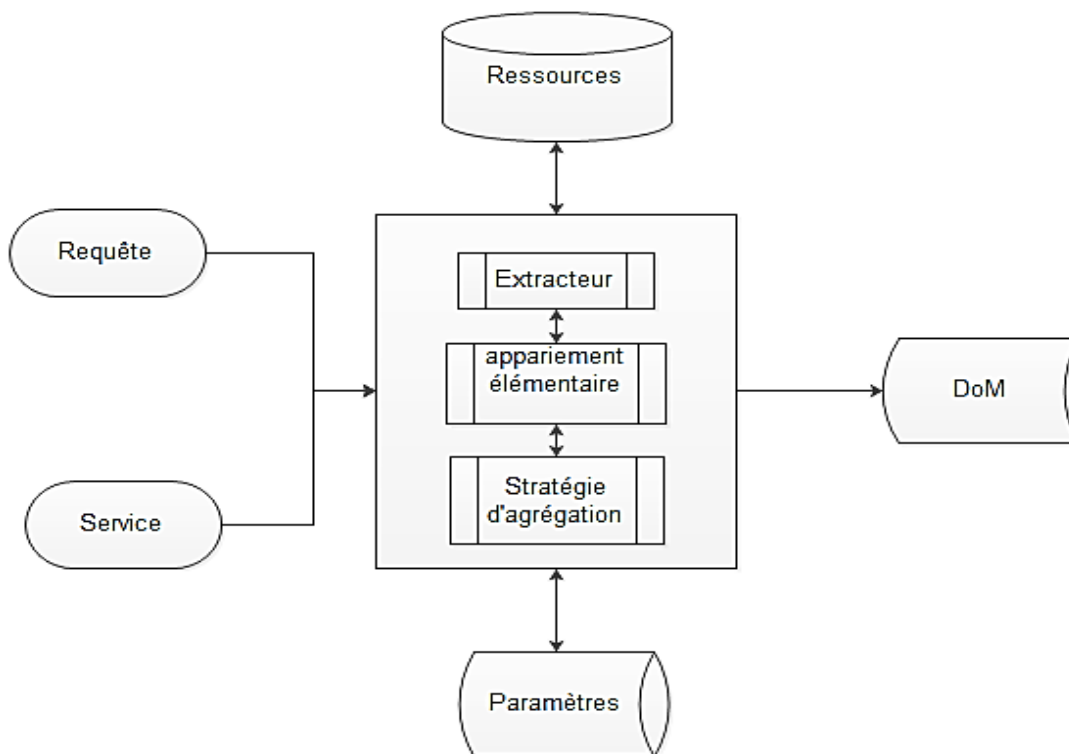


Figure 3.2 : Le processus général de matchmaking des SWS.

3.4. Dimensions d'un processus de matchmaking

Le matchmaking regroupe trois dimensions : l'entrée, le processus de Matchmaking et la sortie.

Également, on peut ajouter la dimension ressources et les paramètres éventuellement utilisés par le processus.

3.4.1 Entrée du Matchmaker

Essentiellement, les descriptions destinées à être appariées et qui peuvent être, une requête(**R**) et un service offert(**S**). Les descriptions peuvent avoir les contraintes suivantes:

- **Hétérogénéité** des langages de description (par exemple pour décrire un service (requis ou offert) on peut utiliser: le langage naturel, WSDL, SAWSDL, OWLS, USDL, BPEL, etc.). les

auteurs de (Baltá, & Fernández, 2010) ont présenté une tentative d'un modèle commun général pour l'alignement entre les langages de description de services.

- L'absence et/ou pauvreté sémantique de la description d'un service (requis ou offert).
- **Hétérogénéité** du modèle sémantique : il est possible, d'avoir une description annotée avec une ontologie et l'autre description annotée avec un autre modèle sémantique comme UML. Généralement, l'ontologie est le seul modèle sémantique utilisé.
- La portée du modèle sémantique au sein de la description du service (si tous les composants d'un service sont annotés sémantiquement ou juste une partie).
- **Hétérogénéité** du langage de description du modèle sémantique : même, si l'ontologie est le seul modèle utilisé, il existe plusieurs langages pour la décrire (par exemple : KIF, OWL, RDFS, UML, F-Logic, etc.)
- Les parties décrites : il est possible d'avoir dans une description, une partie (fonctionnelle, non fonctionnelle ou comportementale), ou la totalité des parties, cela dépend du choix des langages de descriptions.
- Les composants décrits dans une partie : par exemple, une description SAWSDL décrit la partie fonctionnelle d'un service web et n'inclut pas les pré-conditions et les post-conditions d'un service)

La plupart des matchmakers traitent exactement un seul standard service web particulier, en l'occurrence, OWL-S, SAWSDL, WSML. Quelques autres matchmakers appliquent d'autres standards pour formuler les capacités du service. Le choix d'un seul standard par les matchmakers a permis de contourner plusieurs difficultés.

3.4.2 Sortie du Matchmaker

En général, la sortie d'un matchmaker est un degré global de match (GDoM : Global Degree of Match ou tout simplement DoM) entre la requête et le service. La valeur du DoM peut être discrète (exacte) ou continue (numérique).

3.4.3 Ressources & paramètres

Une bonne mise au point de la stratégie d'usage des paramètres et/ou ressources doit être disponible :

- L'usage de ressources externes (r)
Est-il autorisé? Si oui, lesquelles? Et comment elles sont exploitées?
- Un apport humain Est-il autorisé? Si oui, à quels niveaux? Et comment?

- Les paramètres appropriés sont des éléments nécessaires? Lesquels ? Ce point est très important lorsqu'une méthode est très sensible à la variation des paramètres. (Par exemple, le seuil d'acceptation des appariements)

3.5 Caractéristiques du processus de matchmaking

Le processus de matchmaking est Le moteur d'appariement, qui est chargé de calculer le degré de correspondance entre la requête et le service. Il est difficile de cerner l'ensemble des particularités d'un processus matchmaking. Plusieurs facteurs peuvent le caractériser. Les auteurs de (Mohebbi et al.,2010) et (Chabeb, 2011), ont mis en avant quelques caractéristiques, que nous avons repris et ajouté d'autres:

- **Éléments pris en compte**

La description en entrée d'un service est composée de plusieurs parties et chaque partie peut contenir plusieurs éléments ; les matchmakers peuvent exploiter une partie ou la totalité des éléments pour détecter la similarité entre la requête et le service offert, généralement la signature est considérée. La qualité du résultat de l'algorithme de matchmaking est fortement influencée par les éléments pris en compte par le processus d'appariement entre la description du service Web requis et les descriptions des services Web offerts (par exemple dans SAWSDL : on prend les opérations, les inputs et les outputs ou bien quelques éléments seulement). Examiner lesquels des éléments de la description ont été gardés pour être appariés est primordial pour évaluer un certain aspect de la pertinence d'une approche.

- **Support Mono / Multi ontologie**

La plupart des matchmakers supposent que le service et la requête utilisent la même ontologie partagée pour la description. Cependant, dans un environnement véritablement distribué où les services sont autonomes cette hypothèse ne peut pas être vraie. Ainsi, un algorithme doit être capable d'effectuer des correspondances sémantiques à travers les descriptions avec des ontologies hétérogènes.

- **Appariement élémentaire (matching)**

Après identification des éléments potentiels prise en compte par le processus d'appariement, une première phase consiste à appairer les couples de composants. La seconde phase est la recherche des meilleurs appariements entre composants, c.-à-d. la sélection des correspondances optimales en termes de degré d'appariement, cette situation est connue sous le nom du problème d'affectation (Kuhn, 2010). Dans le cas des services SAWSDL, les éléments sont les interfaces, opérations, et les paramètres des messages, le choix de la mesure

peut être une simple mesure syntaxique, une subsomption logique ou une approche hybride. Un appariement élémentaire est appelé sémantique lorsqu'il s'appuie sur les concepts ontologiques. Ci-dessus, nous avons discuté le support (mono ou multi ontologie(s)) d'un algorithme de matchmaking. Dans le premier cas, une mesure de similarité sémantique mono ontologie est nécessaire (examine les diverses relations qui existent entre les concepts dans l'ontologie). Cependant, dans le cas où les inputs par exemple seraient annotés avec des ontologies différentes alors le matchmaker doit prendre en compte l'aspect interopérabilité et intégrer un mécanisme plus complexe. Pour appairer les éléments qui sont des concepts de différentes ontologies, une mesure inter ontologies est nécessaire. Également, les correspondances entre ontologies peuvent être générées par l'usage des techniques d'alignement des ontologies. Le chapitre 4, présente plus en détail les techniques d'appariement sémantique élémentaires.

- **Intervention des utilisateurs**

L'utilisateur doit avoir un certain contrôle sur le processus d'appariement et ne doit pas présenter un frein qui limite une approche de matchmaking. L'algorithme devrait donner la possibilité à l'utilisateur pour régler divers aspects tels que la flexibilité de l'algorithme, la qualité ou de la qualification qui est attendu pour le service défini par l'utilisateur, etc. L'utilisateur contribue à rendre le processus d'appariement plus adapté à ses besoins. Et l'exigence d'une assistance humaine est de plus en plus inacceptable dans un tel domaine (de plus en plus automatique). De plus, une interprétation humaine peut ne pas être assez pertinente, cela nécessite un bon niveau d'expertise. Une intervention humaine, même experte, peut-être plus coûteuse en temps, qu'une approche automatique bien élaborée (Chabeb, 2011).

- **Agrégation des résultats des appariements élémentaires**

Un important critère est l'agrégation des résultats des appariements élémentaires (ou partiels). Comme son nom l'indique, c'est la stratégie ou la manière avec laquelle les résultats partiels d'appariements sont combinés pour calculer le degré final/global d'appariement(DoM). Il faut préciser que souvent, une agrégation des résultats d'appariements dépend du type des degrés d'appariement attendu en sortie. Certaines approches utilisent des formules arithmétiques pour agréger des valeurs numériques (Chabeb, 2011), d'autres utilisent des stratégies plus particulières telles que la valeur minimale, la moyenne voire même des algorithmes complexes pour identifier le degré global.

- **La performance**

On ne peut améliorer que ce que l'on sait mesurer. Pour évaluer, on peut choisir parmi plusieurs indicateurs mesurables qui ont pour vocation de mesurer la performance d'un processus de matchmaking. Ainsi, les indicateurs comme la complexité effective en temps, la précision et le rappel sont des critères fondamentaux pour accepter ou rejeter un algorithme de matchmaking.

3.6 Les travaux de Matchmaking

Un grand nombre de matchmakers existe, Klusch et al. ont répertorié plus de 35 matchmakers des SWS (Klusch, 2008). À partir de nos constatations, nous allons présenter les approches d'appariement qui ont une relation avec notre travail. Dans ce travail, nous sommes préoccupés par l'utilisation de techniques d'alignement d'ontologies dans le matchmaking de services web sémantiques dans un contexte multi-ontologie. Vu nos objectifs, en premier lieu, on présente les travaux de base qui ont fortement influencé la communauté de recherche d'appariement des SWS. Ensuite, quelques travaux d'appariement dans un contexte ontologie unique. Puis on se focalise sur deux autres catégories de travaux qui sont importants pour notre travail à savoir, l'appariement dans un contexte multi-ontologie et les travaux qui concernaient les services web SAWSDL.

3.6.1 Travaux de base

Cette section introduit les contributions de bases qui représentent le fondement pour la plupart des autres matchmakers, même si le nombre d'approches d'appariement des SWS est très important, il y a que quelques contributions élémentaires d'appariement (sémantique) de services web ont été reprises et améliorées par plusieurs autres chercheurs. Pour une étude des matchmakers de services sémantiques en général, nous renvoyons le lecteur intéressé à (Klusch, 2008).

- **LARKS (“Language for Advertisement and Request for Knowledge Sharing”)(Sycara et al., 1999; Sycara et al., 2002)**

Est la première contribution qui a traité le matchmaking dans un contexte similaire à celui des SWS (Sycara et al., 1999), elle a fortement influencé la communauté de recherche de matchmaking des SWS. Les travaux les plus prestigieux sont basés sur LARKS. L'objectif principal de ce travail est la définition et la mise en œuvre d'un langage de description des capacités d'agents logiciels “Language for Advertisement and Request for Knowledge Sharing (LARKS)” et non pas sur les standards des services Web. En effet, LARK facilite la recherche et la mise en correspondance d'agents logiciels sur Internet. Cependant, l'approche est plus générale et peut s'appliquer à des composants logiciels ou encore à des services. Une spécification LARKS est un frame avec la structure suivante: *Contexte:*

Chapitre 3 : Matchmaking des Services Web

contexte de la spécification, *Types*: Déclaration des types de variables utilisées, *Entrée*: Déclaration des variables d'entrée, *Sortie*: Déclaration des variables de sortie, *InConstraints*: Contraintes sur les variables d'entrée, *OutConstraints*: contraintes sur les variables de sortie, *ConcDescriptions*: Description ontologique des termes utilisés.

Donc, les éléments d'une description peuvent également être trouvés d'une manière similaire dans (SA) WSDL ou OWL-S. Le processus d'appariement de Sycara et al. utilise cinq filtres:

1. Un filtre de contexte,
2. Un filtre de profil qui utilise TF-IDF (Term Frequency-Inverse Document Frequency).
3. Un filtre de similarité
4. Les filtres de signatures
5. Les filtres de contraintes

Les filtres se complètent et offrent différents degrés de précision, la sélection des filtres à appliquer est sous le contrôle de l'utilisateur (ou de l'agent demandeur).

Dans LARKS, les types de correspondances suivantes sont définis :

- correspondance exacte, où deux descriptions sont égales,
- correspondance plug-in, qui est essentiellement remplie si un service fourni plus que les fonctionnalités nécessaires,
- un match étendu qui est satisfait si la valeur de similarité pour les descriptions dépasse un seuil prédéfini. Contrairement aux autres matchs, le match étendu ne garantit pas un certain degré de compatibilité, mais "seulement" définit qu'un certain degré de similarité est donné.

Outre la création d'un fondement pour le matchmaking des SWS en général, d'autres aspects du matcher doivent être mentionnés explicitement: l'utilisation du matching de subsumption afin de déterminer les relations entre les concepts et l'approche hybride qui intègre à la fois la sémantique implicite et explicite.

- **Paolucci et al. (2002)**

Dans leur travail sur l'intégration des services DAML-S dans des registres UDDI. Les auteurs discutent le concept d'appariement des services Web sur la base de DAML-S (un prédécesseur de la norme OWL-S). L'appariement est restreint aux entrées et aux sorties (la signature de service). Les auteurs ont défini quatre niveaux : exact, plug in, subsume, et fail. Ces DoMs sont basés sur un appariement

logique de subsomption, à savoir, les relations hiérarchiques entre les concepts dans une ontologie. Un classement des DoMs est défini. Les auteurs présentent également le concept de DoM global pour deux services, qu'ils définissent comme le pire DoM trouvé durant le processus de matching. Une autre hypothèse est que si un service offert dispose d'une certaine sortie / entrée annotée avec un concept sémantique, toute sous-classe de ce concept l'est également. Le classement des services repose sur l'hypothèse qu'une requête favorise un DOM global meilleur pour les sorties plus qu'un DoM Global élevait pour les entrées. Par conséquent, les services sont d'abord triés par le DoM global des sorties. S'il y a deux services ou plus identiques, le DoM global des entrées est également pris en considération. Les auteurs appliquent une approche gourmande (*a greedy approach*) à l'appariement des composants de services. Cette approche gourmande prend le premier concept sémantique d'un ensemble (représentant les entrées ou sorties d'une requête) et détermine le concept correspondant possédant le plus haut Dom à l'ensemble des entrées (ou sorties) d'un service offert. Une amélioration intéressante de cette approche dans (Khater & Malki, 2014), les auteurs ont remplacé l'approche gloutonne par un algorithme du plus court chemin de Dijistra.

3.6.2 Matchmaking des services web dans un contexte ontologie unique

- **Li et Horrocks (2004)**

Bien que l'approche d'appariement dans le travail de Paolucci fût plus conceptuelle, une mise en œuvre rapide et une amélioration de leurs idées ont été menées par (Li & Horrocks, 2004). Dans ce travail, les services sont représentés par des constructions DL. Par conséquent, il est possible de procéder à l'appariement par inférence DL. Les auteurs présentent un moyen de modéliser les services basés sur DAML-S en utilisant des notions de DL; plus précisément, elle est limitée aux entrées et aux sorties, les autres éléments ne sont pas considérés. Par la suite, il est possible d'utiliser Racer DL comme raisonneur qui calcule les degrés de matching sémantiques entre les services offerts et les requêtes. Comme son nom l'indique, Racer DL est adapté pour les concepts sémantiques définis dans une logique de description. Li et Horrocks ont déterminé un DoM entre le service offert et la requête représentée par des paramètres d'entrée et de sortie. A savoir, un match est exacte si les concepts de l'offre et de la requête sont équivalents, *plug in*, si la requête est un sous-concept de l'offre, *subsume* si la requête est un super concept de l'offre *intersection* (non considérée par Paolucci et al.) si l'intersection entre la requête et l'offre est satisfiable, et *disjoint* autrement. En DL, une intersection logique de deux concepts est satisfiable s'il existe des individus appartenant à l'intersection des

individus des concepts. Encore une fois, les Doms sont ordonnés dans une échelle discrète avec exact comme le meilleur DoM et *disjoint* comme le pire.

- **OWLS-iMatcher (Kiefer, 2009)**

OWLS-iMatcher est une approche hybride qui accepte en entrée des descriptions OWL-S. L'algorithme utilise une technique d'appariement logique pour les inputs/outputs et une technique non logique qui utilise la similarité textuelle des noms et des signatures des services. OWLS-iMatcher entame le processus d'appariement par un calcul des valeurs de similarités syntaxiques entre une requête donnée et tous les services disponibles, ensuite, il utilise un modèle mathématique de régression pour prédire l'agrégation d'appariement pour chaque service. Les résultats sont renvoyés, à l'utilisateur, en ordre décroissant d'appariement. Les évaluations expérimentales OWLS-iMatcher ont été réalisées sur la collection de test OWLS-TC.

3.6.3 Matchmaking des services web dans un contexte multi-ontologie

Le groupe pionnier de chercheurs de l'Université Géorgie (Cardoso & Sheth, 2003 ; Oundhakar et al., 2005 ; Cardoso, 2006 ; Cardoso et al., 2008 ; Verma et al., 2007) a abordé ce problème. Ces travaux faisaient partie du projet METEOR-S. Plus tard, d'autres travaux ont porté sur ce problème. Le peu de tentatives sont résumées comme suit :

- **Cardoso et Sheth (2003)**

Ce travail a été réalisé dans cadre du projet METEOR-S. Cardoso et Sheth se sont intéressés par plusieurs aspects de la composition du workflow basé sur les services web, y compris l'appariement basé sur la sémantique des fonctionnalités, et des exigences non-fonctionnelles, et l'interopérabilité des services. Pour la découverte des fonctionnalités de services, à la fois des techniques syntaxiques et sémantiques sont utilisées. Les mesures de similarité syntaxiques sont appliquées aux noms et aux descriptions textuelles de service; les mesures sémantiques sont utilisées pour déterminer la similarité des entrées et des sorties. Toutefois, l'appariement est principalement mené pour l'intégration sémantique, donc, la recherche de la sortie d'un service qui peut être utilisée comme entrée pour un autre service. Comme contribution importante, Cardoso propose un appariement sans ontologie commune. À cet effet, une correspondance entre les classes de concepts provenant de différentes ontologies est effectuée et la distance géométrique entre la similarité entre les domaines des concepts est calculée. Le matching est basé sur la similarité syntaxique. La similarité est une valeur numérique basée sur les signatures des services offerte et du service requis.

- **Cardoso et al. (2008)**

Ce travail présente un algorithme de matchmaking des services web sémantique SAWSDL. L'algorithme est nouveau en trois aspects fondamentaux. Premièrement, la similarité entre les propriétés du service web sémantique, comme les entrées et les sorties, est évaluée à l'aide du modèle de Tversky lequel est basé sur des concepts (Classes), leurs relations sémantiques, et leurs caractéristiques communes et distinctives. Deuxièmement, l'algorithme ne prend en compte seulement les entrées et les sorties du service, mais il tient également compte de la fonctionnalité du service. Finalement, l'algorithme est capable d'opérer un appariement entre une requête de service et des services offerts qui sont annotés avec des concepts qui sont avec ou sans une ontologie commune. En d'autres termes, il peut évaluer la similarité des concepts définis dans un contexte multi-ontologie.

- **Oundhakar et al. (2005)**

Ce travail représente une extension de travail de Cardoso, par une amélioration de la mesure de similarité. Pour la description du service, les auteurs n'utilisent aucun langage standard et proposent un modèle sémantique pour la description du service. L'approche exploite les entrées, les sorties et les opérations. À la fin, un DoM numérique est proposé

- **Usanavasin et al. (2005)**

Propose une approche pour déterminer la similarité sémantique des propriétés des concepts de deux ontologies différentes. Pour résoudre le problème de matchmaking des services web dans un environnement multi-ontologie; les auteurs se sont focalisés sur le côté appariement des ontologies sans arriver à traitement du matchmaking des services.

- **MOD (Multi Ontology Discovery) (Le DuyNgan & Goh, 2006)**

Est un algorithme de découverte des services Web, qui supporte l'appariement de Services Web utilisant différentes ontologies. L'algorithme est composé de quatre étapes, l'appariement des entrées, des sorties, des opérations et un autre défini par l'utilisateur. Les auteurs proposent un algorithme de mise en correspondance des concepts de deux ontologies, la mesure est inspirée de (Oundhakar et al., 2005). Pour l'appariement des concepts, l'algorithme exploite les informations syntaxiques, les propriétés, le domaine et une similarité de voisinage. Un petit test de l'approche est opéré sur OWL-S. On pense que l'approche souffre d'une pauvreté expérimentale.

3.6.4 Approches de Matchmaking pour SAWSDL

À notre connaissance, le nombre de matchmakers sémantiques pour les services Web basés sur SAWSDL est assez faible. Dans ce qui suit quelques approches de matchmaking pour (SA) WSDL seront présentées :

- **Sivashanmugam et al. (2003)**

Dans le cadre des approches d'appariement des annotations sémantiques, pour WSDL, Sivashanmugam et al. ont présenté un mécanisme de découverte qui fait partie du projet METEOR-S. Même si le travail vise une version préliminaire du WSDL-S, il peut également être attribué à SAWSDL (Schulte, 2010). Le Matching est effectué pour les exigences fonctionnelles et la qualité de service. Pour l'appariement fonctionnel, le matching sémantique est effectué pour des opérations, les entrées, les sorties, les préconditions et les effets. Un poids est attribué à chacun des cinq éléments de service, afin d'évaluer le degré chaque partie sémantique pour le calcul de la parité globale de service. La valeur sémantique correspondante est une valeur numérique dans l'intervalle [0,1]. Si les concepts sémantiques dans le service offert et la requête sont identiques, la valeur d'appariement sémantique est 1; si les concepts ne sont pas identiques, une fonction linéaire est utilisée pour calculer la valeur de similarité sur la base d'une hiérarchie de subClassOf.

- **Syeda-Mahmood et al. (2005)**

Une autre approche d'appariement pour WSDL a été présentée par (Syeda-Mahmoodi et al., 2005). Ici, les auteurs combinent les valeurs de similarité en se basant sur des ontologies spécifiques à un domaine. L'appariement est limité aux entrées et aux sorties. Les synonymes sont détectés en utilisant WordNet. La sémantique spécifique au domaine est basée sur les modelReferences de WSDL-S qui sont semblables aux modelReferences défini dans SAWSDL. L'inférence est ensuite menée sur les annotations sémantiques, résultant dans l'une des relations possibles equivalentClass (0.0), subClassOf (0,5), superClassOf (0,5), et RDFType (0.0). En dehors de cette dernière, ces relations peuvent être retracées aux DOM définis par Paolucci et al. Ensuite, une valeur numérique pour les relations sémantiques est déterminée (voir parenthèses); La valeur est de 1,0 si aucune relation n'a pu être détectée. Le résultat global correspondant est la valeur maximale ou de valeur de similarité spécifique.

- **URBE (Plebani & Pernici, 2009)**

Le Matchmaker URBE (ou URBE/URBE-S) de (Plebani et Pernici) utilise une stratégie hybride qui exploite les informations linguistiques et sémantiques d'une description SAWSDL. Dans URBE, les différents niveaux d'abstraction de services sont pris en compte, à savoir les PortTypes (L'équivalent WSDL 1.1 à interfaces dans WSDL 2.0 (Christensen et al., 2001 ; Chinnici et al., 2006), les opérations et les entrées / sorties. La valeur de similarité globale d'une interface de service est la moyenne des valeurs de similarité des différents niveaux d'abstraction de service. Dans une extension, les auteurs présentent également une fonction de similarité sémantique, appelée URBE-S, qui remplace la mesure de similarité des noms si des annotations sémantiques sont disponibles. Cette fonction mesure la longueur du chemin entre les concepts dans une ontologie commune. Dans le système, l'URBE / URBE-S, malgré le mauvais délai de réponse, l'efficacité justifie l'usage de la sémantique dans les algorithmes d'appariement des services web.

- **SAWSDL-MX (Klusch et al., 2009)**

SAWSDL-MX accepte en entrée des services spécifiés en SAWSDL (Farrell & Lausen, 2007). Ce matchmaker est inspiré par les matchmakers OWLS-MX (Klusch & Kapahnke, 2009) et WSMO-MX (Klusch & Kaufer, 2009). Le processus d'appariement utilise une technique d'appariement logique basée sur la subsomption et une technique d'appariement basée sur les techniques de recherche d'information. L'appariement prend en compte les éléments de description suivants : interface/WSDL2.0 ; portType/WSDL 1.1, operation, input, output. Pour réaliser l'appariement des interfaces, le matchmaker effectue des appariements sur des graphes bipartis. Utilise les annotations sémantiques (ModelReference) pour la technique logique, les mesures syntaxiques (Loss-of-Information, Extended Jaccard , Cosine et Jensen-Shannon) ou une technique Hybride. Les degrés d'appariement calculés par l'approche logique : Exact, Plug-in, Subsumes et Subsumed-by et Les degrés d'appariement calculés par l'approche hybride :Subsumed-by (s'il a besoin de calcul de similarité syntaxique additionnelle), Nearest-neighbour.

- **LOG4SWS.KOM (Schulte, 2010)**

LOG4SWS.KOM vise à effectuer une correspondance logique de subsomption, qui prend compte les annotations sémantiques sur les différents niveaux de composantes d'une description de service SAWSDL. Ces composants comprennent des interfaces, Les opérations et les paramètres (c'est-à-dire les entrées et les sorties). Pour chaque niveau de composant, une valeur de similarité individuelle est calculée au cours du processus d'appariement, ensuite, elle est alors agrégée à une valeur de similarité

globale pour l'ensemble du service sur des poids prédéfinis pour chaque niveau. Les valeurs de similitude individuelle résultante sont transformées en représentations numériques entre 0 et 1 pour leurs combinaisons. Pour les composants, qui manquent d'annotations sémantiques LOG4SWS.KOM utilise une autre possibilité. L'alternative consiste à exploiter les informations syntaxiques, c'est-à-dire les noms des composants. La mesure de similarité est alors déterminée en utilisant l'ontologie WordNet.

- **COV4SWS.KOM (Schulte, 2010)**

Les auteurs ont proposé un matchmaker non logique. Il utilise des mesures de similarité basées sur les relations sémantiques dans une ontologie de domaine, à savoir les métriques de Lin (Lin, 1998) et Resnik (Resnik, 1995), pour démontrer que ces métriques ont aussi un avantage significatif par rapport au matchmaker logique LOG4SWS.KOM qui utilise le matching de subsumption. Il fournit nativement une similarité numérique sur une échelle continue. Ainsi, des métriques de relation sémantique peuvent être immédiatement intégrées, avec d'autres mesures de similarité. Cela est spécifiquement utile dans un processus de matchmaking (hybride) avec agrégation pondérée de similarité. Le matchmaker est auto adaptable à une expressivité variable d'une description de service à différents niveaux d'abstraction, par l'utilisation d'un estimateur des moindres carrés ordinaires (OLS : Ordinary least squares estimator) pour déterminer automatiquement le poids des différents niveaux d'abstraction des services.

3.7 Synthèse et conclusion

La table 3.1 présente une synthèse des travaux cités ci-dessus, la récapitulation est basée sur les dimensions du processus d'appariement. Les critères de classification sont l'entrée, la sortie et les paramètres utilisés par le processus de matchmaking, ainsi que ses caractéristiques.

3.7.1 Synthèse de l'état de l'art

Une constatation de l'état de l'art nous a permis de conclure que :

- LARKS (Sycaraet al., 1999 ; Sycaraet al., 2002) peut être considéré comme :
 - la première contribution qui concerne l'appariement dans le cadre des SWS.
 - un travail fondamental et de base pour l'appariement des services web sémantique ;

- le travail qui a fortement influencé la communauté de recherche d'appariement des SWS. Ainsi, les approches d'appariement de Paolucci et de Klusch sont basées sur LARKS.
- Paolucci et al. ont présenté une approche, qui est, à notre connaissance, le travail le plus cité dans ce domaine de recherche (Paolucci et al., 2002).
- Nous avons présenté les approches d'appariement qui ont une relation avec notre travail. Comme il a été mentionné précédemment, un grand nombre de matchmakers existe. Pas toutes les techniques utilisées par les matchmakers ont un intérêt pour notre travail, donc, la présentation a été limitée à des matchmakers qui ont influencé notre travail SAWSDL-MOM ou les travaux, qui ont un impact significatif et une visibilité dans la communauté de la recherche.
- La plupart des apparieurs négligent les aspects non-fonctionnels comme le QoS ou bien utilisent une sélection purement basée sur les paramètres non fonctionnels (Kritikos, 2008 ; Keskes, 2014).
- Cardoso et Sheth (Cardoso & Sheth., 2003), est la première contribution qui prend en compte l'aspect multi ontologie
- Les travaux de bases représentent le fondement pour la plupart des autres matchmakers
- Il n'existe aucun travail, générique, chaque approche accepte un modèle spécifique de service web,
- Aucune approche n'accepte l'appariement entre des modèles sémantiques différents.
- Absence d'une approche de matchmaking dans un contexte multi ontologie avec une expérimentation complète.

Chapitre 3 : Matchmaking des Services Web

Table 3.1: Synthèse des travaux de matchmaking

	Entrée	Sortie	Paramètres et Ressources	Processus			
				Eléments prise en compte	Mono ont	Multi-onto	appariement
LARKS	L.P	D		IOPE	X		Hybride
Paolucci et al	DAML-S	D		IO	X		Logique
Li et Horrocks	DAML-S	D		IO	X		Logique
OWLS-iMatcher	OWL-S	C		IOPE	X		Hybride
Cardoso et Sheth	L.P	D		IO	X	X	Non logique
(Oundhakar et al.	L.P	N		IO+Opération	X	X	Non logique
Cardoso et al.	SAWSDL	N		IO+Opération	X	X	Non logique
Usanavasin	N.S	N		IOPE(TH.)	X	X	Non logique
MOD	OWL-S	N		IO	X	X	Non logique
Sivashanmugam et al.	WSDL-S	N		IOPE	X		Hybride
Syeda-Mahmood	WSDL-S	D	WordNet	IO	X		Logique
URBE	SAWSDL		WordNet	IOPE	X		Hybride
SAWSDL-MX	SAWSDL	D		IO+Opération	X		Hybride
OWLS-MX	OWL-S	D		IOPE	X		Hybride
WSMO-MX	WSMO	D		IOPE	X		Hybride
LOG4SWS.KOM	SAWSDL			IO+Opération	X		Logique
COV4SWS.KOM	SAWSDL		WordNet	IO+Opération	X		Non logique
SAWSDL-MOM	SAWSDL	C	WordNet	IO+Opération	X	X	Non logique

L.P : langage propriétaire ; **N.S**: Non Spécifie ; **IO** : Input et Output ; **IOPE** : Input, Output, Pré condition et effet ; **D** : discret ; **C** : Continue.

Notre approche (SAWSDL-MOM) accepte comme entrée des descriptions SAWSDL et retourne un DoM numérique. Utilise un appariement non logique et supporte le matchmaking multi ontologie. Comme ressource externe, notre approche exploite WordNet.

3.7.2 Conclusion

Comme on peut le voir d'après le présent chapitre, la gamme des techniques possibles est très large et va de méthodes basées sur la logique aux techniques basées sur la similarité de R.I. Cette diversité, des approches est une explication pour la multitude de matchmakers qui ont été proposés pour différents standards de services, ces dernières années. Le Matchmaking de service basé sur l'information sémantique est compté par une communauté de recherche très agile, avec un grand nombre de différentes approches proposées au cours des dernières années. Une enquête sur quelques travaux, nous a permis de définir le matchmaking et ses dimensions, tout en identifiant les caractéristiques importantes du processus. Une synthèse de l'état de l'art a motivée la nécessité de développement d'approches pour le matchmaking des SWS dans un contexte multi ontologie. Également, l'appariement élémentaire entre composants de services est crucial pour la qualité des résultats d'un matchmaker. L'appariement élémentaire est qualifié de sémantique s'il exploite les annotations sémantiques. Dans le cas où la requête et le service seraient annotés par la même ontologie, on parle de Matchmaking dans un contexte ontologie unique, et des techniques d'appariement élémentaire ou similarité mono ontologie peuvent être utilisées. Cependant, dans le cas d'un contexte multi ontologies, des techniques d'interopérabilité sémantique sont nécessaires. Le chapitre suivant, présente un détail sur l'appariement sémantique élémentaire pour le matchmaking des services web sémantiques.

*Chapitre 4 : Appariement sémantique élémentaire dans le
Matchmaking des SWS*

Chapitre 4 : Appariement sémantique élémentaire dans le Matchmaking des SWS

4.1 Introduction

Le développement des services Web soulève ainsi des difficultés similaires à celles qui ont accompagné la croissance du Web. En effet, ces services sont développés par différentes entités et il n'existe pas de consensus sur la manière de faire usage des descriptions de services. L'ajout de la sémantique aux descriptions de service Web par l'utilisation des ontologies permet de rendre explicite, la sémantique implicite des structures dans les descriptions de service, qui sont connus seulement par l'auteur de la description. L'appariement sémantique de services détermine si la sémantique d'un service requis est conforme à celui d'un service offert. Cette question est le noyau de tout cadre de découverte sémantique de services web. L'appariement élémentaire des composants d'un service est la pierre angulaire de toute approche d'appariement sémantique de services web. Plus on réussit à appairer les éléments des services (requis et offerts), plus le résultat de Matchmaking est pertinent, et plus on identifie avec précision la similarité entre les concepts annotant les éléments des services (requis et offerts), plus le résultat de découverte est pertinent. L'aspect mono / multi ontologie(s) des SWS est un critère important pour le choix de la technique d'appariement sémantique élémentaire ou la mesure de similarité sémantique. Dans ce chapitre, on présente les bases de l'appariement sémantique élémentaire, la notion de mesure de similarité entre concepts intra ontologie et inter ontologies. Ensuite, on présente l'interopérabilité entre ontologies et l'alignement des ontologies. Enfin, l'alignement partiel est présenté.

4.2 Similarité sémantique

La similarité sémantique, c'est-à-dire l'aperception de la liaison entre deux concepts, est une capacité de l'homme que les machines ne savent que très mal reproduire. Ainsi, pour un humain, il est évident que les concepts (étudiant et université) sont liés, beaucoup plus que les concepts (neige et calculateur). Mais il est très difficile de le formaliser. Le problème de formalisation et de quantification de la notion intuitive de similarité sémantique entre les concepts a une longue histoire dans la philosophie, la psychologie et l'intelligence artificielle. La similarité sémantique entre les concepts est une mesure quantitative du lien, calculée sur la base des propriétés des concepts et de leurs relations. Avec l'avènement du Web sémantique, les mesures de similarité sémantique sont de plus en plus des composants importants de la Recherche d'Information (RI), extraction d'information (IE) et pour les systèmes à base de connaissances.

4.2.1 Définition

La similarité (Euzenat et al., 2004) est définie par le degré de ressemblance entre deux objets. En effet, tout système ayant pour but d'analyser ou d'organiser automatiquement un ensemble de données ou de connaissances doit utiliser, sous une forme ou une autre, un opérateur de similarité dont le but est d'établir les ressemblances ou les relations qui existent entre les informations manipulées. Il y a plusieurs manières d'évaluer la similarité entre deux entités. La manière la plus commune est d'associer une quantification à cette mesure de similarité. Dans ce qui suit, nous présentons les caractéristiques des mesures de similarités.

Définition plus formelle (Similarité). Une similarité $\sigma:O \times O \rightarrow \mathcal{R}$ est une fonction d'une paire d'entités à un nombre réel exprimant la similarité entre deux objets tels que:

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

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

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

La dissimilarité est sa fonction duale ;

Définition (Dissimilarité). Soit un ensemble O d'entités, une dissimilarité $\delta:O \times O \rightarrow \mathcal{R}$ est une fonction d'une paire d'entités à un nombre réel tels que:

$$\forall x,y \in O, \delta(x,y) \geq 0$$

$$\forall x \in O, \delta(x,x) = 0$$

$$\forall x,y \in O, \delta(x,y) = \delta(y,x)$$

Définition (Distance). Une distance (ou métrique) $\delta:O \times O \rightarrow \mathcal{R}$ est une fonction de dissimilarité qui satisfait :

$$\forall x,y \in O, \delta(x,y) = 0 \text{ Ssi } x=y$$

$$\forall x,y,z \in O, \delta(x,y) + \delta(y,z) \geq \delta(x,z)$$

4.2.2 Classification des mesures de similarité sémantique

Les systèmes de traitement de l'information ont besoin de sources de connaissance pour évaluer la similarité entre concepts. Dans les ontologies, l'utilisation essentielle de la mesure de similarité est de déterminer comment un concept de l'ontologie est similaire à un autre concept de la même ontologie ou d'une autre ontologie. Notre classification des mesures de similarité sémantique suivant cette logique : mesure de similarité intra ontologie (mesure de similarité dans un contexte ontologie unique) et mesure de similarité inter ontologies (mesure de similarité entre différentes ontologies).

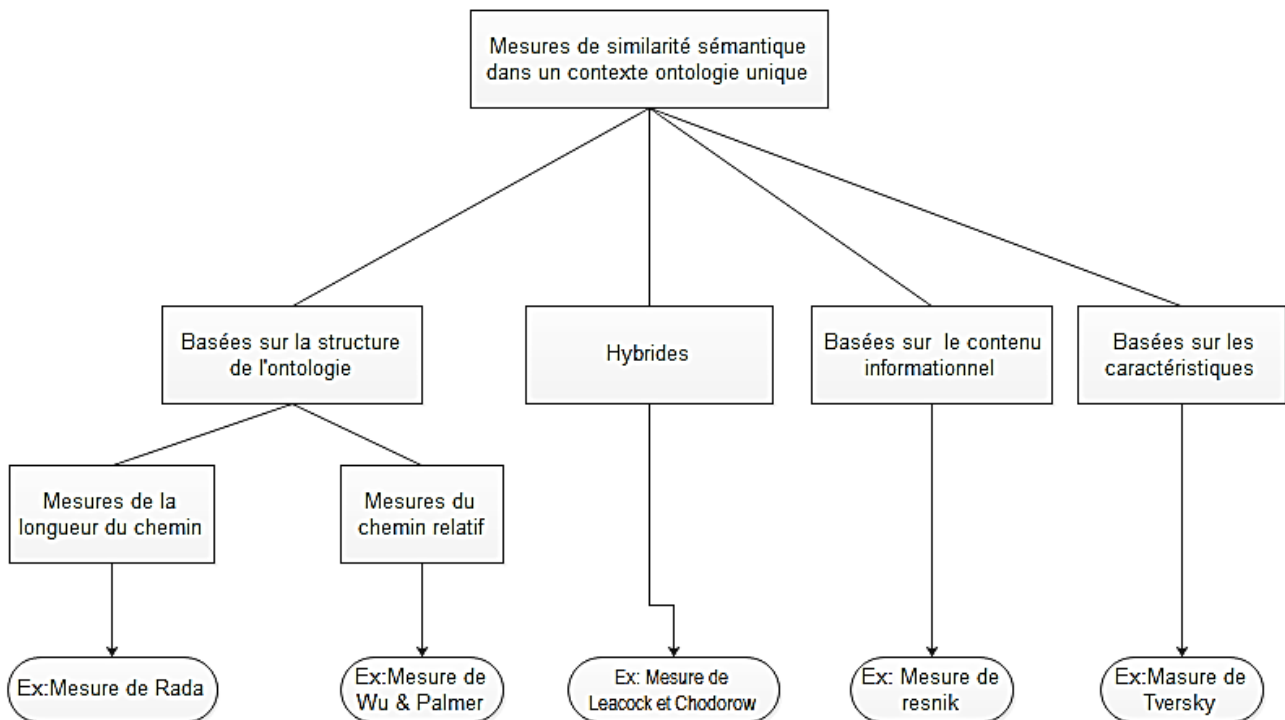


Figure 4.1: Classification des mesures de similarité intra ontologie.

4.2.2.1 Mesures de similarité sémantique intra ontologie

La figure 4.1 inspirée de (Abdelrahman & Kayed, 2015) et présente une classification basée sur la manière avec laquelle est quantifiée cette similarité, la similarité sémantique pourrait utiliser soit la longueur du chemin entre les concepts (mesure de Rada (Rada et al., 1989)) ou le contenu informationnel d'un concept (Mesure de Resnik (Resnik, 1995)) comme base pour la quantification. D'autres approches dites hybrides combinent, la distance du chemin et le contenu informationnel (mesure de Leacock et Chodorow (Leacock & Chodorow, 1998)). Une autre catégorie est basée sur les caractéristiques des concepts pour évaluer la similarité (mesure de Tversky (Tversky & Shafir, 2004)). L'élaboration d'une mesure de similarité sémantique complètement compatible avec évaluation humaine de la similarité est extrêmement difficile. La section suivante décrit les mesures de similarité inter ontologies (similarité entre concepts appartenant à différentes ontologies).

4.2.2.2 Mesures de similarité sémantique inter ontologies

Les mesures de similarité sémantique présentée dans la section précédente sont destinées à mesurer la similarité entre concepts appartenant à une même ontologie. Mais, avec la croissance des sources d'information, il existe un besoin pour le développement de méthodes qui calculent la similarité entre les concepts appartenant à différentes ontologies. Ces méthodes de calcul sont nécessaires pour résoudre les problèmes d'interopérabilité sémantique. La littérature rapporte des approches différentes pour mesurer la similarité entre concepts de différentes ontologies. Une classification, des approches disponibles pour mesurer la similarité entre les

différentes ontologies, est présentée dans la Figure 4.2. Les auteurs (Rodriguez & Egenhofer, 2003) ont proposé une mesure basée sur les caractéristiques et les auteurs (Al-Mubaid & Nguyen, 2009) ont proposé une mesure sur la longueur du chemin. Nous considérons une troisième catégorie de mesure de similarité entre ontologies, représentée par toutes les combinaisons de similarités choisies par les outils d'appariement des ontologies. Dans le présent travail, nous avons proposé une nouvelle approche pour le calcul de la similarité sémantique entre concepts de différentes ontologies.

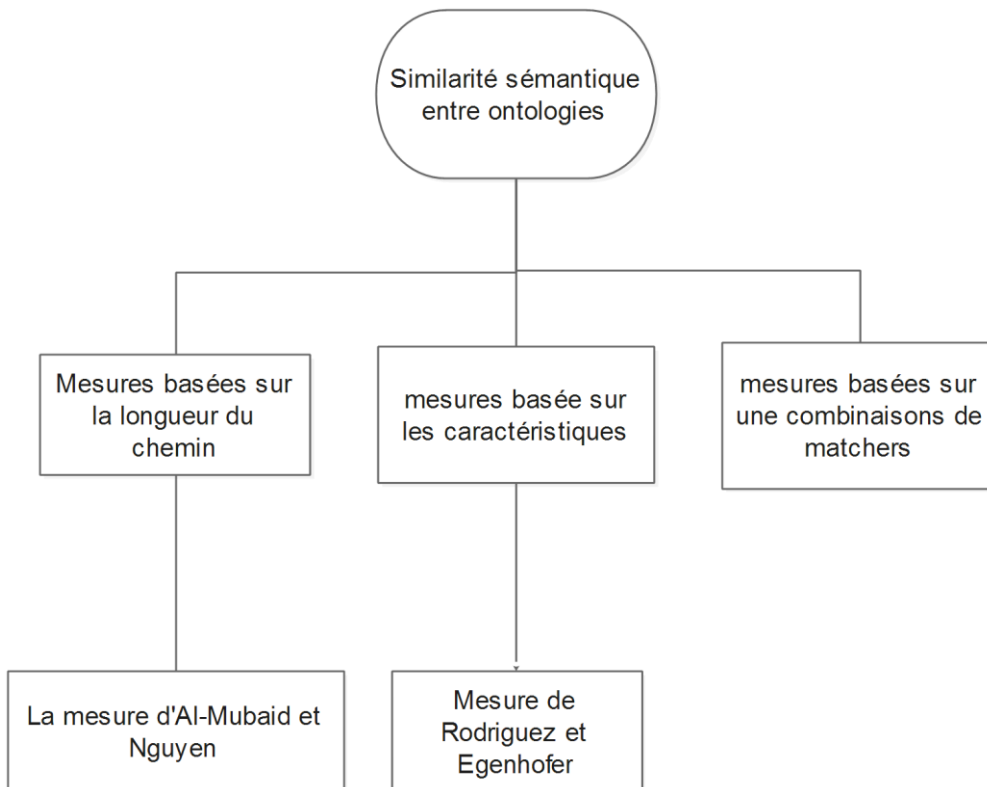


Figure 4.2: Classification des mesures de similarité inter ontologies

4.3 Techniques d'appariement sémantique élémentaire

L'aspect annotation mono / multi ontologie(s) des SWS est un critère important pour le choix de la technique d'appariement sémantique élémentaire ou la mesure de similarité sémantique, lors de l'appariement sémantique des services web. L'appariement élémentaire des concepts adopte la même logique de similarité entre concepts selon le contexte ontologique disponible (mono-ontologie ou multi-ontologie). Donc dans un contexte ontologie unique d'annotation, pour l'appariement élémentaire, une mesure de similarité mono ontologie est nécessaire (la considération des diverses relations qui existent entre les concepts dans l'ontologie). Par contre, dans le cas où la requête et le service offerts seraient annotés avec des ontologies différentes alors le matchmaker doit prendre en compte l'aspect interopérabilité et intègre un mécanisme plus complexe pour appairer les éléments qui sont des concepts de deux ontologies donc une mesure inter ontologies est indispensable, aussi l'usage

des outils d'alignement des ontologies, permet de générer les correspondances possibles entre les entités des ontologies.

4.4 Besoin d'interopérabilité sémantique

Dans un système réparti et ouvert (comme le web sémantique ou les services web sémantiques), l'hétérogénéité ne peut pas être évitée. Les acteurs ont différents intérêts et habitudes, utilisent différents outils, et emploient la connaissance à différents niveaux de détail. La première vague des travaux sur la découverte des services web s'est focalisée sur le cas lorsque les deux services (requis et offert) utilisent la même ontologie. Dans la littérature peu d'approches (Cardoso & Sheth, 2003; Oundhakar et al., 2005) considèrent le cas, lorsque deux services (requis et offert) utilisent différentes ontologies du même domaine. Cette deuxième situation impose le traitement de la problématique d'interopérabilité sémantique des services web sémantique. Une définition est communément admise pour l'interopérabilité sémantique : "elle donne un sens aux informations échangées et s'assure que ce sens est commun dans tous les systèmes entre lesquels des échanges doivent être mis en œuvre" (Jouanot, 2000). Le défi majeur de l'interopérabilité sémantique entre services (requis et offert), est l'hétérogénéité entre les ontologies. Ces diverses raisons d'hétérogénéité mènent à différentes formes d'hétérogénéité qui sont considérées ci-dessous :

4.4.1 Formes d'hétérogénéité

L'hétérogénéité peut se produire à différents niveaux, cependant, pour la définition d'un cadre commun, les auteurs dans (Euzenat et al., 2004) proposent une classification de quatre niveaux : **syntactique, terminologique, conceptuel, sémiotique/pragmatique.**

- **Le niveau syntaxique**

Au niveau syntaxique, nous rencontrons toutes les formes d'hétérogénéité qui dépendent du choix du format de représentation. En effet, il y a plusieurs formats proposés pour la représentation d'ontologie (par exemple : KIF, OWL, RDFS, WSML), et chacun d'eux est basé sur une syntaxe différente. On établissait des équivalences entre les constructeurs des différents langages, on peut traduire les ontologies entre divers langages d'ontologies en préservant le sens.

- **Le niveau terminologique**

Au niveau terminologique, nous rencontrons toutes les formes de disparités liées au processus de nommage des entités (par exemple individus, concepts, propriétés, relations), d'une ontologie. Le nommage est le processus d'associer un objet linguistique d'une langue (utilisée pour l'échange d'information avec d'autres parties) aux entités décrites dans une ontologie. Ce niveau ne devrait pas être confondu avec le niveau conceptuel (voir ci-dessous); en effet, les disparités terminologiques

peuvent se produire dans les situations où les ontologies impliquées sont conceptuellement équivalentes.

Les exemples des disparités au niveau terminologique sont:

- Mots différents sont employés pour appeler la même entité (**synonymie**) ;
- Le même mot est employé pour appeler différentes entités (**polysémie**) ;
- Mots de différentes langues, (anglais, français, italien, espagnol, allemand, grec, etc.) utilisés pour appelaient des entités;
- Variations syntaxiques du même mot (différentes appellations acceptables, abréviations, utilisation de préfixes ou de suffixes facultatifs, etc.).

Dans un sens, les disparités au niveau terminologique ne sont pas aussi profondes que celles qui se produisent au niveau conceptuel. Cependant, la plupart des vrais cas doivent se faire avec le niveau terminologique (les personnes ont différentes manières d'appeler les mêmes entités), et donc ce niveau est aussi crucial que les autres.

• **Le niveau conceptuel**

Au niveau conceptuel, nous rencontrons les disparités qui doivent faire avec la teneur d'une ontologie. Des anomalies à ce niveau peuvent être analysées dans deux classes principales:

- *Différences métaphysiques* expriment la façon dont le monde "est composé de morceaux" (c.-à-d., quelles entités, propriétés et relations sont représentées dans une ontologie);
- *Différences épistémiques* expriment les affirmations qui sont faites au sujet du choix d'entités.

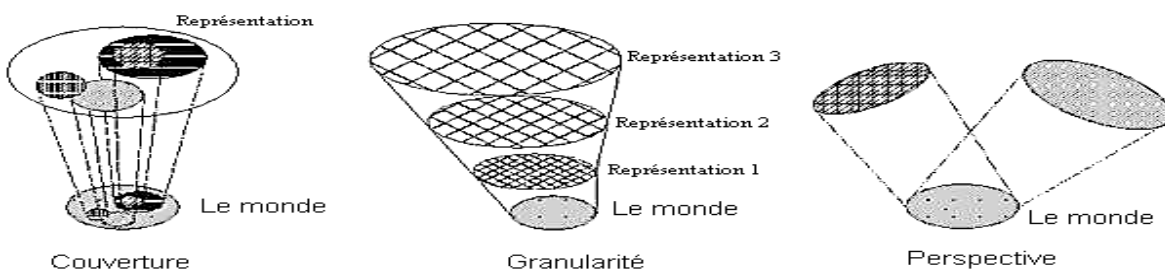


Figure 4.3 : L'hétérogénéité conceptuelle (Euzenat et al., 2004)

Ces deux genres de différences expliquent, par exemple, pourquoi différentes ontologies du même domaine peuvent commencer à partir de différentes classes primitives, ou pourquoi les différentes ontologies peuvent contenir différentes (probablement entités à peu près identiques) affirmations contradictoires. Car les différences épistémiques ne peuvent pas être traitées exclusivement à travers des correspondances.

Chapitre 4 :Appariement sémantique élémentaire dans le Matchmaking des SWS

Les formes pratiques dans lesquelles les différences métaphysiques peuvent surgir sont innombrables. Cependant, suivant la littérature d'intelligence artificielle dans cette matière (Euzenat et al., 2004) sont groupées dans trois types abstraits:

Couverture : une ontologie peut différer des autres pendant qu'elle couvre différentes parties recouvrant probablement le monde (ou même un domaine simple). Par exemple, une ontologie sur le sport peut inclure les courses d'automobiles, tandis que des autres peuvent décider de les ignorer en tant qu'éléments du domaine de sport; l'ontologie peut contenir des propriétés des courses d'automobiles que les autres négligent; et ainsi de suite.

Granularité : une ontologie peut différer des autres pendant que la première fournit plus (ou moins) de détail des descriptions des mêmes entités. Par exemple, des ontologies concernées par la comptabilité et les impôts, ou la livraison, considéreraient seulement le concept générique du document, alors qu'une ontologie pour des bibliothèques distinguerait les types de documents, par exemple livres, biographies ou autobiographies.

Perspective : Les manières individuelles de constater une situation, par exemple, influencées par ses expériences ou considérations. Une ontologie peut fournir un point de vue sur un certain domaine qui est différent du point de vue adopté dans une autre ontologie. Par exemple, deux ontologies peuvent représenter le même domaine au même niveau de couverture et de granularité, mais à différents points de temps. Ce qui signifie que la même propriété peut se tenir au moment où la première ontologie a été conçue et ne tenir pas au moment où l'autre a été conçue, sans vrai désaccord épistémique. Ou encore, d'une perspective spatiale différente. Une perspective qui est du côté droit d'un agent peut-être au côté gauche pour un autre agent faisant face à la direction opposée.

La Figure 4.3 fournit les représentations graphiques de ces trois dimensions dans lesquelles deux ontologies peuvent différer au niveau conceptuel.

- **Le niveau sémiotique ou pragmatique**

Au niveau sémiotique/pragmatique, on rencontre toutes les anomalies liées aux interprétations différentes d'un individu (communautés) pour la même ontologie de différentes manières, dans différents contextes. Par exemple, en contexte lié à la formalisation de la connaissance, un utilisateur peut exprimer la connaissance sous forme de hiérarchies de classes et de clauses du premier ordre et communiquer en employant un langage. Mais si ce dernier langage exprime toute la connaissance avec des clauses (avec préservation la sémantique des affirmations), l'utilisateur initial reconnaîtra à peine (et comprendra à peine) le résultat sémantiquement équivalent. Par conséquent, quand une transformation traduit entre les langages formels, la bonne compréhension ne peut pas être assurée par la conservation de la signification. Dans ce cas, aucune hétérogénéité syntaxique (parce qu'on permet des clauses dans le modèle initial) et aucune hétérogénéité terminologique ni conceptuelle :

seulement un manque d'interprétation de la représentation (équivalente) par son concepteur. L'utilisation prévue a un grand impact sur l'alignement, car il peut être tout à fait risqué d'apparier des entités sur d'autres seulement parce qu'elles sont sémantiquement reliées.

Par exemple, si le concept Europe apparaît dans le schéma de classification d'un dépôt multimédia d'un chemin comme Images/B&W/Europe, nous ne devons pas conclure qu'il est équivalent au concept de Europe dans une ontologie géographique, car la signification pragmatique déterminée de la première doit être un récipient des images noires et blanches de l'Europe, tandis que la signification prévue de la seconde est le continent elle-même (cela ne veut pas dire que les deux choses ne sont pas reliées, mais que les appariements devraient prévoir la prise en compte de chaque structure en considération).

4.4.2 Surmonter l'hétérogénéité

Pratiquement chaque acteur peut utiliser sa (ses) propre(s) ontologie(s) adéquate (s) aux concepts qui annotent les éléments de son service (service requis ou service offert), soit en créant une nouvelle ontologie, soit en réutilisant des ontologies existantes. Pour parvenir à l'interopérabilité sémantique, il faut donc trouver un accord entre toutes ces ontologies. Pour surmonter les problèmes de l'hétérogénéité entre des applications basées sur l'ontologie, il est crucial d'établir des correspondances (Mappings) entre leurs ontologies. La création de tels mappings manuellement est souvent impossible en raison de la taille et la complexité des ontologies. Par conséquent, le problème de la génération automatique de mappings entre ontologies (souvent appelé le problème de mise en correspondance des ontologies (Ontology Mapping), alignement des ontologies (Ontology Alignment) ou appariement des ontologies (Ontology Matching)) a été étudié de manière approfondie ces dernières années. Les correspondances peuvent être utilisées pour différentes tâches telles que fusionner des ontologies, produire des médiateurs, et en général dans toute situation nécessitant l'interopérabilité sémantique.

4.5 Alignement des ontologies comme solution à l'interopérabilité sémantique

L'alignement des ontologies est une tâche essentielle pour réaliser l'interopérabilité sémantique dans de nombreux domaines d'application. La génération manuelle des mappings est une tâche complexe et très difficile. La difficulté persiste même avec des outils d'éditions pour l'assistance de l'expert humain. Selon (Euzenat & Shvaiko, 2007), L'alignement d'ontologies est la mise en correspondance sémantique de leurs entités, il est basé sur un processus d'appariement qui utilise des matchers ou des mesures de similarité. Techniquement, il opère sur différents types d'informations, les noms des éléments, les types de données, la représentation de la structure d'éléments de modèles, caractéristiques des données, etc. Souvent une seule mesure est insuffisante pour détecter les liens sémantiques. Une stratégie de combinaison des matchers est appliquée ainsi que l'usage d'un ensemble de paramètres et de ressources pour obtenir un ensemble de mappings entre les entités des ontologies.

4.5.1 Terminologie

L'état de l'art confirme que, malgré les efforts consentis, les auteurs n'étaient pas parvenus à une terminologie commune. On constate que, dans le domaine d'interopérabilité sémantique des ontologies, différents auteurs, y compris nous-mêmes, utilisent des mots différents pour des concepts similaires et, inversement, des concepts parfois différents ont le même nom (Ehrig & Sure, 2004; Euzenat & Shvaiko, 2007; Euzenat et al., 2004 ; Klusch & Kaufer, 2009 ; Giunchiglia & Shvaiko, 2003 ; Valtchev et al., 2004 ; Xu et al., 2010 ; Cruz et al, 2009 ; Jean-Mary et al., 2009). Cela est particulièrement déroutant, car ces termes, (par exemple mapping, alignment), peuvent être utilisés pour décrire à la fois une action et son résultat. Dans cette section, nous proposons un glossaire de travail.

Appariement(Matching) : est le processus de recherche de relations ou de correspondances entre entités de différentes ontologies.

Alignement (Alignment) est un ensemble de correspondances entre deux ou plusieurs (dans le cas de plusieurs (Par analogie avec l'alignement des séquences moléculaires). L'alignement est la sortie du processus d'appariement.

Correspondance (Correspondance) est la relation tenue, ou censée tenir selon un algorithme d'appariement, entre entités d'ontologies différentes. Ces entités peuvent être aussi différentes que des classes, des individus, des propriétés ou des formules. Certains auteurs utilisent plutôt le terme mapping.

(Mapping) est la version orientée ou dirigée d'un alignement: elle mappe les entités d'une ontologie à au plus une entité d'une autre ontologie. Cela est conforme à la définition mathématique d'une (mapping) plutôt que celle d'une relation générale. La définition mathématique exigerait en principe que l'objet mappé est égale à son image, c'est-à-dire que la relation est une relation d'équivalence. Un mapping peut être considéré comme une collection de règles de mapping toutes orientées dans la même direction, c-à-d d'une ontologie à l'autre, et que les éléments de l'ontologie source apparaissent au plus une fois.

La règle de mapping est une correspondance qui mappe une entité d'une ontologie avec une autre entité d'une autre ontologie.

Notons que, dans plusieurs cas, dans la littérature spécialisée dans ce domaine, le terme appariement des ontologies (Ontology Matching) est utilisé pour désigner l'alignement ontologies (Ontology Alignment) ou la mise en correspondance des ontologies (Ontology Mapping) et les termes sont utilisés de façon interchangeable.

Chapitre 4 :Appariement sémantique élémentaire dans le Matchmaking des SWS

Suivant notre terminologie et constatations : L'alignement des ontologies, ou l'appariement des ontologies, est le processus de détermination des correspondances entre concepts. Un ensemble de correspondances est également appelé alignement. Processus d'appariement Prend comme entrée des ontologies et détermine comme sorties les relations (par exemple, Équivalence, subsomption) entre ces entités. Plus formellement, processus d'appariement détermine les correspondances entre les éléments des ontologies. Une correspondance est un 4-uple: (e, e', n, R) , où:

- e et e' sont respectivement les entités (par exemple, éléments XML, propriétés, classes) de la première et de la seconde ontologie;
- n est une mesure de confiance dans une structure mathématique (généralement dans la gamme $[0,1]$ indique de degré de correspondance entre les entités e et e' ;
- R est une relation (par exemple, l'équivalence ($=$), plus générale (\supseteq), la disjonction) entre les entités e et e' .

Un alignement est un ensemble de correspondances. L'opération d'appariement détermine l'alignement (A) pour une paire d'ontologies $(O$ et $O')$.

4.5.2 Les techniques de comparaison ou matchers

Pour la comparaison entre entités de différentes ontologies, on fait recours à un matcher. Un matcher est une fonction utilisée pour calculer la similarité entre deux entités. Les matchers sont des macros qui peuvent être combinées dans le processus de matching. On retrouve plusieurs méthodes de calcul de la similarité entre les entités de plusieurs ontologies. Dans la présentation des techniques de base, en fait abstraction des hétérogénéités des langages de descriptions des ontologies et l'on suppose que les deux ontologies sont décrites dans le même formalisme de représentation. Les différentes méthodes de comparaison (matchers), utilisées dans le processus de matching sont organisées selon la classification ci-dessous (*Figure 4.4*) (Euzenat et al., 2004). Les différents matchers sont classés suivant leur entrée, c'est-à-dire le type d'entités à comparer et la méthode utilisée pour calculer la similarité. Par exemple, un des matchers syntaxiques compare deux entités (ou deux concepts) suivant l'égalité de leurs noms (par exemple, les labels des concepts).

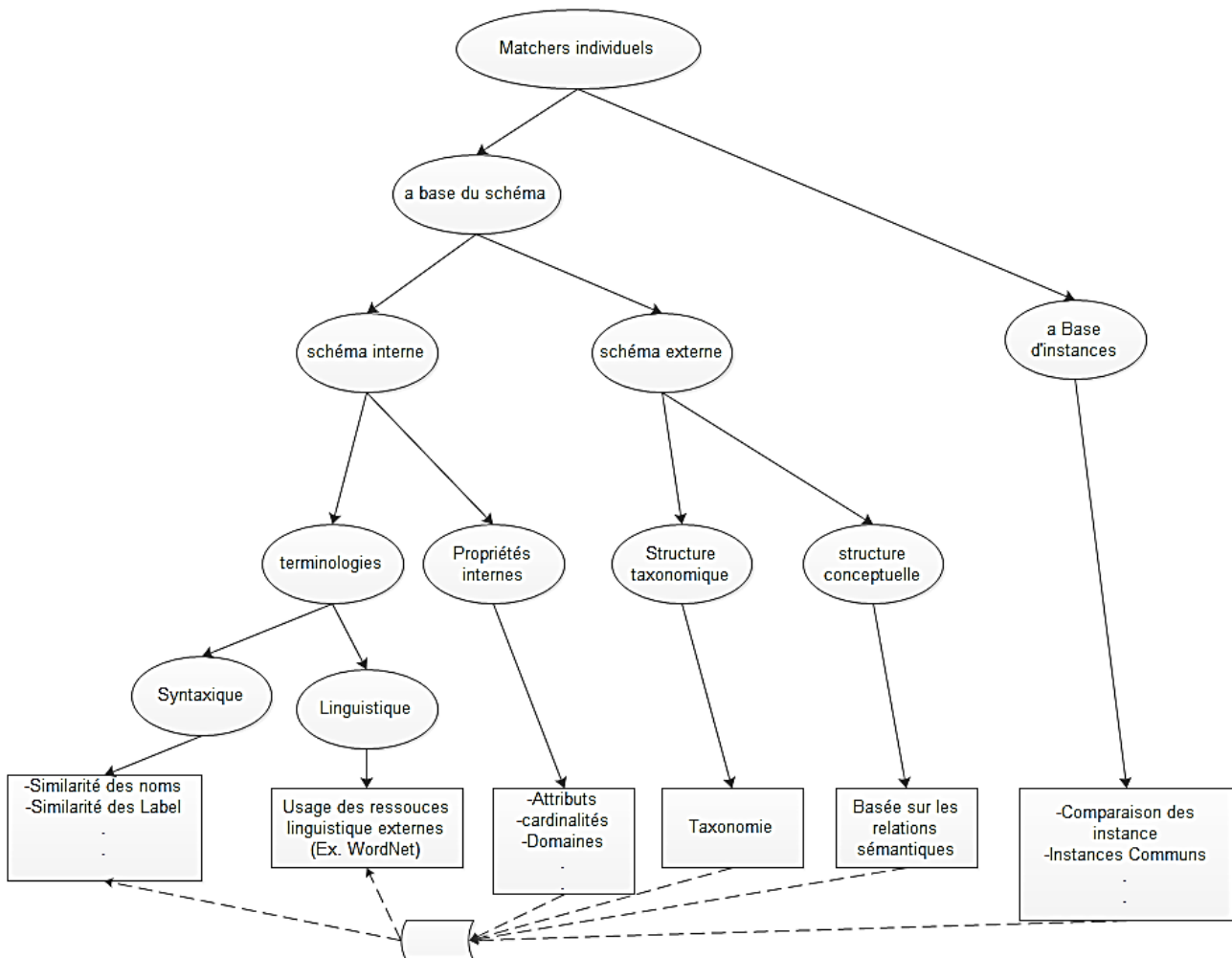


Figure 4.4 : Techniques de base pour l'alignement des ontologies (inspirée de (Euzenat et al., 2004)).

L'appariement entre deux entités ontologiques est basé sur le schéma d'ontologie ou basé sur les instances. Plusieurs niveaux sont présentés ci-dessous:

A) Les méthodes qui sont basées sur les schémas des ontologies

Elles exploitent le schéma (ou une partie) de l'ontologie pour déterminer la similarité entre entités. Le schéma d'une entité peut être divisé en interne et externe ;

-**Les techniques structurelles internes** : elles calculent la similarité entre deux concepts en exploitant les informations relatives à leur structure interne (intervalle de valeur, descriptions des attributs et des cardinalités, etc.) .

- **Les techniques structurelles externes** : elles exploitent la structure des ontologies (comparaison des entités au sein de leurs topologies, par exemple, en comparant leurs pères, fils, etc.). et se basent sur les chemins (par exemple : la longueur d'un chemin) pour déterminer la similarité sémantique entre deux entités.

- **Les techniques terminologiques** : utilisées pour calculer la valeur de similarité entre les descriptions textuelles des entités, telles que les noms, les labels, les commentaires, etc. C'est souvent la première information considérée lors du processus de matching. Les autres techniques peuvent faire recours aux techniques terminologiques. Une étape de normalisation est généralement nécessaire pour procéder à une comparaison terminologique. Si le terme est considéré comme une chaîne de caractères

alors une comparaison terminologie basée sur la syntaxe est choisie. Sinon si l'on considère le terme comme un mot alors une technique terminologique basée sur la linguistique est choisie.

- **Les techniques syntaxiques** effectuent la correspondance à travers les mesures de similarité des chaînes de caractères (par exemple, la similarité de Levenshtein (Levenshtein, 1966).) ;
- **Les techniques linguistiques** utilisant des ressources externes (dictionnaires, taxonomies, etc.) : la similarité entre deux entités représentées par des termes est calculée à partir des liens sémantiques déjà existant dans les ressources externes comme WordNet (Miller, 1995).

B) Les techniques basées sur les instances des ontologies

Lorsque les individus (ou instances) sont disponibles dans les ontologies, il existe une très bonne opportunité pour les systèmes d'appariement. Il est possible de déduire la similarité entre deux entités en analysant leurs extensions (leurs ensembles d'instances). Diverses techniques peuvent être utilisées. Par exemple, la comparaison des extensions communes (Si deux concepts partagent exactement le même ensemble d'individus, sont considérés comme similaires) ou les techniques d'identification des instances (Si un ensemble commun d'instances n'existe pas, il est possible d'identifier la correspondance d'une instance d'un ensemble avec une autre instance d'un autre ensemble).

- Souvent, une seule mesure est incapable de générer des mappings exhaustifs. Les différentes techniques de base sont généralement exploitées ensemble dans un seul algorithme avec une certaine stratégie de combinaison.

4.5.3 Quelques systèmes

Parmi les différents systèmes d'alignement des ontologies apparus ces dernières années une liste non-exhaustive des exemples comprend : PROMPT & PROMPT SUITE (Noy & Musen, 2003 ; Noy & Musen, 2000), QOM (Ehrig & Staab, 2004), S-match (Giunchiglia & Shvaiko, 2003 ; Giunchiglia et al., 2007), OLA (Valtchev et al., 2003) et des travaux plus récents comme SOBOM (Xu et al., 2010), AgreementMaker (Cruz et al., 2009), Eff2Match (Chua & Kim, 2010), GeRMeSMB (Quix et al., 2010) et ASMOV (Jean-Mary et al., 2009).

(Euzenat & Shvaiko, 2007) et (Euzenat et al., 2004) ; présentent un état de l'art complet des systèmes d'alignement des ontologies.

- **PROMPT & PROMPT SUITE** (Noy & Musen, 2003; Noy & Musen, 2000), est l'algorithme semi-automatique pour la fusion, la différence et l'alignement des ontologies. Il gère les ontologies exprimées en OWL et RDFS et produit les alignements / mappages entre deux entrées d'ontologie. Prompt est déjà développé en tant que plug-in Protégé. Il commence par les matchers linguistiques pour la comparaison initiale, mais guide l'utilisateur dans l'exécution d'autres tâches pour lesquelles son intervention est nécessaire (dans le choix des meilleures correspondances).

Chapitre 4 :Appariement sémantique élémentaire dans le Matchmaking des SWS

- **QOM** (Quick Ontology Mapping) (Ehrig & Staab, 2004): est une approche efficace pour identifier les mappings entre deux ontologies. QOM est caractérisée par un temps d'exécution réduit. Afin de réduire la complexité d'exécution QOM utilise une approche de programmation dynamique. Au lieu de comparer toutes les entités de la première ontologie avec toutes les entités de la seconde ontologie, QOM utilise des heuristiques pour abaisser le nombre de candidats au mapping. Le calcul de similarité est effectué avec une large gamme de fonctions de similarité. Une agrégation est opérée sur les mesures de similarité individuelles. Au lieu d'appliquer une fonction d'agrégation linéaire, QOM applique une fonction sigmoïde, qui met l'accent sur les ressemblances et souligne les faibles similarités individuelles. Grâce à plusieurs cycles d'itération, la qualité des résultats augmente considérablement.
- **S-Match**(Giunchiglia & Shvaiko, 2003; Giunchiglia et al., 2007) est une approche pour le matching dit sémantique. Les auteurs mettent en œuvre un opérateur de correspondance qui admet en entrée deux structures sous forme de graphes (par exemple, ontologies) et produit un mapping entre les éléments qui sont en correspondance sémantique. Le matching sémantique calcul une relation, de nature ensembliste, entre les nœuds en tenant compte de la sémantique de chaque nœud. La sémantique d'un nœud est déterminée par son label et par les nœuds qui sont plus haut dans la hiérarchie. Les relations possibles retournées par l'algorithme du matching sémantique sont l'égalité, l'intersection, la disparité, la généralité ou la spécificité. Dans ce cas, le problème du matching est vu comme un problème de satisfaction d'un ensemble de formules du calcul propositionnel. Les graphes et les correspondances à tester sont traduits en formules de la logique propositionnelle en considérant non seulement leurs noms, mais également la position des concepts dans le graphe.
- **OLA** (Valtchev et al., 2003) est un outil d'appariement des ontologies. il est le seul système conçu spécifiquement pour OWL Lite et qui utilise une mesure globale de similarité pour l'alignement d'ontologies. La similarité entre les entités locales dans les ontologies produit un ensemble d'équations résolues de manière itérative pour fournir un ensemble de correspondances. Il est aussi considéré comme un environnement pour manipuler des alignements.
- **SOBOM** (Xu et al., 2010), SOBOM s'occupe d'une ontologie à partir de deux points de vue différents: une ontologie O' avec la structure hiérarchique is-a et une ontologie O'' avec d'autres relations. Premièrement, du point de vue O', SOBOM commence par un ensemble d'ancres fournies par un langage Matcher et ensuite, il extrait des sous-ontologies basées sur les ancres et les rangs. Ces sous-ontologies en fonction de leur profondeur. Deuxièmement, SOBOM utilise un algorithme de simulation inductive sémantique pour calculer la similarité des concepts entre

différentes sous-ontologies dérivées des deux ontologies selon la profondeur des sous-ontologies pour obtenir des alignements conceptuels.

- **ASMOV** (Automated Semantic Mapping of Ontologies with Validation) (**Jean-Mary et al., 2009**) Le système prend comme entrée deux ontologies OWL et un alignement initial optionnel et retourne comme résultat un alignement $n: m$ entre les entités (classes et propriétés) des deux ontologies. Un calcul itératif des mesures de similarité lexicales et structurelles avec l'usage des sources de connaissance comme WordNet et UMLS. L'originalité de cette approche est la phase vérification sémantique de l'alignement généré après chaque itération. Le processus injection l'alignement obtenu dans l'itération suivante comme paramètre d'entrée, jusqu'à la convergence, c'est-à-dire « jusqu'à ce qu'aucune nouvelle correspondance ne soit trouvée ».
- **AgreementMaker** (Cruz et al., 2009): développé à l'université d'Illinois (Chicago), AgreementMaker comprend une large gamme de techniques d'appariement qui sont basées sur le concept ou sur la structure de l'ontologie. Les Matchers basés sur le concept utilisent la comparaison des chaînes, et les matchers structurels comprennent la similarité de pères et des fils. Cet algorithme facilite également l'intervention de l'utilisateur pour la résolution de conflits sémantiques. Les mesures de performances (rappel, précision et temps d'exécution) sont supportées par le système, ainsi que la combinaison pondérée des résultats fournis par ces méthodes. Il intègre une interface utilisateur avec un panneau de commande qui pilote toutes les méthodes d'appariement et des stratégies d'évaluation.
- **Eff2Match** (**Effective and Efficient ontology matching**) (Chua & Kim, 2010), Eff2Match prend comme entrée les URIs, d'une paire d'ontologies a aligné et produit les correspondances entre les entités (concepts ou propriétés) dans l'ontologie source et celles de l'ontologie cible. Le processus d'alignement se compose de quatre étapes: 1) Génération d'ancres, 2) Génération des candidats, 3) Expansion d'ancrage et 4) Augmentation de score itératif. Dans la phase d'ancrage, les entités correspondantes sont identifiées en utilisant une technique d'appariement exacte de chaînes. Dans l'étape génération des candidats, ils énumèrent les candidats pour les entités de l'ontologie source qui n'ont pas été appariées à l'étape précédente en utilisant une approche de modèle d'espace vectoriel (SVM). Dans l'étape d'expansion d'ancrage, des paires d'entités plus équivalentes sont Identifiées en comparant les entités sources avec leurs entités candidates en utilisant des méthodes terminologiques. Dans l'étape finale du processus d'appariement, un processus d'impulsion itératif est utilisé pour identifier plus de paires de concepts équivalents à l'aide de l'ensemble d'ancres
- **GeRoMeSuite** (**Quix et al., 2010**) est un système de gestion de modèle générique qui fournit plusieurs fonctions pour gérer des modèles de données complexes, telles que l'intégration de schéma, la définition et l'exécution des mappings de schéma, la transformation du modèle et

l'appariement. Le système utilise un méta-modèle générique (appelé GeRoMe) pour représenter des modèles, et à cause de cela, il est capable de traiter des modèles dans différents langages de modélisation tels que le schéma XML, OWL, ER, et des schémas relationnels. Un composant pour l'alignement des schémas et l'alignement d'ontologie fait également partie du système. En tant qu'outil de gestion de modèle générique, GeRoMeSuite n'est pas limité à un domaine particulier ou à un langage de modélisation. Par conséquent, l'outil fournit plusieurs stratégies d'appariement bien connues, telles que les matchers de chaînes, Similarity Flooding, les matchers de fils, les matchers des pères et matchers utilisant WordNet, etc. Afin de permettre la combinaison flexible de ces technologies d'appariement de base, il est doté d'une interface utilisateur graphique configurable. En raison de son approche générique, GeRoMeSuite est bien adapté pour les tâches correspondantes à travers des langages de modélisation hétérogènes, comme le fait d'associer XML Schema à OWL.

4.6 Alignement partiel des ontologies

Un alignement entre les deux ontologies (en général entre deux modèles sémantiques) apparie rarement tous les concepts du modèle avec tous les concepts de l'autre. Aussi dans des environnements avec des applications dynamiques tel que les agents ou les services web, il est préférable d'apparier en temps réel que des parties (utiles) des modèles sémantiques. Donc, généralement un alignement perdre quelques informations et peut être partiel ou incomplet (Kalfoglou & Schorlemmer, 2003). La définition formelle de l'alignement partiel des ontologies proposée par (Kalfoglou & Schorlemmer, 2003), stipule qu'un alignement total entre l'ontologie $O_1 = (S_1, A_1)$ et l'ontologie $O_2 = (S_2, A_2)$ est un isomorphisme $f: S_1 \rightarrow S_2$ des signatures ontologique, tel que, $A_2 \models f(A_1)$, ce qui signifie que toutes les interprétations qui satisfont O_2 satisfont O_1 . Bien sûr, dans la réalité, il est difficile d'atteindre ces mappings totaux et, par conséquent, il y a un alignement partiel d'ontologie entre $O_1 = (S_1, A_1)$ et $O_2 = (S_2, A_2)$, s'il existe une sous ontologie $O'_1 = (S'_1, A'_1)$ (S'_1 sous-ensemble de S_1 et A'_1 est un sous-ensemble de A_1), tel qu'il y ait un alignement total entre O'_1 et O_2 (Kalfoglou & Schorlemmer, 2003). L'existence d'une autre sous ontologie $O'_2 = (S'_2, A'_2)$ (S'_2 sous-ensemble de S_2 et A'_2 est un sous-ensemble de A_2) qui vérifie un alignement total entre O'_1 et O'_2 .

Dans des environnements dynamiques, avec des contraintes temps réel, il toujours souhaitable et pragmatique de généré qu'un alignement partiel.

4.7 Synthèse sur les techniques d'appariement sémantique élémentaire

Dans les études comparatives, des travaux de matching des ontologies se limitent aux systèmes de matching. Nous avons traité un contexte plus large, notamment l'appariement individuel dans le matchmaking des services web sémantiques. Pour généraliser, nous avons pris en compte un

Chapitre 4 :Appariement sémantique élémentaire dans le Matchmaking des SWS

appariement sémantique, abstraction du contexte mono ou multi ontologie. La *Table 4.1*, présente une comparaison entre les techniques d'appariement sémantique élémentaire, qui représente le noyau de toute approche de matchmaking des services web sémantiques. La comparaison est basée sur les critères suivants :

- ✓ Intra ontologie : indique si la technique traite l'appariement au sein de la même ontologie
- ✓ Inter ontologie : indique si la technique traite l'appariement entre des ontologies différentes. Il est évident que le critère inter ontologie subsume le critère intra ontologie. Dans la table 4.1, nous n'avons pas mentionné qu'une méthode inter ontologie est également intra ontologie, pour garder l'objectif principal de la méthode.
- ✓ Techniques de base : présente les techniques de base utilisées par la méthode
- ✓ Niveau : montre es que la technique d'appariement cherche l'appariement simplement de concept (niveau concept) ou toutes les entités de l'ontologie (niveau TEO)
- ✓ Exploite : le critère indique les parties des caractéristiques exploitées d'un élément pour opérer un appariement.
- ✓ Utilise ressources : présente les ressources exploitées par la techniques, par exemple les ontologies elles-mêmes sont des ressources ou encore une ressource lexicale.
- ✓ Entrée : les éléments acceptés comme entrée pour la technique.
- ✓ Sortie : Indique les résultats retournés par la technique.

Dans le tableau les abréviations suivantes indiquent :

BC : basée sur les chaînes de caractères,

BL : basée sur le langage

T.E.O : toutes les entités de l'ontologie

Chapitre 4 :Appariement sémantique élémentaire dans le Matchmaking des SWS

Critère Technique	Intra- ontologie	Inter- ontologies	Techniques de base	Niveau	Exploite	Utilise ressources	entrée	sortie
Mesures basées sur les chaînes de caractères	X	X	BC	Concept	noms	Les noms et labels des concepts	02 concepts	1 mapping + DM
Wu & Palmer (mesure basée sur la subsumption)	X		Longueur du chemin	Concept	Hiérarchie	Ontologie	02 concepts	1 mapping + DM
Rodriguez & Egenhofer(2003)		X	Appariement des mots(ou leurs synonymes), Appariement des caractéristiques, Appariement du voisinage	Concept	Noms, structure	Ontologies	02 concepts	1 mapping + DM
(Al-Mubaid & Nguyen, 2009)		X	Longueur de chemin entre deux concepts;	Concept	structure	Ontologies	02 concepts	1 mapping + DM
PROMPT & PROMPT SUITE		X	BC,Structure, proposition utilisateur	T.E.O	Noms, structure	Ontologies	02 ontologies	Alignement
QOM		X	BC,SE,	T.E.O	Nom, Structure	Ontologies	02 ontologies	Alignement
S-Match		X	BC, BL, SAT propositionnelle	T.E.O	Noms, la structure	WordNet	02 ontologies	Alignement
OLA		X	BC, BL, calcul itératif du point fixe, appariement du voisinage	T.E.O	Noms, hiérarchie Data type,	WordNet	02 ontologies	Alignement
SOBOM		X	BC, réseaux bayésiens naïfs, Similarité structurelle basée sur is-a et part-of	T.E.O	Noms, la structure	WordNet	02 ontologies	Alignement
ASMOV		X	lexicale, structurelle, matchers extensionnels, vérification sémantique	T.E.O	Noms et structure, instances	WordNet, UMLS	02 ontologies	Alignement
AgreementMaker		X	BC, BL, Similarité du voisinage	T.E.O	Noms, structure	WordNet	02 ontologies	Alignement
Eff2Match		X	BC, SVM, simulation itérative	T.E.O	Nom, structure		02 ontologies	Alignement
GeRoMeSuite		X	CB,Similarity flooding, fils	T.E.O			02 ontologies	Alignement
Notre mesure		X	BC,BL, Structurelle	Concept	Noms, la structure	Wordnet+ontologies	02 concepts	1 mapping + DM
Align_partiel		X	BC,BL, Structurelle, meilleurs affectation	Ensemble de concepts	Noms, structure	Wordnet +ontologies	02 Ensembles de concepts	Alignement partiel

Table 4.1 : Synthèse sur les techniques d'appariement sémantique élémentaire

Chapitre 4 :Appariement sémantique élémentaire dans le Matchmaking des SWS

Le plus simple appariement sémantique (première ligne : mesure basée sur les chaînes de caractères) s'appuie seulement sur l'information syntaxique, avec l'usage d'une simple mesure terminologique.

Un autre niveau supérieur est l'usage de l'information sémantique dans une seule ontologie. Ce point concerne tous les matchmakers logiques. Une comparaison de subsomption peut utiliser la mesure de longueur du chemin qui sépare deux concepts par exemple. Les travaux de matchmaking hybride combinent les deux niveaux d'appariement déjà décrits.

Pour effectuer un appariement sémantique élémentaire dans un contexte services web sémantiques multi ontologie, nous avons besoin d'un algorithme d'appariement qui prend en compte que les concepts avec lesquels la requête et le service sont annotés. Le cas idéal est un algorithme de similarité inter ontologie qui traite l'appariement entre les ontologies d'une manière partielle. Nous avons présenté deux mesures de similarité inter ontologies, à savoir, la mesure de Rodriguez & Egenhofer(2003) et Al-Mubaid & Nguyen(2009). Dans un environnement multi ontologie, il est toujours question d'interopérabilité des ontologies d'où la description de quelques travaux d'appariement des ontologies. Si l'on opte pour le choix d'un outil d'alignement dans un contexte services web sémantiques multi ontologie, il faut apporter des adaptations à des contraintes, comme la dynamique et l'aspect partiel de l'appariement. D'où l'idée et la nécessité de proposer une mesure de similarité inter ontologie et un algorithme d'alignement partiel.

4.8 Conclusion

Après cet aperçu de l'état de l'art, nous concluons que:

Pour atteindre un matchmaking sémantique, il faut appairer les annotations sémantiques de la requête et du service offert. La plupart des approches traitent le matchmaking des SWS dans le cadre d'une ontologie unique. Le matchmaking mono ontologie devient un goulot d'étranglement pour les services web, qui se traduit par un manque de scalabilité. D'une part, il est inévitable que plusieurs ontologies coexistent dans un environnement ouvert comme le web et l'interopérabilité sémantique des services Web dans un contexte multi-ontologie est obligatoire. D'autre part, l'alignement des ontologies représente une solution solide pour l'interopérabilité sémantique. Le chapitre suivant présente notre proposition qui utilise l'alignement des ontologies comme une solution pour l'appariement des services Web sémantiques.

Partie II : Contribution et Evaluation

- *Chapitre 5: Matchmaking des SWS dans un environnement multi-ontologie*
- *Chapitre 6 : Evaluation & Expérimentation*

Chapitre 5: Matchmaking des SWS dans un contexte multi-ontologie

Chapitre 5: Matchmaking des SWS dans un contexte multi-ontologie

5.1 Introduction

Comme il a été expliqué, le Matchmaking est le composant de base du processus de découverte des services Web. Le meilleur contexte possible serait l'existence d'une seule et unique ontologie avec laquelle, tous les services Web sont annotés. Une telle solution est cependant très peu pratique compte tenu de la variété d'applications et l'intention derrière leur développement. Même si on se limite à une ontologie globale par un seul domaine, les différents experts du domaine, des utilisateurs, des groupes de recherche et les organisations peuvent conceptualiser les mêmes entités du monde réel différemment, conduisant à un environnement multi-ontologie de domaine. Cependant, les différences et les hétérogénéités entre les ontologies utilisées peuvent toujours exister. En outre, les accords qui sont capturés par rapport à un domaine peuvent différer vu la différence des perspectives et des niveaux de détails. Par conséquent, quand une personne ou une organisation choisit une ontologie pour annoter, elle peut choisir l'une de ses ontologies disponibles qui enrichit au mieux le service web. En outre, le service Web lui-même peut englober plusieurs domaines. Ainsi, plus d'une ontologie peut être nécessaire tout en annotant un service Web unique. Dans un tel environnement où les requêtes et les offres respectent différentes ontologies, il se pose la nécessité d'un algorithme d'appariement qui peut surmonter les hétérogénéités entre différentes ontologies. Notre travail se focalise sur l'appariement de services Web dans un contexte multi-ontologie, il s'agit essentiellement d'aborder certains problèmes d'hétérogénéité sémantique en présence de plusieurs ontologies. Notre démarche, pour traiter cette problématique, est regroupée en deux grandes parties : En premier lieu, on propose un cadre général d'appariement de services web (web services matchmaking), ainsi on montre l'importance cruciale de mesure de similarité entre ontologies ; tout en présentant une version d'une similarité entre concepts de différentes ontologies. Ensuite, dans une deuxième phase, on présente une instantiation de notre Framework sur les services web SAWSDL ; avec une amélioration de notre mesure de similarité qui a été intégrée dans un algorithme d'alignement partiel des ontologies de domaines. Cet algorithme d'alignement partiel est la pierre angulaire de notre démarche de matchmaking des services web SAWSDL. Un outil nommé SAWSDL-MOM est développé.

5.2 Problématique

Pour une invocation automatique de service, il est nécessaire de détecter des services qui fournissent les capacités requises d'une manière très précise. Comme il a été déjà décrit, les services basés

seulement sur une norme syntaxique comme WSDL ne fournissent pas suffisamment d'information pour décrire les capacités de service en détail. Cela est valable aussi bien pour une description fonctionnelle que pour une description non fonctionnelle d'un service. Les lacunes des langages de description syntaxique de services concernant la découverte de service sont l'une des principales raisons de l'intégration de l'information sémantique dans les descriptions de service Web. Il est évident, que la simple existence des descriptions sémantiques ne permet pas l'amélioration de la découverte de service. Les informations sémantiques doivent être intégrées au cours du processus d'appariement. Pour aboutir à cet objectif, les chercheurs ont proposé une multitude d'approches d'appariement (voir chapitre 4).

Les présuppositions sur la nature et le degré de couverture des descriptions des services Web par des informations sémantiques ont fortement conditionné les techniques sur lesquelles est basée une approche de Matchmaking. En général, ces approches sont basées, sur des techniques syntaxiques ou sémantiques, et utilisent un appariement élémentaire (une mesure de similarité) non logique (le plus souvent pour les approches syntaxiques), un appariement basé sur la logique ou une combinaison hybride entre les techniques logiques et les techniques non logiques. Selon les résultats du concours annuel S3 (International Semantic Service Selection Contest) sur la sélection de services sémantique, les performances des Matchmakers dépendent très souvent des techniques d'appariement élémentaires. La constatation évidente de l'état de l'art est que la quasi-totalité des Matchmakers présupposent l'existence d'une seule ontologie partagée entre la requête et le service offert. Dans le contexte multi-ontologie, on fait face à de nouvelles difficultés, en particulier le traitement des hétérogénéités des modèles sémantiques, et la prise en charge de l'aspect interopérabilité sémantique. Le problème pivot à résoudre est la définition d'un mécanisme de médiation entre la requête et le service. Dans un premier lieu, un cadre générique pour l'appariement des services web dans un contexte multi-ontologie est nécessaire pour déterminer les composants essentiels d'une telle solution. En second lieu, le traitement de l'interopérabilité des ontologies différentes. L'usage direct d'une solution développée dans la littérature pour mesurer la similarité demande plusieurs adaptations, et l'usage des outils de Matching des ontologies n'est pas vraiment adéquat pour un contexte temps réel et dynamique comme les services web.

5.3 Architecture de la découverte

Une architecture de découverte sémantique des services web (voir *Figure 5.1*), doit incorporer un mécanisme (Multi-Ontologie Matchmaker) capable d'apparier une requête de service (enrichie par un modèle sémantique, une ontologie le plus souvent) et un ensemble de services offerts (enrichis aussi par des modèles sémantiques), pour déterminer des degrés de correspondance en exploitant les

informations sémantiques. Un classement des résultats (Ranking), en fonction du degré de correspondance génère une liste des services pertinents pour la requête.

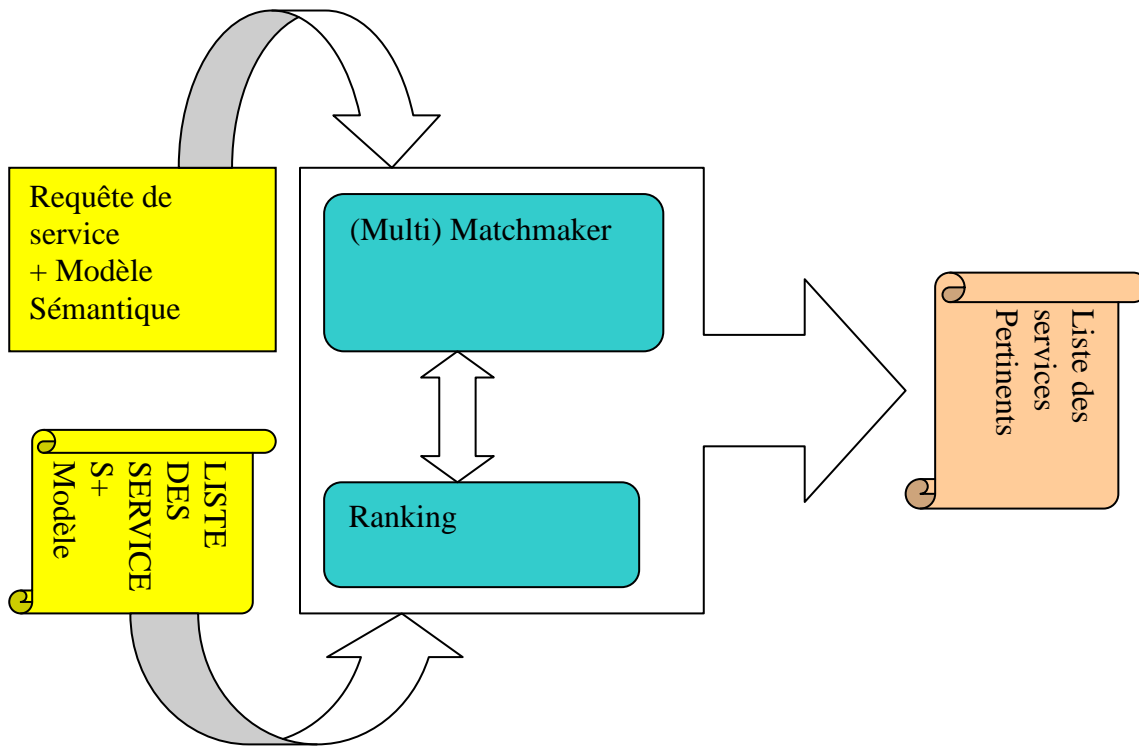


Figure 5.1: Découverte de service web sémantiques

La section suivante présente notre première contribution, un cadre général qui détail la partie matchmaking d'une requête et un service.

5.4 Framework général

Dans cette première proposition (Fellah et al., 2016a) nous avons tenté de décrire un cadre général d'appariement des services web sémantique susceptible de gérer l'interopérabilité sémantique dans un contexte multi-ontologie. Nous présentons un Framework générique (Figure 5.2), pour l'appariement des services web dans un environnement multi-ontologie. Pour rendre notre Framework le plus général que possible, nous n'imposons aucune restriction sur la description de la requête ou du service, que ce soit sur le langage de description et/ou les parties décrites du service (c.-à-d. La partie fonctionnelle et la partie non fonctionnelle).

Par exemple si l'on opte pour SAWSDL (Farrell & Lausen, 2007) pour représenter les services web. La spécification SAWSDL est limitée à l'annotation des paramètres fonctionnels comme les entrée/sorties avec des concepts d'une ontologie via l'attribut modelReference. Pour les paramètres non fonctionnels, il est possible de faire leurs descriptions avec un langage enrichi avec la sémantique

Chapitre 5 : Matchmaking des SWS dans un contexte multi-ontologie

des ontologies, comme dans le travail de (Kritikosr, 2008) qui se focalise sur le paramètre Qos, exprimé avec un langage appelé OWL-Q ou encore le travail de (Keskes, 2014), le paramètre QoS(Quality of service) peut être annoté par les concepts d'une ontologie de contexte.

Cependant, pour appairer une requête à un service, principalement, on a besoin d'une mesure de similarité qui s'emploie conjointement avec un algorithme de stratégie d'agrégation locale (qui peut utiliser un algorithme d'alignement partiel et un algorithme d'affectation comme la méthode hongroise (Kuhn, 2010) et une stratégie de matchmaking.

Cette vision de matchmaking de service est présentée dans la *Figure 5.2*. L'approche suit plusieurs pas. Premièrement, les éléments d'annotations (par exemple Inputs_A représente deux vecteurs des annotations sémantiques des entrées et de même Outputs_A dans la *Figure 5.2*) qui sont des concepts ontologiques sont extraits et classifiés par catégorie (c.-à-d. partie fonctionnelle (les catégories comme les opérations, les inputs et les outputs) et les autres parties). Par la suite, les concepts sont appariés par catégories avec l'utilisation de l'algorithme de Matching des concepts proposé dans la *Figure 5.4*), les résultats des similarités entre les annotations sémantiques, sont stockés dans des matrices (par exemple inputs_S pour les entrées. La matrice des similarités des opérations est implicite pour une souplesse de la propagation des similarités. Par la suite, on procède à une agrégation locale pour chaque catégorie de similarités calculées et présentées sous forme de matrice (par exemple Input_S aggregation est l'agrégation des similarités des inputs (requête et service). On procède à un calcul d'une agrégation des similarités du niveau fonctionnel et une autre du niveau non fonctionnel. Finalement, un degré d'appariement global et final (matching score dans la *Figure 5.2*)) est calculé. Une stratégie d'agrégation précise est sélectionnée pour le calcul de chaque degré d'appariement (local ou global). La stratégie peut varier d'une simple pondération arithmétique à un algorithme complexe.

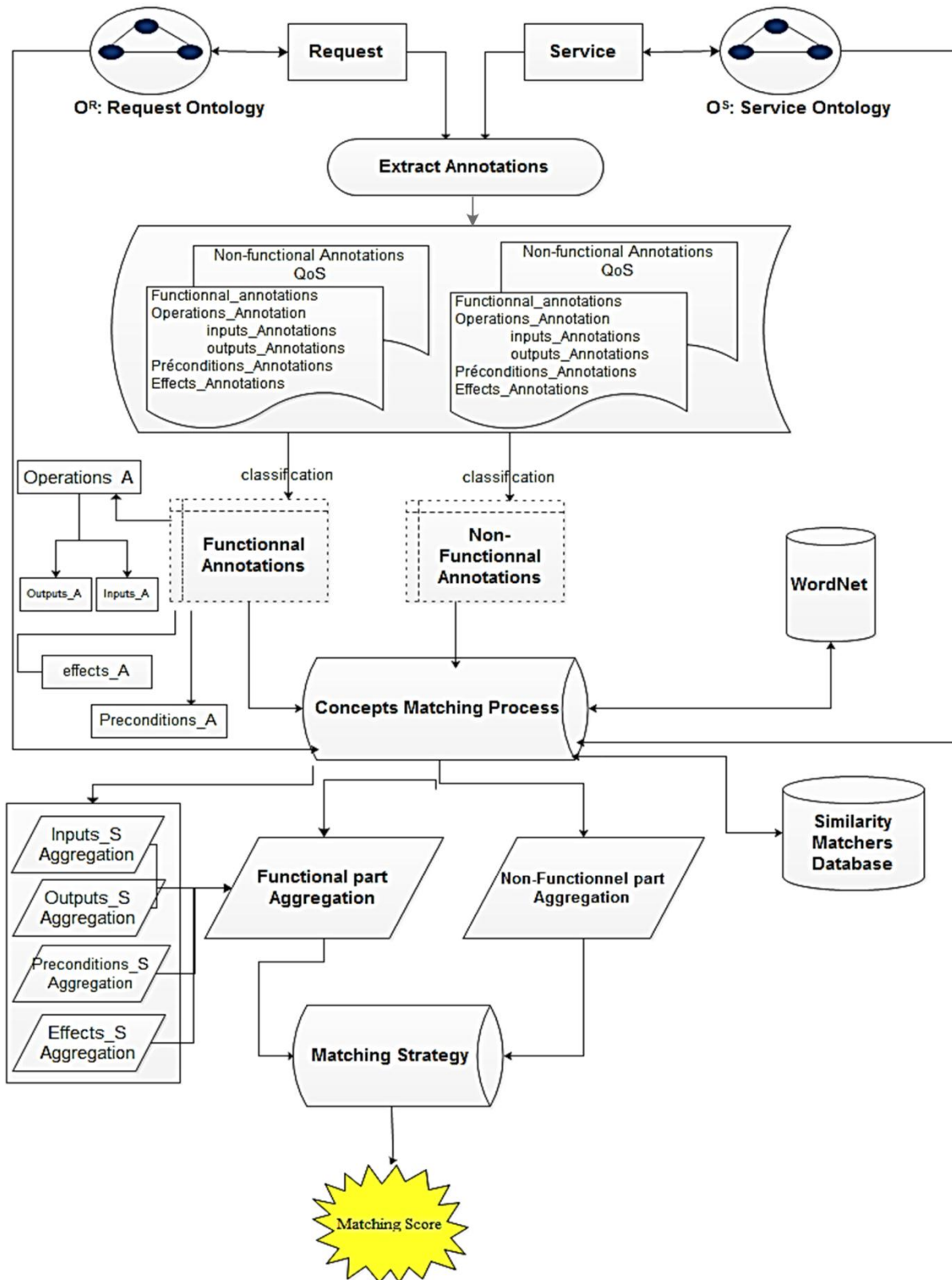


Figure 5.2: Framework de Matchmaing des SWS multi ontologie.

Notre architecture d'appariement des services (de la Figure 5.2), est composée essentiellement de :

- Un module d'extraction des annotations
- On suppose l'existence d'un parseur universel ou un transformateur du langage de description vers un langage pivot pour :
 - L'abstraction du langage de description de la requête
 - L'abstraction du langage de description du service
- Un outil d'extraction et de catégorisation des annotations sémantiques
- Un élément fondamental pour l'appariement élémentaire, entre les descriptions sémantiques. Nos investigations nous ont conduit à déduire que la mesure de similarité est la pierre angulaire de toute architecture de découverte. L'usage direct d'une mesure de similarité développée dans la littérature demande des adaptations, et l'usage des outils de matching des ontologies n'est pas vraiment adéquat pour un contexte temps réel et dynamique comme les services web. La similarité inter ontologie est détaillée dans la section suivante.
- Un ensemble d'outils et de ressources (Dictionnaires, thésaurus, Bibliothèque de Matchers)
- Un Module d'agrégation des résultats intermédiaires de l'appariement élémentaire par partie des composants du service web sémantique (i.e. la partie fonctionnelle et la partie non fonctionnelle)

5.5 Mesure de Similarité inter Ontologies

Une des parties les plus importantes d'un processus d'appariement est le calcul de similarité (ou l'appariement élémentaire); car elle "décide" comment deux concepts sont similaires. Pour notre approche, nous appliquons plusieurs types de similarité simples. Notre mesure de similarité peut normalement être utilisée dans n'importe quelle instanciation de notre Framework. Le noyau de notre solution qui est détaillé dans la suite, composé principalement :

- Un algorithme pour mesurer la similarité entre concepts de différentes ontologies, et
- Un ensemble de mesures de similarité ou des apparieurs élémentaires.

5.5.1 Algorithme d'appariement des concepts

Pour comparer une requête R (sa description sémantique, contient un concept $CR \in O^R$ qui annote un élément ER qui est une entrée par exemple) avec un service S (qui possède une entrée, annotée par un concept $CS \in O^S$); le calcul de la similarité entre le concept CR et le concept CS est fondamental pour la comparaison entre la requête et le service.

Pour aligner le concept CR avec le concept CS appartenant à deux ontologies différentes, nous commençons par comparer leurs URIs. S'ils sont identiques, le mapping est parfait, sinon nous calculons une mesure de similarité globale, qui représente une agrégation de plusieurs similarités,

Chapitre 5 : Matchmaking des SWS dans un contexte multi-ontologie

notamment, la similarité des noms, et la similarité des contextes internes et la similarité des contextes externes. La *Figure 5.4* présente le processus d'appariement des concepts. L'algorithme (*Compute_Similarity*) est l'implémentation du mode de calcul de la similarité.

```
(1) Algo Compute_Similarity
(2) Inputs: CR :Request Concept ; CS:Service Concept ;
(3) Outputs: Result :Overall similarity;
(4) {
(5) If (CR, CS) in Alignment_Cache
(6) return(result from Alignment_Cache)
(7) Else
(8)   {
(9)     Sim_URI=Compute_sim_uri(CR,CS) ;
(10)    If(Sim_URI =1)
(11)      return(Result=sim_uri)
(12)    Else
(13)      Result=Compute_Averall_Sim(CR,CS);
(14)    }
(15) return(Result)
(16) }
```

Figure 5.3: Calcul de similarité entre deux concepts.

Le calcul de la similarité globale est détaillé par l'algorithme (*Compute_Averall_Sim* , *Figure 5.12*).

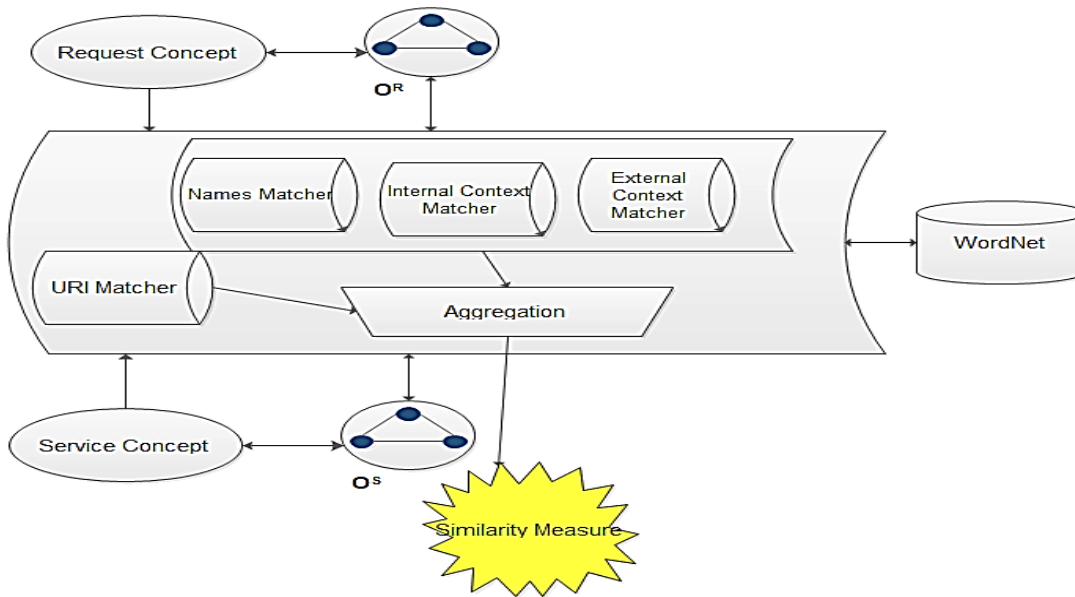


Figure 5.4: Le Processus de Matching des Concepts.

L'algorithme `Compute_Averall_Sim` invoque les fonctions suivantes :

- `Compute_Sim_Name` : est l'implémentation de la mesure de similarité des noms présentée dans la section (5.5.3).
- `Compute_Sim_Prop` : est l'implémentation de la mesure de similarité des propriétés ou du contexte interne présentée dans la section (5.5.4).
- `Compute_Sim_V` : est l'implémentation de la mesure de similarité du contexte externe présentée dans la section (5.5.5).

(1) **Algo** `Compute_Averall_Sim`

(2) **Inputs** :CR, CS: two concepts

(3) **Outputs**:result

(4) **Parameters**:WName,WProp,WFather

(5) {

(6) `Sim_Name=Compute_Sim_Name(CR.name,CS.name);`

(7) `Sim_Prop=Compute_Sim_Prop(CR.prop, CS.prop);`

(8) `Sim_Vois=Compute_Sim_V(CR, CS) ;`

(9) `Result=WName * Sim_Name+WProp * Sim_Prop+WV *SimV;`

(10) **Return**(Result)

(11) }

Figure 5.5 : Calcul de similarité globale.

➤ Exemple de comparaison entre deux concepts

L'exemple dans la Figure 5.6 est une adaptation de (Le DuyNgan & Goh, 2006) qui présente le matching des deux concepts *Network Node* and *Network Equipment*. Les deux concepts ont comme super concept « *Equipment* »; et, leurs sous concepts sont *Computer*, *Switch Equipment*, et *Computer*, *Central Hub*, respectivement.

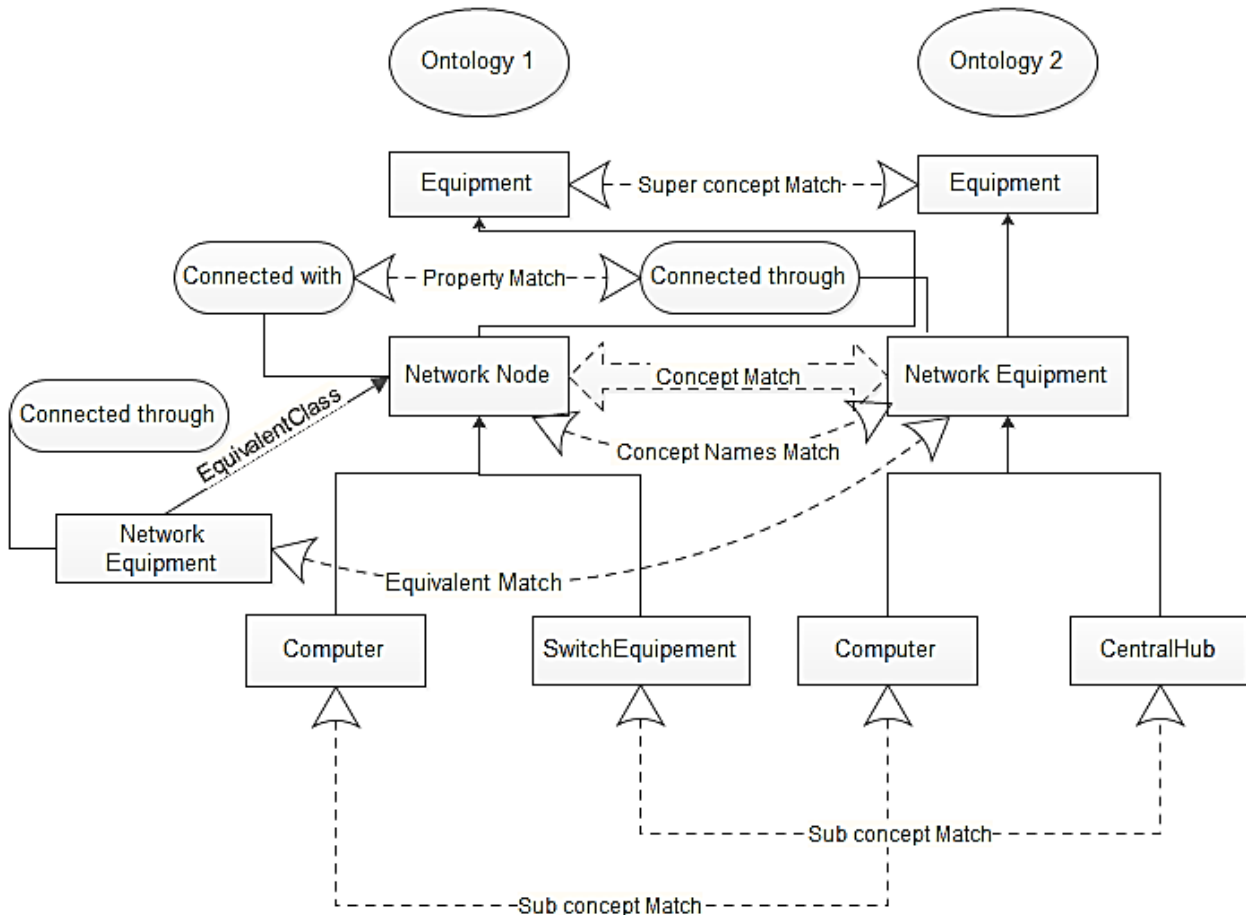


Figure 5.6: Comparaison entre deux concepts: *Network Node* and *Network Equipment*

Pour calculer la similarité entre *Network Node* et *Network Equipment*; plusieurs dimensions doivent être prise en considération:

- la similarité entre URI (*Network Node*) et URI (*Network Equipment*),
- la similarité syntaxique entre $Names(Network Node)$ et $Names(Network Equipment)$.
- la similarité entre les propriétés: *Connected With* of *Network Node* et la propriété: *Connected through* of *Network Equipment*;
- la similarité entre le contexte externe: *Equipment*, *Computer*, *CentralHub* de *Network Node* et le the contexte externe: *Equipment*, *Computer*, *Switch* de *Network Equipment*.

Le mode de calcul de chaque similitude est formalisé dans les sections suivantes.

5.5.2 La similarité des URI (URI Matcher)

Deux concepts sont équivalents (100% similaires) si leurs URIs sont les mêmes :

$$Sim_{URI}(CR, CS) = \begin{cases} 1 & \text{if } URI(CR) = URI(CS) \\ 0 & \text{Autrement} \end{cases} \quad (1)$$

Par exemple dans *Figure 5.6*, si on prend pour CR le concept *Equipment* de *Ontology1* et pour CS le concept *Equipment* de *Ontology2* alors $Sim_{URI}(CR, CS) = 0$

5.5.3 La similarité des noms (Names Matcher)

$$Sim_{Names}(CR, CS) = \frac{Sim_{Syn}(T_{CR}, T_{CS}) + Sim_{Ling}(T_{CR}, T_{CS})}{2} \quad (2)$$

Où:

$Sim_{Syn}(T_{CR}, T_{CS})$: La similarité Syntaxique entre les termes T_{CR} et T_{CS} respectivement associés aux concepts CR et CS .

$Sim_{Ling}(T_{CR}, T_{CS})$: Similarité Linguistique entre les termes T_{CR} et T_{CS} respectivement associés aux concepts CR et CS .

5.5.3.1 La similarité Syntaxique (Levenshtein Matcher)

La similarité de deux chaînes de caractères sur l'intervalle [0,1] on se basant sur edit distance de Levenshtein, ed (Levenshtein, 1966).

$$Sim_{Syn}(CR, CS) = \frac{Max(|T_{CR}|, |T_{CS}|) - ed(T_{CR}, T_{CS})}{Min(|T_{CR}|, |T_{CS}|) + ed(T_{CR}, T_{CS})} \quad (3)$$

Où:

$|T_{CR}|$: est le nombre de caractères dans le terme T du concept CR ; également pour $|T_{CS}|$.

$ed(T_{CR}, T_{CS})$: La distance d'édition qui représente le nombre de transformations (addition, suppression, modification) nécessaires pour obtenir T_{CS} à partir de T_{CR} .

Notre mesure est une amélioration du travail (Ehrig & Staab, 2004). Par exemple, si T_{CR} ='Professor' and T_{CS} ='Assistant Professor'; leur mesure est limitée à 0, par contre la nôtre retourne 0.47, (c'est une valeur significative).

$|T_{CR}|=9$; and $T_{CS} = 19$; ed(le nombre des transformations, addition dans ce cas) =10 alors, on a :

$$Sim_{Syn}(CR, CS) = \frac{19-10}{9+10} = 0.47$$

5.5.3.2 WordNet Matcher (Linguistic Matcher)

WordNet est un thésaurus de l'anglais, développé sous la direction de Miller à l'université de Princeton (Miller, 1995). Un ensemble de synonymes cognitifs appelés synsets, chacun représente un concept différent, sont formés en regroupant les noms, les verbes, les adjectifs et les adverbes. Les synsets sont créés en utilisant des relations sémantiques et lexicales conceptuelles. WordNet peut également être considérée comme ontologie des termes de l'anglais.

Notre proposition de mesure de similarité linguistique (Sim_{Ling}) est formulée comme suit :

$$Sim_{Ling}(CR, CS) = \frac{|Syn(T_{CR}) \cap Syn(T_{CS})|}{|Syn(T_{CR}) \cup Syn(T_{CS})|} \quad (4)$$

Où:

$Syn(T_{CR})$: Tous les synsets du terme T_{CR} ,

$Syn(T_{CS})$: Tous les synsets du terme T_{CS} , et

$|Syn(T_{CR}) \cap Syn(T_{CS})|$: Le nombre de termes communs.

Deux concepts sont complètement similaires, si leurs termes associés possèdent les mêmes synsets. La similarité retourne 0, s'ils n'ont aucun synset commun.

Dans des cas spéciaux où le nom est composé de plusieurs termes alors la stratégie suivante est adoptée :

Si $T_{CR} \in \{T1_{CR}, T2_{CR}, \dots, Tn_{CR}\}$ ou $T_{CS} \in \{T1_{CS}, T2_{CS}, \dots, Tn_{CS}\}$

$$\text{Alors } Sim_{Ling}(CR, CS) = \text{Max}(Sim_{Ling}(CR_{Ti}, CS_{Tj}))$$

5.5.4 La similarité de contexte interne:

Le contexte interne est composé des propriétés internes du concept. Chaque propriété est caractérisée par un nom, rang et un domaine. La matrice des similarités entre les propriétés des deux concepts CR et CS, notée M_p est construite comme suit:

$$M_p = \begin{bmatrix} Sim_p(CR_{P1}, CS_{P1}) & \cdots & Sim_p(CR_{P1}, CS_{Pm}) \\ \vdots & \ddots & \vdots \\ Sim_p(CR_{Pn}, CS_{P1}) & \cdots & Sim_p(CR_{Pn}, CS_{Pm}) \end{bmatrix} \quad (5)$$

Où:

$Sim_p(CR_{Pi}, CS_{Pj})$: est la similarité entre les propriétés P_i du concept CR et la propriété P_j du concept CS, calculée selon l'équation 7;

n: le nombre de propriétés du concept CR,

m: le nombre de propriétés du concept CS.

soient Q_1, Q_2, \dots, Q_k , $k = \min(n, m)$, où Q_i est calculée comme suit:

Algorithme affectation

1. $i=1$
 2. $Q_i = \max(M_p) = m_{ij}$
 3. élimination de la ligne i et de la colonne j
 4. $i=i+1$
 5. répéter 2 jusqu'à M_p est vide
-

L'algorithme cherche les affectations optimales binaires de similarités meilleures, dans la matrice de similarité (il cherche le vecteur des couples de concepts avec une mesure de similarité maximale). Finalement, la similarité des propriétés ou du contexte interne est donnée par :

$$Sim_{prop}(CR, CS) = \frac{\sum_{i=1}^k Q_i}{\max(n, m)} \quad (6)$$

5.5.4.1 Matcher de propriétés

Un concept peut avoir une ou plusieurs propriétés. Similaire au concept, une propriété possède aussi un nom et une description. En plus, elle contient un domaine, et une cardinalité. Pour calculer la similarité des propriétés, toutes les informations des deux propriétés peuvent être appariées. Pour calculer la similarité des noms et des descriptions, on adopte la similarité syntaxique introduite dans la section (5.5.4). La similarité finale des propriétés est la combinaison des deux composants Nom et rang.

$$Sim_p(CR_{P_i}, CS_{P_j}) = W_N * Sim_{Names}(CR_{P_i}, CS_{P_j}) + W_R * Sim_{Ranges}(CR_{P_i}, CS_{P_j}) \quad (7)$$

Où:

$Sim_{Names}(CR_{P_i}, CS_{P_j})$: La similarité des noms Eq. (2),

$Sim_{Ranges}(CR_{P_i}, CS_{P_j})$: La similarité des rangs Eq. (8),

5.5.4.2 Matcher du rang de propriété

- La gamme d'une propriété est soit un type de données primitif ou à un autre concept. Si les deux gammes sont des types de données primitives, alors la similitude entre deux plages est telle que représentée ci-dessous dans l'équation (8). Si une propriété de gamme à un type primitif et l'autre a un concept, les gammes sont incompatibles; donc la gamme similitude est 0.

Si deux propriétés de la gamme sont des concepts, la mise en correspondance est effectuée de manière récursive comme avec le calcul de similarité pour deux concepts. (Usanavasin et al. (2005))

$$Sim_{Ranges}(CR_{Pi}, CS_{Pj}) = \begin{cases} 1 & \text{Si } Range(CR_{Pi}) = Range(CS_{Pi}) \\ 0.9 & \text{Si } Range(CR_{Pi}) = \text{anytype et } Range(CS_{Pi}) = \text{String} \\ 1 & \text{Si } Range(CR_{Pi}) = \text{integer et } Range(CS_{Pi}) = \text{Float} \\ 1 & \text{Si } Range(CR_{Pi}) = \text{integer et } Range(CS_{Pi}) = \text{Double} \\ \frac{2}{3} & \text{Si } Range(CR_{Pi}) = \text{Float et } Range(CS_{Pi}) = \text{integer} \\ \frac{1}{3} & \text{Si } Range(CR_{Pi}) = \text{Float et } Range(CS_{Pi}) = \text{Double} \\ \frac{1}{3} & \text{Si } Range(CR_{Pi}) = \text{Double et } Range(CS_{Pi}) = \text{integer} \\ \frac{1}{2} & \text{Si } Range(CR_{Pi}) = \text{Double et } Range(CS_{Pi}) = \text{Float} \\ 0 & \text{Autrement} \end{cases} \quad (8)$$

5.5.5 La similarité de contexte externe

Les techniques structurelles consistent en exploitation de la structure des ontologies à comparer, souvent représenté sous forme de graphiques, et comparés. La similarité entre deux entités de deux ontologies peut être basée sur la position des entités dans leurs hiérarchies (Euzenat et al., 2004).

Ces techniques mettent en œuvre diverses heuristiques et sont basés sur l'hypothèse suivante:

Si deux entités dans deux ontologies sont semblables, leurs voisins sont aussi en quelque sorte. Cette remarque peut être utilisée de plusieurs façons.

Exemple de critères pour décider que les deux entités sont similaires:

C1: deux concepts sont similaires si leurs «super-concepts» («pères») sont similaires;

C2: deux concepts sont similaires si leurs «sous-concepts» ("fils") sont similaires;

C3: deux concepts sont similaires si leurs «voisins ou frères» sont similaires.

Dans ce cadre, nous proposons une nouvelle mesure de similarité structurelle multi critères en améliorant les travaux de Abolhassani (Abolhassani et al., 2006) et Fella (Fella et al., 2008).

Le contexte externe (ou le voisinage) d'un concept dans notre vision est composé de:

-Les Frères: Tous les concepts pour lesquels il existe une relation symétrique; formellement:

$$Frere(c) = \{c' \in O / \exists R \in O \text{ and } R(c', c) \text{ and } \exists R(c, c')\}$$

-Fils: Tous les concepts pour lesquels il existe une relation asymétrique; formellement:

$$Fils(c) = \{c' \in O / \exists R \in O \text{ and } R(c, c') \text{ and } \nexists R(c', c)\}$$

-Pères: Tous les concepts pour lesquels il existe une relation asymétrique; formellement:

$$Pere(c) = \{c' \in O / \exists R \in O \text{ and } R(c', c) \text{ and } \nexists R(c, c')\}$$

5.5.5.1 les bases théoriques et intuitive

A partir de des critères C1, C2 et C3, la similarité entre le concept CR de l'ontologie O^R et le concept CS de l'ontologie O^S dépend de la similarité entre leurs contextes relatifs. En d'autres termes, si le contexte du concept CR noté $V(CR)$ est similaire au contexte $V(CS)$ du concept CS alors CR et CS sont probablement similaires. Pour notre étude, nous considérons le voisinage directs du concept qui se réduit à l'ensemble {Père, Frère, Fils}.

Définition (le contexte externe du concept c)

Le contexte du concept $c \in$ ontologie O est défini par:

$$V(c) = \{\text{les concepts } c' / c' \in O \wedge (c' \in \text{Père}(c) \vee c' \in \text{Fils}(c) \vee c' \in \text{Frère}(c))\}$$

Définition (le Contexte Externe Relatif)

Le contexte externe relatif du couple (CR, CS) (avec : $CR \in O^R$ and $CS \in O^S$) est défini par les couples (ce^R, ce^S) tel que les conditions suivantes sont satisfaites:

$$\forall ce^R \in O^R \text{ and } \forall ce^S \in O^S$$

$$REC(CR, CS) = \{(ce^R, ce^S) \text{ tel que, } ce^R \in V(CR) \text{ et } ce^S \in V(CS) \text{ et}$$

$$Cat(V(CR)) = Cat(V(CS))\} , \text{Où}$$

$$Cat(V(c)) \in \{\text{Père}(c), \text{Fils}(c), \text{Frère}(c)\}$$

La similarité entre le concept $CR \in O^R$ et le concept $CS \in O^S$ dépend de la similarité des couples du Contexte Externe Relatif $(CR_i^{CRE}, CS_i^{CRE}) \in REC(CR, CS)$ avec $i=1..|REC(CR,CS)|$ et aussi de la force de lien (relatedness) entre CR et CR_i^{CRE} et la force de lien entre CS and CS_i^{CRE} . Cette force de lien indique le niveau de granularité, plus de détail entre les deux concepts s'exprime par une faible valeur. On quantifié cette mesure par prise en compte de la similarité entre les contextes des concepts et la connectivité entre le concept et son contexte. La connectivité entre CR et CR_i^{CRE} est la distance sémantique entre CR et CR_i^{CRE} . Il est difficile de choisir entre les mesure de similarité sémantiques et spécialement laquelle peut retourner le meilleur résultat dans notre cas. Nous comptons ici sur la mesure de Wu et Palmer (Wu & Palmer, 1994) appliquée à WordNet, et elle est choisie en raison de sa simplicité de mise en œuvre.

La similarité de Wu & Palmer (Sim_{wp}) est exprimée comme suit :

$$Sim_{wp}(C1, C2) = \frac{depth(LCA(C1, C2))}{depth(C1) + depth(C2)} \quad (9)$$

$$\text{et } Sim_{Str}(C1, c2) = Sim_{wp}(C1, C2)$$

Où, LCA (C1,C2) est Le plus bas ancêtres communs (Lowest Common Ancestors of C1 and C2) et depth (C1) est la longueur du chemin entre c1 et la racine en passant par LCA(C1, C2).

5.5.6.2 Apparieur des pères (Fathers Matcher)

Nous calculons la similarité des pères entre ontologies comme suit:

$$Sim_F(CR, CS) = \frac{\sum_{l=1}^{k_F} (Q_l * (1 - |Sim_{Str}(CR, CR_{Fl}) - Sim_{Str}(CS, CS_{Fl})|))}{k_{Max_F}} \quad (10)$$

Q_l : représente la similarité des noms entre les couples des pères $(CR_l^{CRE}, CS_l^{CRE}) \in CRE(CR, CS)$ qui contient uniquement les pères et $Sim_{Str}(C1, c2) = Sim_{Wp}(C1, C2)$

$CR_{Fl} \in Père(CR)$ et $CS_{Fl} \in Père(CS)$

Par exemple : le calcul de la similarité des pères entre les concepts CR de l'ontologie1 et CS de l'ontologie2 :

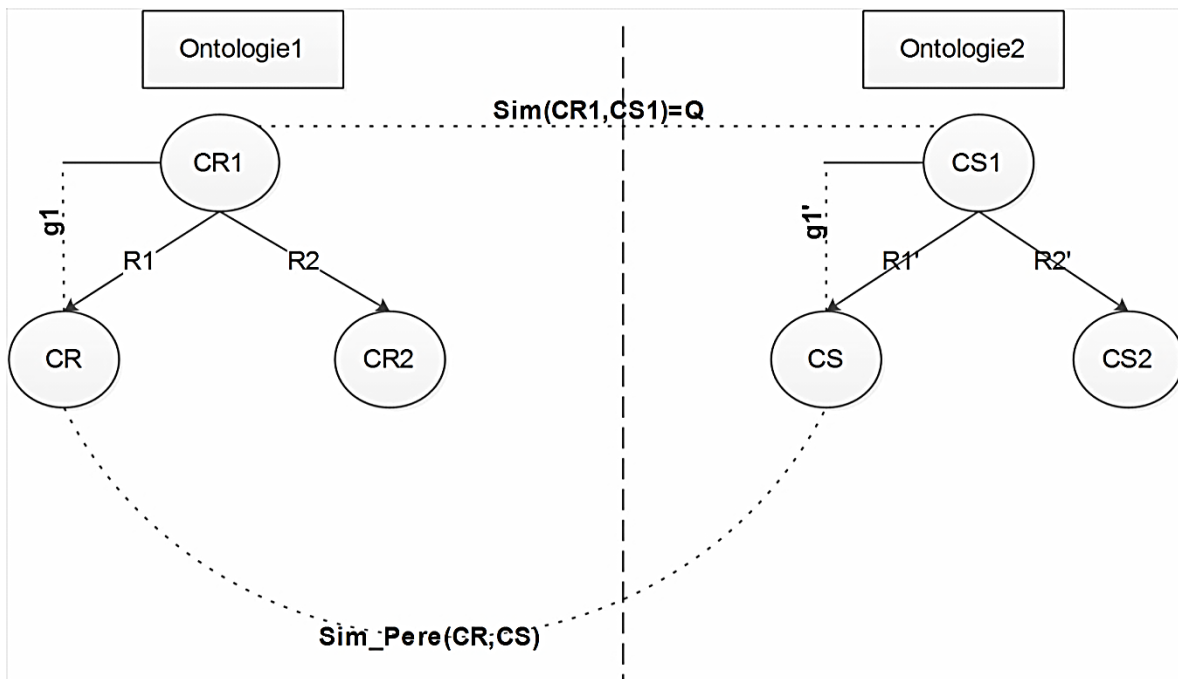


Figure 5.6: La similarité des contextes externes (pères) des deux concepts: CR & CS.

$g1$: La granularité de CR par rapport au père CR1,

$g1'$: La granularité de CS par rapport au père CS1,

$Sim(CR1, CS1)$: la similarité des pères CR1 et CS1, souvent la similarité des noms.,

Chapitre 5 : Matchmaking des SWS dans un contexte multi-ontologie

$Sim_père(CR,CS)$: la similarité des pères entre CR et CS ; intuitivement, elle dépend de la granularité de chaque concept (le lien entre concept et son père c.-à-d. $g1$ et $g1'$) et aussi de la similarité entre les pères $Sim(CR1,CS1)$ noté Q , selon l'équation (10) :

$$Sim_père(CR2,CS2) = Q * (1 - |g1 - g1'|) / 1$$

Dans le cas de plusieurs pères pour un concept alors on calcul une matrice de similarité entre les pères des deux concepts. $Q_{ij} = Sim(CR_i,CS_j)$: Représente la similarité des noms entre les couples des pères $(CR_i^{CRE}, CS_i^{CRE}) \in CRE(CR,CS)$ qui contient uniquement les pères. On obtient la matrice suivante:

$$M_P = \begin{bmatrix} Q_{11} = Sim(CR_1,CS_1) & \dots & Q_{1m} = Sim(CR_1,CS_m) \\ \vdots & \ddots & \vdots \\ Q_{n1} = Sim(CR_n,CS_1) & \dots & Q_{nm} = Sim(CR_n,CS_m) \end{bmatrix}$$

Et ensuite, à partir de la matrice M_P un vecteur des meilleurs appariements Q_1, Q_2, \dots, Q_{KF} entre couples de pères est généré par l'algorithme d'affectation suivant :

Algorithme affectation

1. $l=1$
 2. $Q_l = \max(M_P) = m_{ij}$
 3. élimination de la ligne i et de la colonne j
 4. $l=l+1$
 5. répéter 2 jusqu'à M_P est vide
-

Avec :

n : le nombre des pères du concept CR et

m : le nombre des pères du concept CS.

$KF = \min(n,m)$.

Il est facile de vérifier que l'équation (10) satisfait les propriétés de mesure de similarité:

- 1) $\forall C \in \text{Ontology}, Sim_F(C,C)=1$
- 2) $\forall x, y \in \text{Ontology}, Sim_F(x,y) = Sim_F(y,x)$
- 3) $1 \geq Sim_F \geq 0$

5.5.6.3 Appariement des frères

La similarité entre les frères entre ontologies est calculée de la même manière que pour les pères :

$$Sim_B(CR, CS) = \frac{\sum_{l=1}^{k_B} (Q_l * (1 - |Sim_{Str}(CR, CR_{Bl}) - Sim_{Str}(CS, CS_{Bl})|))}{kMax_B} \quad (11)$$

Où:

$$kMax_B = Max(n, m)$$

Q_l : Représente la similarité entre les couples des frères $(CR_l^{CRE}, CS_l^{CRE}) \in CRE(CR, CS)$ qui contient uniquement les frères. $CR_{Bl} \in Frères(CR)$ et $CS_{Bl} \in Frères(CS)$

5.5.6.4 Apparieur des fils

La similarité entre les fils entre ontologies est calculée de la même manière que pour les frères :

$$Sim_C(CR, CS) = \frac{\sum_{l=1}^{k_C} (Q_l * (1 - |Sim_{Str}(CR, CR_{Cl}) - Sim_{Str}(CS, CS_{Cl})|))}{kMax_C} \quad (12)$$

Q_l : représente la similarité entre les couples de fils $(CR_l^{CRE}, CS_l^{CRE}) \in CRE(CR, CS)$ qui contient seulement les fils.

$$CR_{Cl} \in Fils(CR) \quad \text{et} \quad CS_{Cl} \in Fils(CS)$$

→ Finalement, la similarité du contexte externe ou du voisinage est donnée par :

$$Sim_V(CR, CS) = \alpha_F * Sim_F(CR, CS) + \alpha_C * Sim_C(CR, CS) + \alpha_B * Sim_B(CR, CS) \quad (13)$$

$$\alpha_F + \alpha_C + \alpha_B = 1$$

$$Si \text{ Père}(CR) = \emptyset \vee \text{Père}(CS) = \emptyset \quad \text{Then } \alpha_F = 0$$

$$Si \text{ Fils}(CR) = \emptyset \vee \text{Fils}(CS) = \emptyset \quad \text{Then } \alpha_C = 0$$

$$Si \text{ Frère}(CR) = \emptyset \vee \text{Frère}(CS) = \emptyset \quad \text{Then } \alpha_B = 0$$

Exemple

Nous allons démontrer cela par un exemple simple, étant donné dans ce qui suit deux ontologies et deux concepts pour aligner:

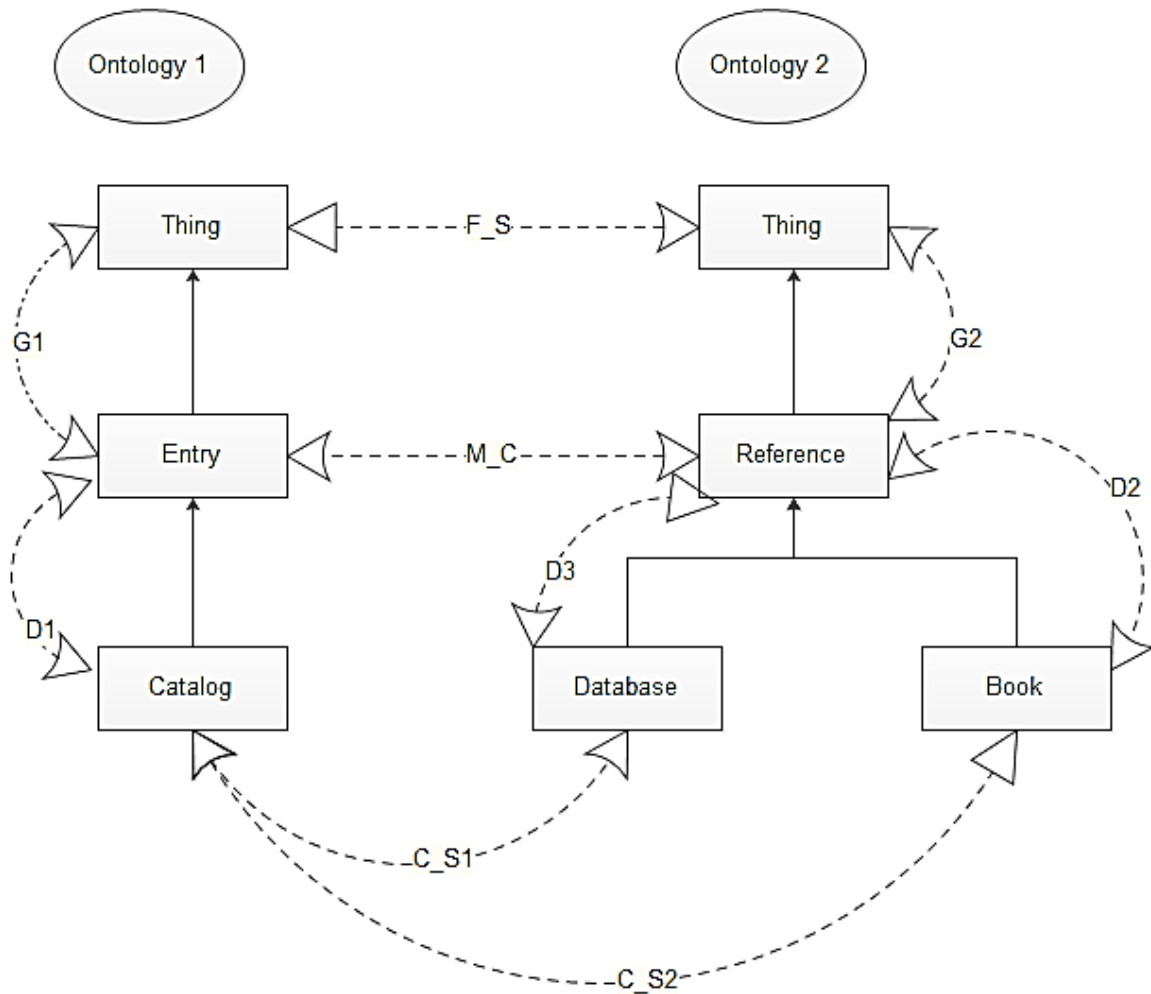


Figure 5.7: La similarité des contextes externes des deux concepts: Entry & Reference.

CR =le concept Entry et CS =le concept Reference;

$V(CR)=\{Thing, Catalog\}$;

$V(CS)=\{Thing, Database, Book\}$;

$P\grave{e}re(CR)=\{Thing\}$; $P\grave{e}re(CS)=\{Thing\}$

$Fils(CR)=\{Catalog\}$; $Fils(CS)=\{Catalog, Book\}$

$REC(CR, CS)=\{(Thing, Thing), (Catalog, Database), (Catalog, Book)\}$

La matrice de similarités des pères (MF note dans l'équation 12) est réduite à un simple élément

$Q1 = Sim_{Names}(Thing, Thing)=F_S$ (Dans la Figure 5.5) =1;

La matrice de similarités entre les fils est:

$$MC = \begin{pmatrix} Sim_{Names}(Catalog, Database) \\ Sim_{Names}(Catalog, Book) \end{pmatrix}$$

$Q2 = Sim_{Names}(Catalog, Database)=CS1$ (Dans la Figure 5.5) ($CS1 > CS2$)

La mesure des granularités (mesure entre deux concepts au sein de la même ontologie, représente le niveau de détail qui sépare les deux concepts) est comme suit:

$G1 = Sim_{Str}(Entry, Thing)$ et $G2 = Sim_{Str}(Reference, Thing)$ pour les pères et

$D1 = Sim_{Str}(Entry, Catalog)$ et $D3 = Sim_{Str}(Reference, Database)$ pour les fils

Où $Sim_{Str}(CR, CS) = Sim_{WP}(CR, CS)$

Finalement:

$Sim_F(Entry, Reference) = Q1 * (1 - |Sim_{WP}(Entry, Thing) - Sim_{FWP}(Reference, Thing)|)/1$

$Sim_C(Entry, Reference) = Q1 * (1 - |Sim_{WP}(Entry, Catalog) - Sim_{WP}(Reference, Database)|)/1$

$Sim_V(Entry, Reference) = 0.5 * Sim_F(Entry, Reference) + 0.5 * Sim_C(Entry, Reference)$

Selon les données de la Figure 5.5:

$Sim_V(Entry, Reference) = 0.5 * (F_S * (1 - |G1 - G2|) + 0.5 * (CS1 * (1 - |D1 - D2|))$

Où:

$\alpha_F = \alpha_C = 0.5$ et $\alpha_B = 0$.

5.5.6 Similarité globale (Global Matcher)

L'agrégation de toutes les similarités dans une mesure de similarité globale est donnée par l'équation suivante :

$$Sim_{overall_sim}(CR, CS) = W_{Names} * Sim_{Names}(CR, CS) + W_{Prop} * Sim_{Prop}(CR, CS) + W_V * Sim_V(CR, CS) \quad (14)$$

$$W_{Names} + W_{Prop} + W_V = 1$$

Si $Prop(CR) = \emptyset \vee Prop(CS) = \emptyset$ Alors $W_{Prop} = 0$

5.6 Matchmaking des services web basé sur l'alignement partiel des ontologies

L'efficacité de l'algorithme d'appariement entre une requête et les services offerts est un paramètre crucial pour une bonne découverte sémantique. Le processus d'appariement sémantique entre un service requis et service offert commence par un appariement élémentaire à chaque niveau d'éléments de la description (c.-à-d. appariement entre interfaces (ou porttypes) requises et offertes, appariement entre opérations requises et offertes, etc.). Les résultats de la première phase alimentent l'appariement global qui est basé sur une stratégie déterminée pour générer le degré de match global. Notre deuxième proposition consiste en trois parties essentielles imbriquées:

Un algorithme d'appariement des services web SAWSDL. Les services web peuvent être annotés par des ontologies différentes. Le composant noyau de notre approche est un algorithme d'alignement partiel, qui utilise une variété de mesures de similarités. Nous supposons que le nombre de services offerts sont déjà enregistrés dans un registre central. Soit une requête de service, alors les services

correspondants sont identifiés. La liste des offres sera ordonnée en respectant la similarité relative au service offert.

La proposition (Fellah et al., 2016b) consiste à :

- 1- Un multi Matchmaker pour les services web SAWSDL appelé SAWSDL-MOM, Avec une architecture qui instancie le Framework de notre première proposition
- 2- Le noyau du multi Matchmaker incorpore un algorithme d'alignement partiel de deux ontologies. À notre connaissance, il n'y a aucune autre approche qui utilise la technique d'alignement partiel des ontologies pour l'appariement des services web. Dans le présent travail, nous sommes concentrés sur l'usage des techniques d'alignement dans l'appariement des services web.
- 3- Une mesure de similarité entre concepts de deux ontologies qui représente une amélioration de notre première mesure de similarité
- 4- L'usage de l'algorithme hongrois dans les situations de besoin d'affectation 1:1 au lieu d'une approche gourmande (a greedy approach)
- 5- L'utilisation d'une stratégie simple pour l'obtention du degré de match global

5.6.1 Spécification

L'idée de base de l'approche de Matchmaking utilisée dans SAWSDL-MOM est l'interopérabilité entre les modèles sémantiques qui annotent la requête et les services offerts disponibles, principalement par la dérivation des relations entre les éléments ontologiques spécifiés dans la requête et les services offerts. Les DoMs sont numériques dans la gamme 0-1. SAWSDL-MOM renforce cette approche en appliquant des appariements à différents niveaux d'abstraction de service, l'intégration d'une approche basée sur l'alignement partiel des ontologies, qui utilise des similarités : syntaxique, linguistique et structurelle ; enfin génère un ensemble de groupe de mappings. La valeur de similarité globale est le résultat d'agrégation des résultats partiels.

➤ Restrictions et principes

- La requête est décrite avec le langage SAWSDL
- Les services offerts sont décrits avec le langage SAWSDL
- Les ontologies qui annotent la requête et le service peuvent être différentes
- Les ontologies sont décrites uniquement avec le langage OWL
- Seulement, une seule et première annotation est considérée.

5.6.2 Identification des éléments à appairer

SAWSDL possède une partie abstraite qui définit ce que fait un service. Cette définition abstraite est constituée à partir des PortTypes (Interfaces pour ((SA)WSDL 2.0) et les opérations correspondantes et des messages d'entrée et de sortie, qui sont habituellement définies en utilisant des types XSD. Comme il a été mentionné, la spécification SAWSDL ne précise pas laquelle des parties d'une description WSDL qui doit être annotée sémantiquement. Dans la situation, où tous les niveaux d'abstraction de service web peuvent être parfaitement décrits par des concepts sémantiques, un Matchmaker basé purement sur la sémantique pourrait être suffisant. Dans les autres cas, il est fort probable que des grandes parties de la description de service ne sont pas annotées sémantiquement. Dans cette situation, d'autres techniques de Matchmaking complémentaires aux techniques sémantiques pourraient être nécessaires. Seuls certains éléments (dans la majorité des cas, la signature service, à savoir, les entrées et sorties) sont utilisés, dans la plus part des approches. Les approches se différencient sur la façon d'intégrer et de combiner les similarités des différents niveaux dans une valeur globale de similarité pour le service tout entier. L'agrégation de similarité les valeurs de différents niveaux abstraction de service n'est souvent pas explicitement considérées. Une autre différence est la manière de combiner les différentes mesures de similarité pour chaque niveau abstraction de service. Il semble difficile de proposer une manière universellement valable pour combiner les différents niveaux d'abstraction de service. La restriction à certains niveaux de description de services pourrait être approprié dans des situations et non pas dans des autres. Pour notre cas, au niveau conceptuel, il est très facile d'inclure tous le niveau de descriptions d'un service SAWSDL. D'une part, notre objectif prioritaire est le traitement de l'interopérabilité sémantique. D'autre part nos tests sont limités à une collection de tests qui n'est pas natives dont les annotations sont seulement aux niveaux des messages. Donc, nous avons limité notre approche aux contraintes exposées.

5.6.3 Instanciation du Framework pour SAWSDL

Pour développer une approche de Matchmaking des services web SAWSDL, En premier lieu, nous avons instancié le cadre général présenté dans la *Figure 5.2* aux services web spécifiques SAWSDL (voir *Figure 5.8*). L'architecture contribue à la résolution de différentes issues discutées dans les sections précédentes.

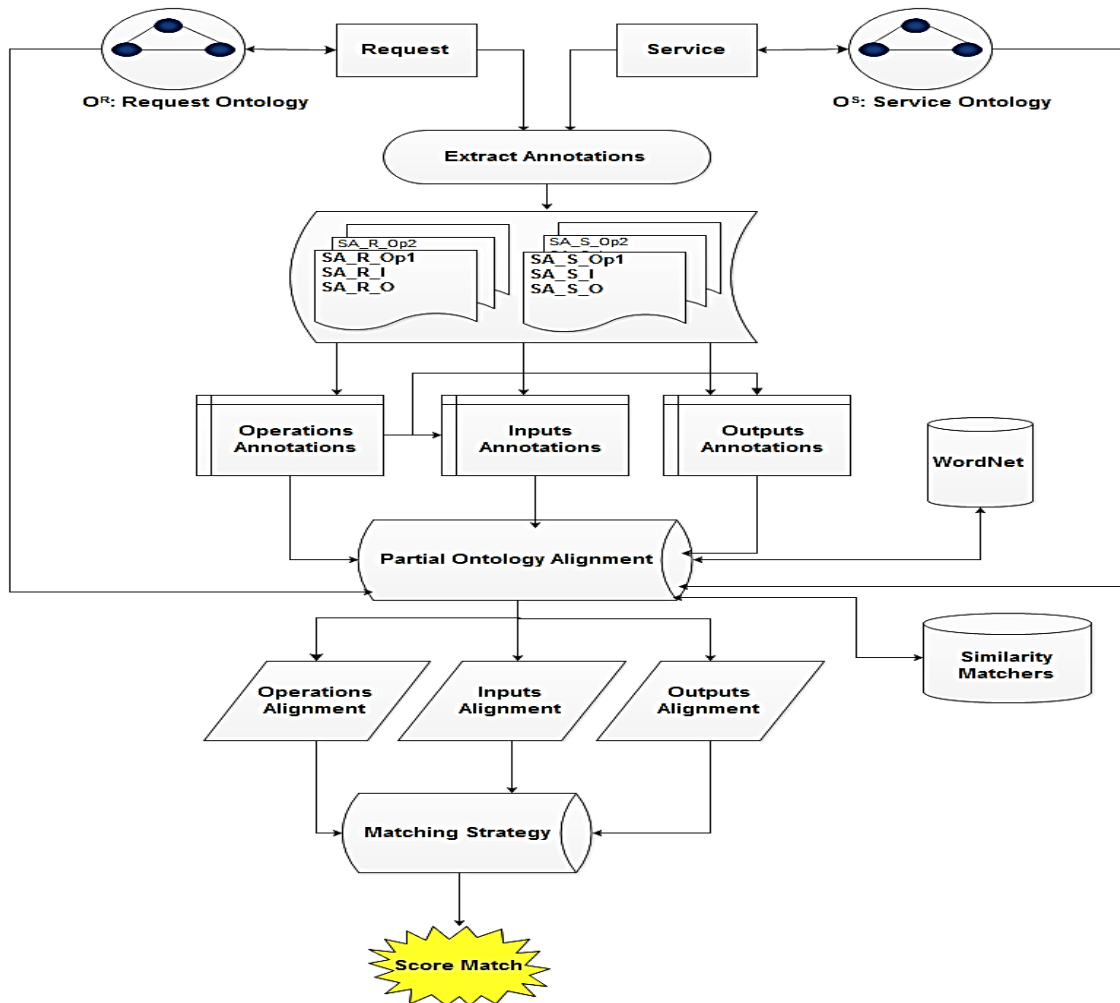


Figure 5.8: Le processus de Matchmaking SAWSDL-MOM.

L'appariement des services prend comme entrées deux descriptions (une requête et un service offert) et retourne un degré d'appariement (DoM) entre eux.

Les deux descriptions (requête et service offert) sont exprimées avec le modèle commun unifié SAWSDL.

On suppose que les modelReferences dans les requête et services SAWSDL pointent sur des concepts ontologiques, exclusivement définies dans OWL (La transformation du modèle de description du service dans un modèle de services commun est hors de notre travail).

Après un parsing des descriptions en entrée, le premier pas dans le Matchmaking de deux services consiste à extraire les annotations sémantiques de chaque modèle de description (SAWSDL).

Une fois l'extraction est complétée, les concepts utilisés pour les annotations sont groupées par catégories,

Chaque catégories (c.-à-d. : les opérations, les entrées, sorties) contient deux listes (pour la requête et pour l'offre) et pour chaque couple de deux listes seront considérée comme deux sous ontologies à aligner avec un algorithme d'alignement partiel,

L'algorithme d'alignement utilise les ontologies de base comme connaissance externe dans le processus, ainsi que d'autres ressources externes comme WordNet. Une bibliothèque de mesures de similarité est utilisée avec une stratégie par l'algorithme d'alignement partiel. Pour un usage optimal des ressources, il est possible d'avoir une cache pour ne pas recalculer les correspondances déjà trouvées.

Chaque résultat d'appariement d'une catégorie est agrégé dans un seul résultat unique, par exemple, on obtient un degré global pour la catégorie des inputs. La stratégie d'appariement des services détermine comment on combine les scores locaux pour obtenir un score final. On utilise l'algorithme hongrois pour déterminer les correspondances 1:1. Cette vision est implémentée par l'algorithme de la section suivante :

5.6.4 Algorithme d'appariement des services web SAWSDL (SAWSDL Matchmaking)

La *figure 5.9* présente le détail d'un processus de Matchmaking des services web SAWSDL, qui considère tous les niveaux de description d'un service : Au départ un calcul de similarité élémentaire est opéré au niveau individuel. Selon une stratégie choisie au préalable, le résultat des agrégations de similarité à chaque niveau de description est propagé à un autre niveau, pour une agrégation globale au niveau du service tout entier.

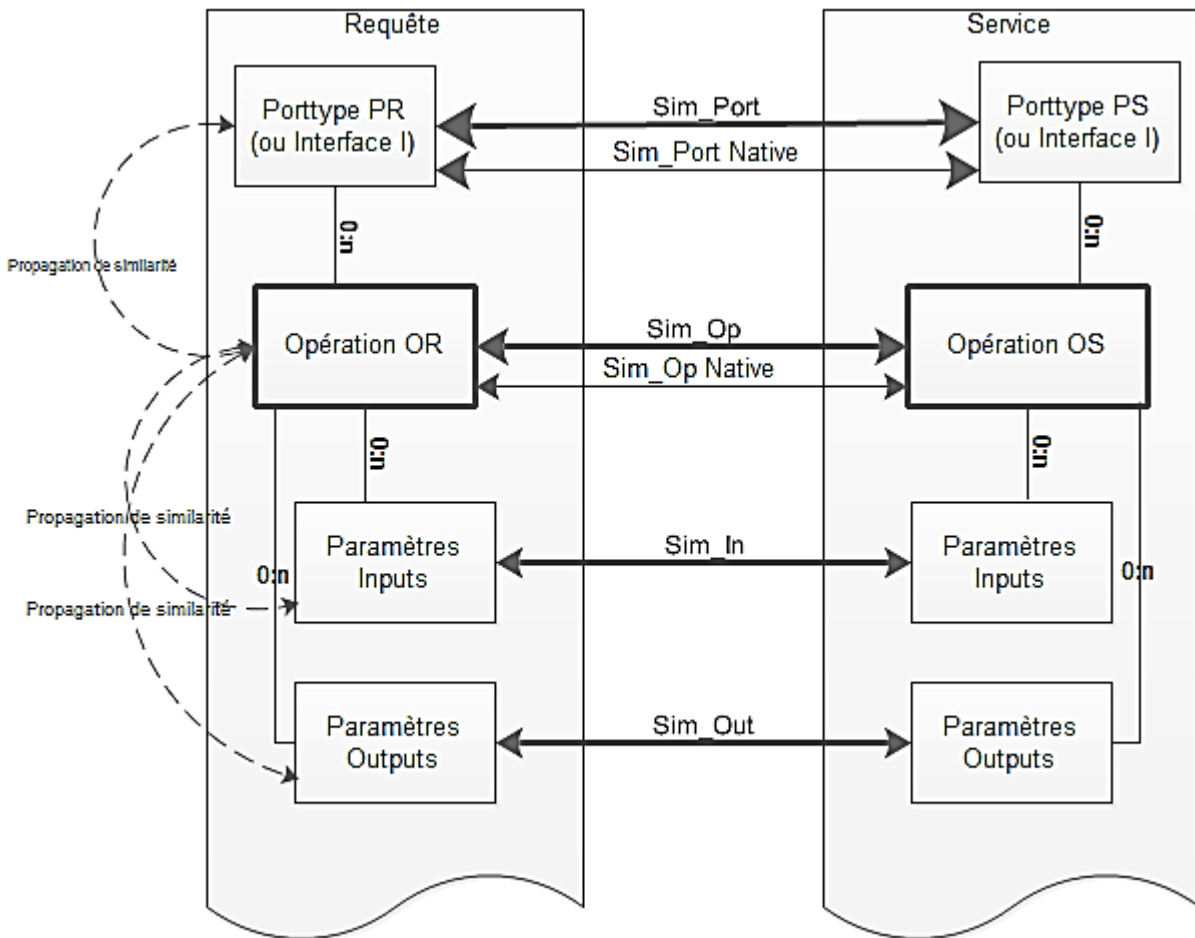


Figure 5.9 : Détail du processus de Matching

Par exemple si une approche se concentre sur les opérations alors, la propagation des similarités natives des Porttypes, des paramètres Inputs et des Paramètres Outputs seront propagée et agrégée avec la similarité native de l'opération pour obtenir une similarité de l'opération.

Dans la collection de test SAWSDL-TC, les annotations sont au niveau des messages et les services sont atomiques, et notre préoccupation principale est de surmonter les hétérogénéités des annotations entre la requête et le service par l'application de similarité inter ontologies. Donc nos idées peuvent s'appliquées seulement au niveau des messages.

Trois procédures ont été conçues pour le processus d'appariement de services web :

- (i) Extraction des concepts : l'Extraction des concepts est conçue pour minimiser le nombre de concepts à mapper entre deux ontologies telles que l'alignement des ontologies en ligne efficace peut être atteint. Il est particulièrement important pour l'alignement de grandes ontologies. L'extraction des concepts de la requête (les lignes (11) et (13) pour les inputs et les lignes (14) et (17) pour les outputs). Les concepts du service sont extraits (les lignes (22) et (24) pour les inputs et les lignes (25) et (27) pour les outputs).

Chapitre 5 : Matchmaking des SWS dans un contexte multi-ontologie

Il est possible de remplacer les commentaires des lignes (10 et 21), afin d'extraire les annotations des opérations s'ils existent.

- (ii) Le calcul de l'alignement partiel (les lignes (28) et (29)) entre catégorie des messages. Le vecteur `I_Match` est le résultat de l'appariement entre les concepts d'entrée de la requête et les concepts d'entrée du service. Egalement, `O_Match` est le résultat d'alignement des sorties.
- (iii) Agrégation des résultats de l'alignement partiel (`I_Match` et `O_Match`) pour le calcul des similarités locales (les lignes (30) et (31)) aux niveaux des messages entrée et sortie.
- (iv) Calcul de la similarité globale (GDOM : Global Degré Of Match ; la ligne (32)) du service tout entier.

Le pseudo code de l'algorithme d'appariement est le suivant:

```
(1) Algo Matchmaking
(2) Inputs : Request, offert
(3) Outputs : Results
(4) Parameters: Alignment_Cache
(5) {
(6) //Si le test de cardinalité OK
(7) For each Operation in Request do
(8)   {
(9)     // Si les operations ne sont pas annotées sinon extraction des annotations
(10)    //Extraction des annotations sémantiques de la requête
(11)    For each input in Operation do
(12)      // SALRI :Semantic Annotations List of Request Inputs
(13)      SALRI.add(input.ModelReference)
(14)    For each output in Operation do
(15)      // SALRO :Semantic Annotations List of Request Outputs
(16)      SALRO.add(output.ModelReference)
(17)   }
(18) For each Operation in Offert do
(19)   {
(20)     // Si les operations ne sont pas annotées sinon extraction des annotations
(21)     //Extraction des annotations sémantiques du service offert
(22)    For each input in Operation do
```

```
(23) // SALRI :Semantic Annotations of Offert Inputs
(24) SALOI.add(input.ModelReference)
(25) For each output in Operation do
(26)     SALOO.add(output.ModelReference)
(27) }
(28) I_Match=Partial_alignment(SALRI,SALOI);
(29) O_Match=Partial_alignment(SALRO,SALOO);
(30) Sim_Inputs=agregate(I_Match);
(31) Sim_Outputs=agregate(O_Match);
(32) Result=(Sim_Inputs+Sim_Outputs)/2;
(33) // Si les operations sont annotées alors propagation des similarités
(34) Return(Result)}
(35) }
```

Figure 5.10 (Listing): Algorithme de Matchmaking

Le commentaire de la ligne (9) suppose l'absence des annotations sémantiques des opérations par des concepts ontologiques. Dans le cas contraire, un calcul préalable de similarités natives entre les opérations est nécessaires (on procède à un alignement partiel entre les annotations des opérations en premier lieu, puis on propage les résultats des similarités des paramètres d'entrée et des paramètres de sorties). Le commentaire de la ligne (6) suppose aussi, que le test de cardinalités est vérifié :

- Nombre (paramètres inputs du service offert) \geq Nombre (paramètres inputs de la requête) et
- Nombre (paramètres outputs de la requête) \geq Nombre (paramètres outputs du service offert)

Le détail de l'algorithme d'alignement partiel est donné dans la section suivante.

5.6.5 Alignement partiel

L'appariement des modèles sémantiques est une opération importante dans les applications traditionnelles, telles que l'intégration des ontologies, l'intégration des schémas ou les entrepôts de données. Généralement, ces applications sont caractérisés par des modèles structurels hétérogènes qui sont analysés et appariés soit manuellement ou semi automatiquement au moment de la conception. Par contre, un autre type d'applications qui peuvent être caractérisées par leur dynamique, par exemple, les agents, les systèmes peer-to-peer, les services Web ; nécessitent une opération d'appariement en temps réel (Euzenat et al., 2004). Notre travail propose un mécanisme d'appariement partiel des

ontologies en ligne pour répondre aux défis de l'appariement des ontologies dans un environnement informatique dynamique. Si les concepts à comparer sont définies dans des ontologies différentes, alors un alignement sémantique est nécessaire.

Dans notre travail, nous sommes concentrés sur les techniques d'alignement des ontologies et en même temps l'usage des alignements. Donc la nature des outils d'appariement des ontologies est différente par rapport à notre mécanisme d'appariement. Le fait que le service ou la requête ne sont pas annotés par la totalité des concepts des ontologies alors l'appariement des concepts ne prend en compte que les concepts réellement impliqués dans l'annotation. Dans la littérature un alignement entre deux ontologies qui considère toutes les entités appartenant aux ontologies en entrée d'un processus de matching est désigné par alignement total, par contre la considération de deux sous-ensembles des ontologies est appelé alignement partiel.

Dans un contexte d'annotation des services web avec plusieurs ontologies, il est possible d'annoter un service unique par plusieurs ontologies. Dans ce cas, une vision plus pragmatique du problème d'appariement de deux descriptions de services web peut être considérée comme un problème de multi alignement ; nous nous limitons dans notre étude au problème d'alignement de deux ontologies.

Dans notre appariement on suppose :

- Un appariement partiel qui considère seulement quelques éléments des ontologies. Seulement les concepts sont utilisés pour l'annotation de la requête et du service. Donc on aura l'entrée de l'algorithme deux listes de concepts (la ligne (2) de l'algorithme d'alignement partiel).
- La cardinalité des correspondances résultats entre les entités est un-un, ce qui nous oblige à filtrer les correspondances après calcul des similarités (la ligne (8), la fonction `Compute_similarity`) entre les entités des deux sous ontologies en entrée, pour garder que les meilleurs appariements, on utilise une fonction d'affectation optimale, qui implémente l'algorithme Hongrois (la ligne (09), la fonction `optimal_assignment`) qui sera détaillée dans la section suivante.

```
(1) Algo partial_alignment
(2) Inputs :LCS,LCC: two lists of concepts
(3) Outputs:Result:one to one mappings
(4) Parameters: ontologies, wordnet,basic matchers, Alignment_Cache
(5) {
(6) For each cs in LCS do
(7) For each cc in LCC do
(8)     {
(9)         Sim(i,j)=Compute_similarity(cs,cc);
(10)        Result=optimal_assignment(Sim);
(11)    }
(12) }
```

Figure 5.11(Listing) : *Algorithme d'Alignement Partiel*

Le mode de calcul de similarité (ligne (8), Fonction Compute_similarity) est détaillé dans la section (5.6.6) qui est une amélioration de la mesure présentée dans la section (5.6).

5.6.5.1 Affectation optimale avec l'algorithme Hongrois (ALGORITHME DE KÜHN)

L'algorithme hongrois ou méthode hongroise (Kuhn, 2010), (parfois appelé aussi algorithme de Kuhn-Munkres) est un algorithme d'optimisation combinatoire, qui résout le problème d'affectation en temps polynomial. Il a été proposé en 1955 par le mathématicien américain Harold Kuhn, qui l'a baptisé « méthode hongroise » parce qu'il s'appuyait sur des travaux antérieurs de deux mathématiciens hongrois : Dénes König et Jenő Egerváry. James Munkres a revu l'algorithme en 1957 et a prouvé qu'il s'exécutait en temps polynomial.

L'algorithme Hongrois est un algorithme permettant de résoudre le problème d'affectation. Il s'agit d'une procédure itérative qui transforme une matrice des coûts en une suite de matrices équivalentes, jusqu'à ce que l'obtention d'une solution optimale soit évidente. La matrice finale est telle que toutes les entrées sont soit positives ou nulles et qu'une affectation faisant appel seulement aux entrées nulles est possible. Cette affectation de coût 0 est alors nécessairement optimale.

L'algorithme hongrois minimise les couts, pour l'affectation. Dans notre cas d'affectation des alignements avec des degrés de correspondances, l'objectif est de maximiser les similarités entre couples. D'où l'idée de soumettre à la méthode non une matrice de similarité M , mais une matrice de dissimilarité $(1 - M)$. Donc, de cette manière on aura une affectation optimale qui est une

maximisation des degrés de similarité entre couples de concepts et par conséquents entre couples d'éléments de service web à apparier.

5.6.6 Mesure de similarité

L'algorithme d'alignement partiel utilise notre mesure de similarité proposée dans la section 5.5. Toutefois, dans un souci d'amélioration de la performance de notre mesure de similarité inter ontologies, nous avons fait un certain nombre de constatations :

H1: le besoin de déterminer les affectations 1 :1, peut être résolu efficacement par l'application de la méthode Hongroise

H2: Le contexte externe d'un concept est souvent connu par ses pères (Super concept) et Fils et frères du concept dans son ontologie. Souvent, un concept peut avoir ou non des sous concepts ou des frères mais toujours possède des super-concepts. Ceci veut dire que lors du calcul de la similarité contextuelle entre deux concepts, la similarité entre leurs supers concepts devrait être considérée seulement.

A partir nos hypothèses (H1, H2) et pour satisfaire nos critères sur la similarité ; plusieurs améliorations ont été opérée sur notre mesure présentée dans la première proposition (section 5.5).

Une première amélioration de la similarité consiste en usage de l'algorithme Hongrois pour la détermination des vecteurs d'affectations optimales à partir des matrices de similarités des propriétés internes des concepts ou des contextes externes des concepts.

A partir de la constatation C2, le père d'un concept est toujours existant. Donc une deuxième amélioration introduite dans la mesure est la réduction du contexte externe d'un concept à l'ensemble de ses pères (c.-à-d. : en prend en compte les pères directs et indirects).

Pour aligner le concept *CR* avec le concept *CS* appartenant à deux ontologies différentes, nous avons utilisé l'algorithme `Compute_Similarity` (section 5.5.1) qui invoque l'algorithme `Compute_Averall_Sim` (section 5.5.1). Pour introduire les améliorations introduites sur la mesure de similarité des contextes externes, une nouvelle version de l'algorithme nommée `Compute_Averall_Sim.1` est développée.

```
(1) Algo Compute_Averall_Sim.1
(2) Inputs :CR, CS: two concepts
(3) Outputs:result
(4) Parameters:WName,WProp,WFather
(5) {
(6) Sim_Name=Compute_Sim_Name(CR.name,CS.name);
(7) Sim_Prop=Compute_Sim_Prop(CR.prop, CS.prop);
(8) Sim_Vois=Compute_Sim_Father(CR, CS) ;
(9) Result=WName * Sim_Name+WProp * Sim_Prop+WFather *SimFather;
(10) Return(Result)
(11) }
```

Figure 5.12(Listing) : Calcul de similarité globale.

L'algorithme Compute_Averall_Sim.1 invoque les fonctions suivantes :

- Compute_Sim_Name : est l'implémentation de la mesure de similarité des noms(section 5.5.3).
- Compute_Sim_Prop : est l'implémentation de la mesure de similarité des propriétés ou du contexte interne, (section 5.5.4). Egalement, l'affectation optimale qui utilise l'algorithme gourmand est remplacée par l'algorithme Hongrois.
- Compute_Sim_Father : est l'implémentation de la mesure de similarité du contexte externe, présentée dans la section (5.5.5) avec l'introduction de l'amélioration citée ci-dessus (c.-à-d. : le contexte externe réduit seulement aux pères, avec la considération de tous les pères). Egalement avec l'usage de l'algorithme Hongrois pour l'affectation.

5.7 Vers une extension de SAWSDL-TC

SAWSDL-TC(2009) est l'unique collection de test disponible actuellement pour l'évaluation des matchmakers SAWSDL. Afin de procéder à des évaluations plus solides et vu les insuffisances de la base de test SAWSDL-TC (section 2.4.5 et section 6.3.2) dans un contexte multi-ontologie. Nous proposons des idées pour le développement d'une nouvelle base de test pour les multi-matchmaker SAWSDL. Afin de nous permettre de procéder à une évaluation dans une bonne exhaustivité par rapport à un réel environnement multi ontologies. Le principe de notre extension est inspiré de la base de test utilisée, pour évaluer les algorithmes d'alignement d'ontologies (OAEI(2014)). L'idée est de procéder à des altérations de l'ontologie (ontologie qui annote un service Web) qui sera considérée

Chapitre 5 : Matchmaking des SWS dans un contexte multi-ontologie

comme l'ontologie de base, générant de nouvelles ontologies, ce qui affecte les services annotés par l'ontologie de base et implique de nouveaux services Web annotés par les nouvelles ontologies. Tous les services Web ainsi générés, vont devenir des nouveaux tests. Le résultat sera une base de test adéquate pour le multi appariement des services web. Compte tenu de l'importance, de la taille et de la complexité de ce projet, le présent travail ne le traite pas dans sa totalité, seulement pour clarifier les idées, nous considérons un choix simplifié de SAWSDL-TC comme suit:

La requête

- BookPriceService

Et les services:

- NovelPriceService
- Short-storyAuthorbook-typeService

Exemple l'ontologie de base Books.owl et le service web BookPriceService annoté avec le concept #Book appartenant à l'ontologie Books.owl. Nous procédons aux modifications ou altérations suivantes:

Altération 1:

Construction d'une nouvelle ontologie books1.owl, exactement semblable à books.owl

Altération 2:

Construction d'une nouvelle ontologie books2.owl exactement semblable à books.owl renommant le concept #Book par #Text_Book.

Ces changements engendrent un second service Web BookPrice1 qui est annoté avec le concept #Text_Book

Altération 3:

La construction d'une nouvelle ontologie books3.owl, exactement semblable à l'ontologie books.owl renommant le concept #Book avec #hthtfh (une chaîne quelconque).

Et ainsi de suite, de la même façon que les altérations ont été opérées dans la base de test (OAEI, 2014). Nous avons utilisé pour créer nos ontologies OWL Protégé 4.0 et une version d'essai de SyncRO Soft, S. R. L. Oxygen XML Editor Version 17.1 (SyncRO Soft, 2015) pour modifier et annoter des services Web. Donc, nous pouvons exprimer toutes les hétérogénéités possibles entre ontologies pour tester l'efficacité des outils d'appariement. Par conséquent, les trois modifications de book.owl génèrent une ontologie pour chaque service de base de test SAWSDL-TC, trois autres services annotés avec le concept des variantes de #Book, et ainsi, la base sera enrichie par un ensemble

complet de cas d'études pour nous permettre de faire un bon jugement des matchmakers. Les expériences réalisées sur certains changements montrent clairement l'efficacité de nos idées et l'utilisation de l'alignement des ontologies par rapport à d'autres outils d'appariement centrés sur l'utilisation d'une ontologie commune. Un de nos projets dans le futur proche est le développement d'une base de données de test pour SAWSDL assez riche pour satisfaire tous les aspects et contraintes de l'interopérabilité des services Web dans un environnement multi-ontologie.

5.8 Conclusion

Dans ce chapitre, nous avons présenté un cadre général pour l'appariement des services web, lesquels peuvent être annotés par des ontologies différentes, nous avons mettons en exergue l'importance d'une mesure de similarité inter ontologies comme composant clé de n'importe quelle architecture de découverte de service web dans un contexte multi ontologie, également, nous avons proposé une mesure de similarité entre concepts de différentes ontologies. On se basant sur l'état de l'art d'appariement des services web sémantique, et notre première vision générale, nous avons proposé une architecture de Matchmaking des services web SAWSDL qui est une instance de la première proposition. Nous avons proposé un algorithme pour le Matchmaking des services web qui utilise un algorithme d'alignement partiel entre ontologies, ce dernier utilise une mesure de similarité inter ontologies, donc, une critique objective de la similarité proposée, nous a permis de dégager une amélioration de ladite similarité. La sélection des meilleures correspondances, entre concepts, est un problème d'affectation optimale. Nous avons utilisé une approche gourmande (*a greedy approach*) dans la première proposition et pour améliorer les performances, nous avons utilisé la méthode Hongroise dans la deuxième proposition.

Chapitre 6 : Evaluation & Expérimentation

Chapitre 6 : Evaluation & Expérimentation

6.1 Introduction

Afin de vérifier et de valider nos propositions en matière de performances de : la mesure de similarité, l'alignement partiel des ontologies et le Matchmaking de services Web sémantiques. Nous avons mis en œuvre un système adéquat. Le recours à des collections de tests standards de la communauté du web sémantique et du service Web a permis de vérifier nos propositions. La base de test OAEI de la compétition d'alignement des ontologies pour tester la mesure de similarité, ainsi que l'algorithme d'alignement partiel, et la base de tests standards SAWSDL-TC (SAWSDL-TC, 2009) pour tester le matchmaking des services web SAWSDL. Ce chapitre est organisé comme suit : on commence par présentation des indicateurs et des mesures clés de l'évaluation. Puis une description des bases standards utilisées dans notre évaluation. Ensuite, on décrit la philosophie et les parties de nos tests. Finalement, on se focalise sur le détail de nos tests, en terminant par une discussion de nos résultats.

6.2 Principe et indicateurs de l'évaluation

L'évaluation consiste à mesurer la différence entre un résultat attendu et un résultat obtenu. Aucune métrique n'y est associée mais on utilise en général un indicateur compris entre [0, 1], pour faciliter l'interprétation. Certains éléments sont très subjectifs et difficilement automatisables. Il faut mettre en avant l'importance du contexte sur la pertinence, ainsi que l'utilisateur(le juge). Par conséquent, si l'un de ces facteurs évolue, la notion de pertinence change également, et donc l'évaluation.

6.2.1 Les indicateurs et mesures

Un système peut répondre à une requête selon le modèle simple suivant :

	Pertinent	Non pertinent	Total
Retrouvé(ou proposé)	A	b	a+b
Non retrouvé(ou non proposé)	C	d	c+d
Total	a+c	b+d	a+b+c+d=N

A partir de ce tableau de contingence, la communauté scientifique calcul diverses mesures :

$$\text{la précision} = \frac{a}{a+b}; \text{ le rappel} = \frac{a}{a+c}; \text{ la pertinence} = \frac{a+d}{a+b+c+d}; \text{ l'erreur} = \frac{b+c}{a+b+c+d}$$

$$\text{le taux de chute} = \frac{b}{b+d}; \text{ le silence} = (1 - \text{rappel}) = \frac{c}{a+c}; \text{ la spécificité} = \frac{d}{b+d}$$

$$\text{le bruit} = \frac{b}{a+b}; \text{ l'overlap} = \frac{a}{a+b+c}; \text{ le généralité} = \frac{a}{a+b+c+d}$$

Au final, 4 mesures (a,b,c,d) génèrent 10 indicateurs de base, eux-mêmes combinés pour donner d'autres mesures. Dans la grande majorité des cas, on se limite à la précision et au rappel. À partir de ces différentes mesures plusieurs indicateurs de synthèses ont été créés, mais le plus célèbre est la F-mesure.

$$F - \text{mesure} = \frac{(1+\beta^2) \times (\text{Précision} \times \text{Rappel})}{(\beta^2 \times \text{Précision}) + \text{Rappel}} \quad \text{Avec } \beta^2 = 1$$

Nous observons que cette mesure ne prend pas en compte la pertinence et fonctionne sur un modèle binaire : une réponse est bonne ou fautive. Dans le cas de l'évaluation d'un algorithme de classification, la F-mesure est fournie comme un résultat global. On distinguera la macro moyenne (macro average) de la micro moyenne (micro average). La macro moyenne calcule une moyenne globale, tandis que la micro moyenne calcule la moyenne des mesures de chaque catégorie.

6.3 Bases de tests standards

Comme d'autres domaines scientifiques, en informatique, toute évaluation nécessite des données de test. Les données d'essai sont explicitement générées ou collectées pour être utilisées dans des évaluations.

6.3.1 La base de test de l'alignement des ontologies OAEI(2014)

La proposition de notre expérimentation est différente des travaux de matching des ontologies, lesquels utilisent généralement la collection des tests (Ontology Alignment Evaluation Initiative Benchmark Test Library) (OAEI, 2014). L'ontologie de base se compose d'un ensemble de références bibliographiques. Elle est une version allégée avec le nombre d'entités ontologiques par rapport à l'ontologie réelle. Chaque cas de test de l'indice de référence de base met en évidence une caractéristique de la seconde ontologie alignée avec la base de données de test. L'objectif fondamental de ces tests est d'envisager tous les aspects qui existent dans l'ontologie OWL et ensuite examiner son impact sur les résultats de l'alignement.

6.3.2 La base de test SAWSDL –TC

Pour tester notre algorithme SAWSDL-MOM d'appariement des SWS, on a utilisé la version 3 de la base de test SAWSDL-TC3 (SAWSDL-TC, 2009). Actuellement c'est l'unique collection de test qui existe pour la comparaison et l'évaluation des approches de matchmaking des services web SAWSDL. SAWSDL-TC3 consiste en une collection de services web WSDL 1.1 qui couvre différents domaines annotés sémantiquement. La base de test inclut les requêtes et les services. Les offres et les requêtes sont encodées en utilisant SAWSDL. Les concepteurs de la base ajoutent un ensemble de relevance

binaire graduée pour chaque requête, cette relevance peut être utilisée pour calculer les métriques d'évaluation. Dans SAWSDL-TC, les annotations sémantiques existent seulement au niveau paramètres du message.

6.4 Evaluation expérimentale

Pour valider nos propositions, nous avons procédé à trois types d'évaluations :

6.4.1 Evaluation de la mesure de similarité

Notre approche d'évaluation est différente des tests conventionnels qui utilisent les métriques conventionnelles telles que la précision et le rappel. Nous avons conduit une série de tests en utilisant le benchmark *Ontology Alignment Evaluation Initiative OAEI* (2014).

Nous avons exécuté 10 cas de tests; chaque test se réduit à un concept source de l'ontologie 101, et un autre concept dans une ontologie cible, laquelle peut être l'ontologie : 103, 104, 204, 224, 225 ou 304.

Les résultats sont donnés dans la Table 6.1 qui indique les concepts testés et le résultat d'appariement correspondant. Les liens constatés, entre les concepts dans notre test, en se basant sur le principe que les mesures de similarité doivent avoir une forte corrélation avec le jugement humain. Pour l'humain, trois catégories des liens entre les concepts ont été identifiées:

- **Low semantic relatedness (LSR):** proximité sémantique faible : deux concepts sont faiblement connectés et une bonne mesure de leur similarité doit être inférieure à 0.5
- **Strong semantic relatedness (SSR):** forte proximité sémantique : deux concepts très fortement liés alors une bonne mesure de leur similarité converge vers 0.75
- **Semantic equivalence (SE):** équivalence sémantique : deux concepts sont équivalents alors une bonne mesure de leur similarité converge vers 1.

Cas	Concept de la requête	Concept du service	103	104	203	204	224	225	304	Expected Results	NER
1	Book	Book	0,99	0,99	0,99	0,65	0,99	0,99	0,59	SE	1
2	Article	Article	0,99	0,99	0,99	0,99	0,99	0,99	0,99	SE	1
3	MastersThesis	MastersThesis	0,84	0,84	0,84	0,00	0,84	0,84	0,64	SE	1
4	Book	Article	0,75	0,75	0,75	0,45	0,75	0,75	0,45	SSR	0.75
5	Part	Article	0,80	0,80	0,80	0,49	0,80	0,80	0,38	SSR	0.75
6	PhdThesis	MastersThesis	0,68	0,68	0,68	0,72	0,68	0,68	0,52	SSR	0.75
7	Report	Deliverable	0,48	0,48	0,48	0,25	0,48	0,48	0,25	LSR	0.55
8	Reference	School	0,37	0,37	0,37	0,37	0,37	0,37	0,53	LSR	0.4
9	Part	Misc	0,44	0,44	0,44	0,22	0,44	0,44	0,22	LSR	0.4
10	TechReport	Journal	0,38	0,38	0,38	0,31	0,38	0,38	0,38	LSR	0.4

Table 6.1: Les résultats des tests de notre mesure.

NER (Numeric Expected Results) : Résultats numériques attendus.

En se basant sur les résultats (la table 6.1) des évaluations en matière de convergence entre les valeurs de similarité et le jugement humain, nous avons dressé un graphe (Figure 6.1) illustrant le comportement de la mesure par rapport à l'avis de l'expert.

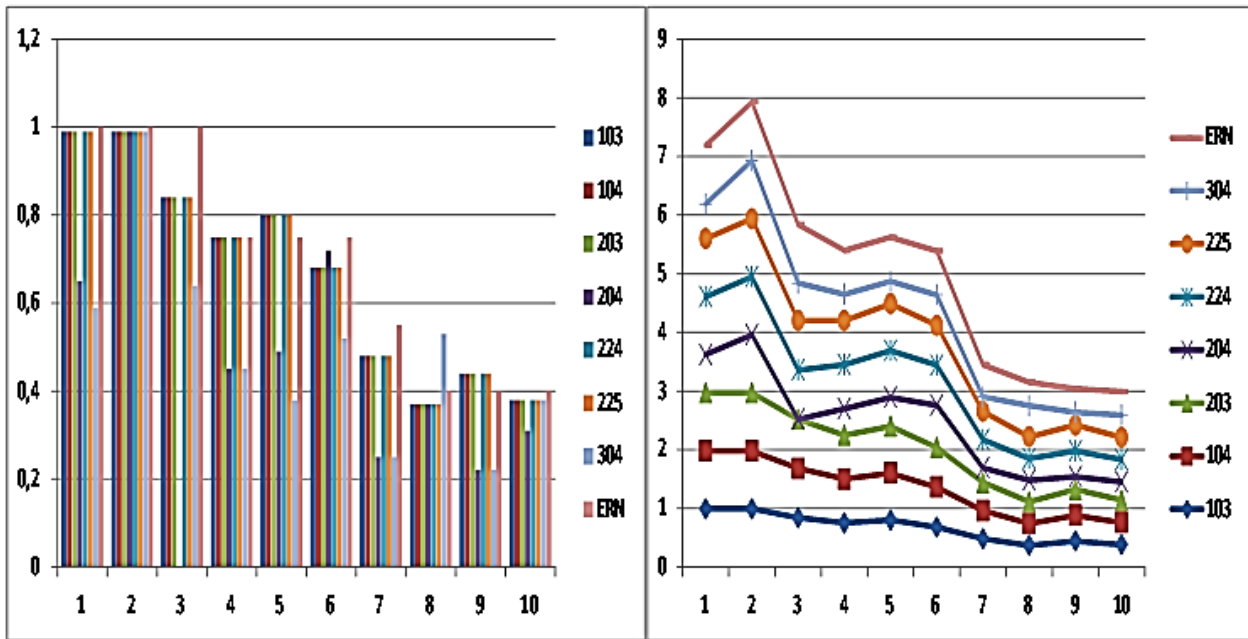


Figure 6.1 : Résultats graphiquement (tests de notre mesure)

On remarque que les résultats ont un comportement similaire à celui du résultat escompté (corrélation avec le jugement humain).

Dans les résultats Table 1 couleur verte indique les concepts sémantiquement équivalents selon la base de test. La couleur jaune exprime un lien étroit entre les concepts qui peut être une subsumption; les concepts avec la couleur bleue ont une distance relativement grande; la couleur rouge pour les concepts sémantiquement divergents.

Comme on le voit sur la Table 6.1 et la figure 6.1, les résultats expérimentaux confirment la viabilité de la mesure de similarité sémantique proposée dans le présent travail.

Dans la figure 6.1, les courbes de nos expérimentations (Sept courbes parmi huit) ont la même tendance, que la huitième courbe qui représente le jugement humain.

➤ Evaluation expérimentale de la deuxième proposition

Pour valider le matchmaking des services web SAWSDL, notre expérimentation est divisée en deux parties, la première est réservée à l'évaluation de l'algorithme d'alignement partiel et la deuxième pour l'évaluation de notre algorithme de matchmaking avec SAWSDL-TC.

6.4.2 Evaluation de l'algorithme d'alignement partiel

La nature des outils de matching des ontologies est différente de notre mécanisme. Une comparaison équitable de performance est difficile à atteindre. Notre mécanisme de matching tentera de trouver les

correspondances seulement pour les concepts apparaissant dans la requête, tandis que les outils d'alignement cherchaient les mappings pour tous les concepts (voire toutes les entités)

À partir de cette dernière remarque, les expériences doivent être bien conçues. Les résultats des évaluations sont présentés dans la dernière partie de cette section. Notre objectif est de concevoir un algorithme d'alignement partiel et efficace pour un environnement dynamique. Efficacité peut être mesurée par la vitesse et la précision des résultats proposés. La vitesse de l'algorithme est mesurée par le temps nécessaire pour l'exécution du matching. Pour mesurer la précision d'alignement partiel entre deux ontologies, des correspondances des références calculées manuellement sont nécessaires. On a procédé à une série de tests avec la base de tests « Ontology Alignment Evaluation Initiative » OAEI(2014). L'ontologie de base consiste en un ensemble de références bibliographiques. La base de test se compose d'un ensemble d'ontologies de références bibliographiques. L'ontologie de base est une version allégée en nombre d'entités ontologiques par rapport à l'ontologie réelle. Chaque cas de test de la collection met en évidence une caractéristique de la seconde ontologie à aligner. L'objectif fondamental de ces tests est d'assumer tous les aspects qui existent dans l'ontologie OWL et ensuite examiner quel impact cela pourrait avoir sur les résultats de l'alignement. Nous avons exécuté 5 cas de tests; chaque test considère un ensemble d'un certain nombre de concepts (par exemple 2, 4, 6, 8 et 10 concepts) de l'ontologie 101 et un autre ensemble de concepts dans une ontologie cible (l'ontologie cible peut être l'une des ontologies 101,103, 104, 202, 230, 301 et 304).

Dans ce qui suit les différents tests et leurs résultats. *La figure 6.3* représente un aperçu complet des résultats. Le graphe résume l'affichage d'un tableau de valeurs de précision, rappel et f-masure. Pour chaque couple d'ontologies, ensuite, la moyenne des performances obtenues dans chaque cas de test est calculée.

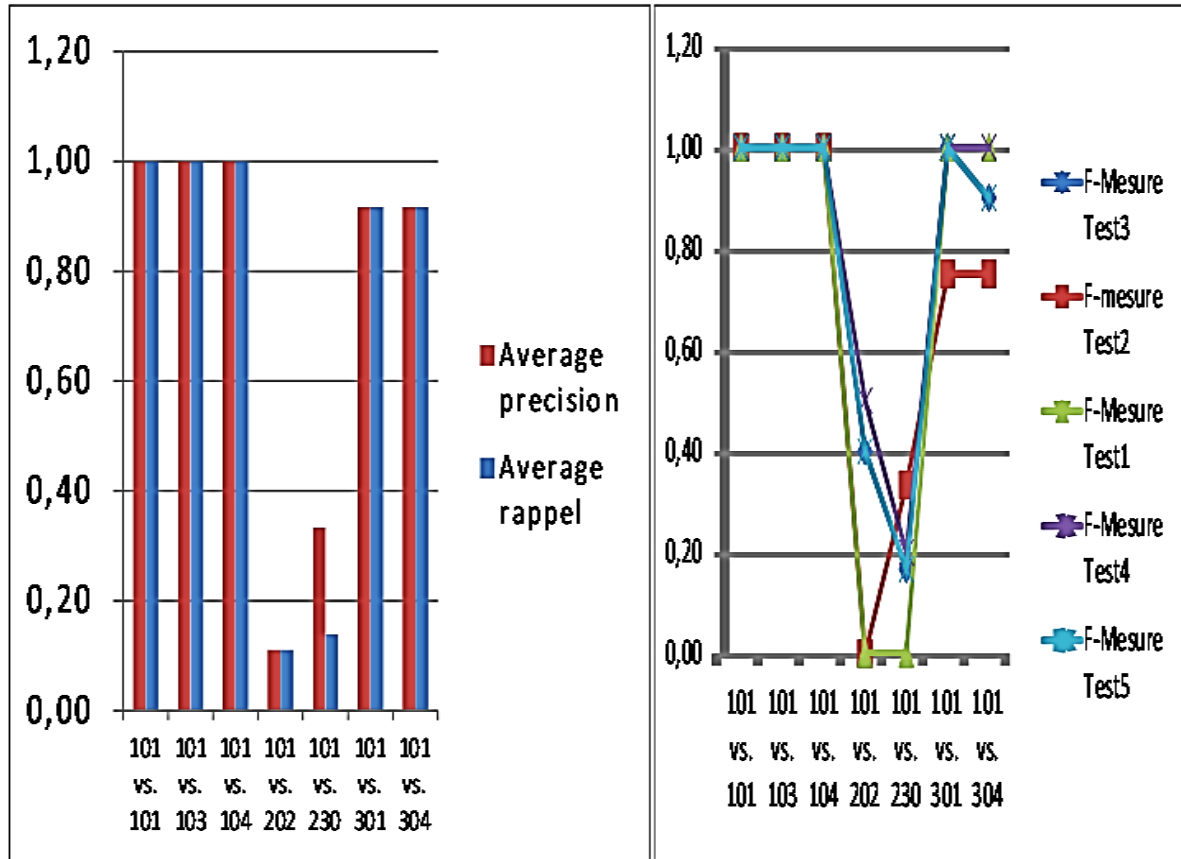


Figure 6.2: Les tests de l'Alignement Partiel.

L'alignement produit à chaque test est comparé à l'alignement de référence. Ainsi, des valeurs de mesure de qualité (précision, rappel et F-mesure) sont déterminées. Dans la *figure 6.2*, les moyennes des performances (précision, rappel et F-mesure) des 5 cas de tests sont représentées. Les meilleurs résultats sont obtenus lorsque les structures des ontologies sont similaires (ou identiques), à savoir les familles de tests 10x et 30x. Ainsi, notre méthode obtient des valeurs de précision pour ces tests qui sont égaux à 1,00. Cela est dû au fait que notre approche explore les structures des entités pour aligner par la similarité structurelle.

6.4.3 Evaluation de l'algorithme de Matchmaking SAWSDL-MOM

Nous avons effectué une série de tests à l'aide de la collection de test SAWSDL-TC3. Actuellement, elle est la seule collection de test standard pour le matchmaking de services SAWSDL et elle est utilisée pour la comparaison entre approches. SAWSDL-TC3 se compose d'une collection services Web WSDL 1.1, sémantiquement annotés, qui couvrent des domaines différents. La collection de test contient les et les offres. Les requêtes et les offres sont codées en utilisant SAWSDL. Les concepteurs de base ont ajouté, une pertinence binaire et graduée définie pour chaque requête, qui peut être utilisée pour calculer des métriques d'évaluation. Dans SAWSDL-TC, les annotations sémantiques existent

uniquement au niveau du message de paramètres. SAWSDL-MOM intègre des informations provenant du niveau des paramètres de message SAWSDL et il est possible de l'étendre à d'autres niveaux. Dans ce qui suit, nous allons présenter le plus important de notre expérimentation. L'approche de multi matchmaking, présentée a été mise en œuvre dans SAWSDL-MOM (Muli Ontology Matcher); utilisant un parser pour SAWSDL ainsi que l'usage de l'API OWL et du raisonneur Pellet pour les ontologies et JWNL comme interface pour WordNet. La collection de test, SAWSDL-TC3 a été adoptée. Conformément à la procédure dans le concours S3, nous avons évalué les mesures classiques adoptées en RI(Recherche d'Information), les mesures sont automatiquement calculées par SAWSDL-MX V2.0 comme dans l'outil SME2(Service Semantic Matchmaker Evaluation Environment (Klusch, 2012): fait partie du Framework d'évaluation S3. Il permet à l'utilisateur d'effectuer une évaluation de la performance de recherche des Matchmakers sémantiques de services sur les collections de tests). Dans les résultats de l'évaluation, nous avons adopté les mesures de performances suivantes :

- ✓ **La précision moyenne** (AP : Average Precision), qui se calcul comme suit

$$AP = \sum_{R \in RS} \frac{1}{|Rel_R|} \sum_{r=1}^{|L_R|} isrel(r) \frac{count(r)}{r}, \quad \text{Avec :}$$

R : Une requête ;

RS : Ensemble de requêtes ;

$|Rel_R|$: Nombre de services relevant pour la requête R ;

L_R : La liste ordonnée des services découverts par la requête R ;

$$isrel(r) = \begin{cases} 1 & \text{si le service est dans } L_R \text{ au Rang } r \text{ est relevant} \\ 0 & \text{else} \end{cases}$$

$$count(r) = \sum_1^r isrel(i)$$

- ✓ Un deuxième indicateur important est le temps de réponse moyen de requête (AQRT : Average Query Response Time).

Nous avons comparé nos résultats avec des variantes de (SAWSDL-MX1) qui sont :

- SAWSDL-M0: Logic Based Matcher,
- SAWSDL-M1: Hybrid Matcher using Loss of Information Similarity,
- SAWSDL-M2: Hybrid Matcher using Extended Jaccard Similarity,

- SAWSDL-M3: Hybrid Matcher using Cosine Similarity,
- SAWSDL-M4: Hybrid Matcher using Jensen Shannon Similarity.

La table suivante produit une comparaison entre plusieurs matchmakers selon des critères de performance :

Table 6.2: Résultats de comparaison des matchmakers.

Matchmaker	1 Requête		2 Requêtes		4 Requêtes	
	20 Services		20 Services		20 Services	
	AP	AQRT	AP	AQRT	AP	AQRT
SAWSDL-M0	0.500	2255	0.667	3698	0.75	1873
SAWSDL-M1	0.883	1502	0.922	3068	0.941	1866
SAWSDL-M2	0.840	1512	0.894	3604	0.942	1892
SAWSDL-M3	0.849	1770	0.899	3814	0.925	1632
SAWSDL-M4	0.840	1925	0.894	3416	0.920	1542
SAWSDL-MOM	0.905	5551	0.937	10202	0.952	12088

Notre matchmaker présenté complémente et poursuit les travaux déjà existant en combinant des techniques acceptées et de nouvelles idées conduisant à de bons résultats en ce qui concerne les mesures IR, tel que le rappel et la précision. Notre algorithme fournit un bon AP, mieux que toutes les variantes du matchmaker de SAWSDL-MX1.

6.5 Discussion

À la lumière des résultats obtenus et nous rappelons que l'objectif de nos recherches était le matchmaking des services web sémantiques dans un environnement multi-ontologie. Compte tenu de notre problématique et nos objectifs, l'aspect interopérabilité sémantique des ontologies est fondamental. L'exploration des travaux de matchmaking, nous a permis de dégager notre proposition pour le matchmaking des SWS. Pour valider nos travaux, nous avons expérimenté les composantes principales, la mesure de similarité, l'alignement partiel, et le matchmaking des services web sémantiques SAWSDL, Nous pouvons avancer que :

- Le but de l'évaluation de la mesure de similarité est d'examiner sa corrélation avec le jugement humain (expert). Il est très difficile d'aboutir à une stratégie efficace de test. La

mesure de similarité retourne une valeur numérique. Pour comparer le résultat avec un jugement humain (expert), l'idée était de transformer le jugement en une valeur numérique (NER : Numeric Expected Results (Résultats numériques attendus)) et de procéder à une comparaison. Dans le premier test, le concept Book de l'ontologie 101 est considéré ; un autre concept Book de l'ontologie 103 est pris pour la comparaison. Pour l'humain les deux concepts sont similaires, aussi la mesure de similarité retourne un 1, ce qui indique une corrélation. Le test 7 de Report et Delivrable retourne une faible proximité sémantique. On remarque que les résultats ont un comportement similaire à celui du résultat escompté (on peut dire qu'il y a une corrélation avec le jugement humain). Les résultats expérimentaux confirment la viabilité de la mesure de similarité sémantique proposée dans le présent travail.

- L'algorithme d'alignement partiel tente de trouver les correspondances pour les concepts apparaissant dans la requête, avec ceux du service offert. L'idée des expérimentations est de simuler cette situation sur plan ontologies. Nous avons exécuté 5 cas de tests; chaque cas est un alignement d'un ensemble de concepts d'une première ontologie avec un autre ensemble de concepts d'une deuxième ontologie. Pour réaliser nos test, nous avons choisi l'ontologie 101 comme première ontologie et à chaque fois l'une des ontologies (101,103, 104, 202, 230, 301 et 304) comme deuxième ontologie. Pour mesurer la précision d'alignement partiel entre deux ontologies, la moyenne des performances des tests est considérée. Les meilleurs résultats sont obtenus lorsque les structures des ontologies sont similaires (ou identiques), à savoir les familles de tests 10x et 30x. Ainsi, notre méthode obtient des valeurs de précision pour ces tests qui sont égaux à 1,00. Cela est dû au fait que notre approche explore les structures des entités pour aligner par la similarité structurelle.
- SAWSDL-MOM inclut les paramètres de message SAWSDL et il est possible de l'étendre à d'autres niveaux. Pour évaluer notre algorithme, nous avons adopté une procédure conventionnelle, introduite dans le concours S3, en effet la mesure d'une précision moyenne AP et le temps de réponse moyen AQR. Notre algorithme fournit un bon AP, mieux que toutes les variantes du matchmaker de SAWSDL-MX1. Il est évident que notre approche est un peu plus coûteuse en temps, vu la complexité de la procédure mise en œuvre pour l'appariement individuel et le type d'expérimentation réalisé (1, 2, 4 requête(s)) pour le même nombre de services (20 services). Cependant, dans un scénario réel, une seule requête sera satisfaite à la fois. Les résultats expérimentaux montrent que, l'approche peut faire un compromis entre la précision moyenne et le temps de réponse moyen.

6.6 Conclusion

Dans le présent chapitre, nous avons présenté les expérimentations de nos travaux de recherche. En premier lieu, nous avons décrit les principes et les indicateurs de notre évaluation. Nous avons utilisé la base de test OAEI pour évaluer les aspects interopérabilité sémantique entre ontologies. En effet, nous avons procédé à des tests de la mesure de similarité. L'algorithme d'alignement partiel a été évalué en utilisant la même collection de test. Pour réaliser l'évaluation de l'algorithme de matchmaking, nous avons utilisé la collection de test SAWSDL-TC, tout en respectant les indicateurs SME2. Enfin, nous avons présenté une discussion des résultats des évaluations des performances de notre approche.

Conclusion & Perspectives

Conclusion & Perspectives

Conclusion

Vu le nombre important de services Web disponibles aujourd'hui, la découverte automatique est reconnue comme une tâche cruciale. Pour promouvoir l'automatisation de la découverte des services sémantiques, différents langages ont été créés permettant de décrire les fonctionnalités de services dans une forme interprétable par la machine en utilisant les technologies du Web sémantique. Les travaux de recherche menés sur la description de services Web utilisent de plus en plus les ontologies pour fournir une représentation de l'information sémantique. L'appariement (matchmaking) sémantique est une phase cruciale dans le processus de découverte de services Web sémantiques.

Pour réussir une découverte sémantique, il faut proposer un algorithme d'appariement efficace entre une requête de service et les descriptions de services. Il se déroule sur deux étapes : 1. Calcul des valeurs d'appariements élémentaires entre éléments sémantiques des descriptions de services, 2. Agrégation des valeurs d'appariements élémentaires et calcul d'un degré global d'appariement.

Plusieurs travaux ont discuté la découverte sémantique, lorsque les deux services, requis et offert, utilisent une ontologie commune. On pense qu'une approche basée sur l'usage d'une ontologie unique est très peu pratique. Rares, sont les approches qui considèrent le cas, où les deux services, requis et offert appartiennent à différentes ontologies du même domaine. Pour rendre un système de découverte plus réaliste et améliorer sa scalabilité, il est important de prendre en compte l'aspect multi ontologie.

Notre travail s'inscrit dans le cadre le plus général de l'interopérabilité sémantique, qui est une exigence pour augmenter le degré d'automatisation et la scalabilité. En particulier, nos travaux de recherche s'intéressent aux difficultés inhérentes à l'appariement de deux services web sémantiques requis et offert, dans un contexte multi-ontologie. En l'occurrence, l'usage de l'alignement des ontologies pour surmonter l'hétérogénéité sémantique, qui reste une gageure informatique.

Dans la première contribution, nous avons proposé un cadre général de matchmaking des services web sémantiques dans un contexte multi-ontologie, dont le noyau est une nouvelle mesure de similarité sémantique entre concepts de deux ontologies. L'algorithme de mesure similarité des concepts entre ontologies utilise une combinaison de plusieurs similarités de bases qui considèrent différents aspects des concepts tels que la structure des deux ontologies source et cible.

Dans la seconde proposition, nous avons développé un algorithme d'appariement (matchmaking) des services web SAWSDL nommé SAWSDL-MOM, qui est une instance de notre cadre général et intégrait un algorithme d'alignement partiel des ontologies. Ce dernier utilise une version améliorée de notre mesure de similarité. Les résultats obtenus en expérimentations sont satisfaisants et encouragent les futurs travaux dans cette piste de recherche.

Conclusion & Perspectives

Perspectives

Plusieurs améliorations de notre travail sont possibles, à savoir:

- Amélioration de la mesure de similarité, par exemple :
 - La considération de la fermeture transitive pour le calcul du voisinage,
 - L’usage de ressources autres que WordNet
 - L’usage de l’algorithme des courts chemins de Dijkstra au lieu de l’algorithme Hongrois.
- Appliquer notre algorithme d’alignement partiel à d’autres contextes comme l’informatique pervasive.
- Ajout des optimisations à l’algorithme pour le rendre plus adapté aux contraintes temps réel,
- Exploration d’autres méthodes d’agrégation des résultats élémentaires,
- L’utilisation d’un estimateur pour déterminer automatiquement le poids des différents niveaux d’abstraction des services comme l’estimateur des moindres carrés ordinaires (OLS : Ordinary Least Squares estimator) par exemple. Exploration de méthodes d’agrégation des résultats élémentaires,
- Extension de l’approche pour couvrir l’aspect comportemental des services,
- Appliquer nos idées à d’autres modèles de représentations des services web comme OWL-S et WSMO,
- Concevoir une collection de tests qui supporte les services web annotés avec plusieurs ontologies. Par le développement d’outil d’annotation (Automatique !) sémantique des services web SAWSDL,
- Rendre le Matchmaker générique, par l’adoption d’un langage de description pivot pour les services web sémantiques et la résolution des problèmes de transformation d’un langage de description vers un autre.

Finalement, il faut souligner la double gageure de ce travail de thèse. D’une part, les services web sémantiques et d’autres part l’interopérabilité des ontologies. Nous pensons que même avec l’émergence de nouvelles thématiques de recherche comme le Cloud Computing, il reste encore des pistes à explorer dans le contexte des services web sémantiques, comme:

- La description sémantique des services web,
- Le processus de matchmaking,
- Les collections de tests qui sont synthétiques et ne reflètent pas la réalité,

Conclusion & Perspectives

- Les plates-formes d'évaluation des matchmakers,
- Et bien, d'autres.

En outre, la compréhension du sens est une aptitude humaine que les machines ont du mal à imiter. Malgré les efforts déployés par les chercheurs, l'interopérabilité des ontologies est toujours une problématique d'actualité avec des challenges toujours aussi complexes.

Bibliographie

- Abdelrahman, A. M., & Kayed, A. (2015). A survey on semantic similarity measures between concepts in health domain. *American Journal of Computational Mathematics*, 5(02), 204.
- Abolhassani H. , B.B. Hariri, S. H. Haeri. (2006).On Ontology Alignment Experiments, *Webology*, Volume 3, Number 3.
- Al-Mubaid, H., & Nguyen, H. A. (2009). Measuring semantic similarity between biomedical concepts within multiple ontologies. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 39(4), 389-398.
- Altova (2009) ‘SemanticWorks Semantic Web tool - Visual RDF and OWL editor’, http://www.altova.com/products/semanticworks/semantic_web_rdf_owl_editor.html, Website accessed Sep. 10, 2016.
- Akkiraju, R., Farrell, J., Miller, J. A., Nagarajan, M., Sheth, A., & Verma, K. (2005). Web service semantics-wsdl-s.
- Anderson, S., Bohren, J., Boubez, T., Chanliau, M., Della-Libera, G., Dixon, B., ... & Zolfonoon, R. (2005). Web services trust language (ws-trust). *Public draft release, Actional Corporation, BEA Systems, Computer Associates International, International Business Machines Corporation, Layer*, 7, 68.
- Andrews, T., Curbera, F., Dholakia, H., Golland, Y., Klein, J., Leymann, F., ... & Trickovic, I. (2003). Business process execution language for web services.
- Andrieux, A., Czajkowski, K., Dan, A., Keahey, K., Ludwig, H., Nakata, T., ... & Xu, M. (2004, August). Web services agreement specification (WS-Agreement). In *Global Grid Forum* (Vol. 2).
- Baltá, A., & Fernández, A.(2010). Web Service Description Alignment. *Proceeding of the thirth international Symposiom on Web Services*, 212-218.
- Bellwood, T., Clément, L., Ehnebuske, D., Hately, A., Hondo, M., Husband, Y. L., ... & von Riegen, C. (2002). UDDI Version 3.0. Published specification, *Oasis*, 5, 16-18.
- Berners-Lee,T., J.Hendler, and O.Lassila. (2002) .The Semantic Web. A new form of Web content that is meaningful to computers willuleash a revolution of new possibilities. *Scientific American*, 284(5): 34–43.

Bibliographie

- Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M., Ferris, C., & Orchard, D. (2004). Web services architecture, w3c working group note, 2004. URL: <http://www.w3.org/TR/2004/NOTE-ws-arch-20040211>.
- Borst, W, N. (1997). Construction of Engineering Ontologies. PhD thesis, Institute for Telematica and Information Technology, University of Twente, Enschede, The Netherlands
- Box, D., Ehnebuske, D., Kakivaya, G., Layman, A., Mendelsohn, N., Nielsen, H. F., ... & Winer, D. (2000). Simple object access protocol (SOAP) 1.1.
- Brickley, D., & Guha, R. V. (2000). Resource Description Framework (RDF) Schema Specification 1.0: W3C Candidate Recommendation 27 March 2000.
- Brown, A., Johnston, S., & Kelly, K. (2002). Using service-oriented architecture and component-based development to build web service applications. Rational Software Corporation, 6.
- Cabrera, F., Copeland, G., Cox, B., Freund, T., Klein, J., Storey, T., & Thatte, S. (2002). Web services transaction (WS-transaction). *Microsoft, IBM etc.*
- Cardoso, J., & Sheth, A. (2003). Semantic e-workflow composition. *Journal of Intelligent Information Systems*, 21(3), 191-225.
- Cardoso, J. (2006, September). Discovering semantic web services with and without a common ontology commitment. In *Services Computing Workshops, 2006. SCW'06. IEEE* (pp. 183-190). IEEE.
- Cardoso, J., Miller, J. A., & Emani, S. (2008). Web services discovery utilizing semantically annotated WSDL. In *Reasoning Web* (pp. 240-268). Springer Berlin Heidelberg.
- Cerami, E. (2002). *Web services essentials: distributed applications with XML-RPC, SOAP, UDDI & WSDL*. O'Reilly Media, Inc..
- Chabeb, Y. (2011). Contributions à la description et la découverte de services web sémantiques (Doctoral dissertation, Institut National des Télécommunications).
- Cherifi, C. (2011). Classification et Composition de Services Web: Une Perspective Réseaux Complexes (Doctoral dissertation, Université Pascal Paoli).
- Chinnici, R., Moreau, J. J., Ryman, A., & Weerawarana, S. (2006). Web services description language (wsdl) version 2.0 part 1: Core language. *W3C Recommendation*, 26.
- Christensen, E., Curbera, F., Meredith, G., & Weerawarana, S. (2001). Web services description language (WSDL) 1.1.

Bibliographie

- Chua, W.W.K., Kim, J.J. (2010). Eff2Match Results for OAEI 2010. In Proceeding of 5th International Workshop on Ontology Matching (OM 2010), Shanghai, China, 7–11 November 2010.
- COMPOSER, T. (2011). TopBraid Composer getting started guide version 3.0, created by TopQuadrant. US. Retrieved from <http://www.topquadrant.com/docs/marcom/TBC-Getting-Started-Guide.pdf>
- Corcho, Ó., Fernández-López, M., Gómez-Pérez, A., & Vicente, Ó. (2002, October). WebODE: An integrated workbench for ontology representation, reasoning, and exchange. In International Conference on Knowledge Engineering and Knowledge Management (pp. 138-153). Springer Berlin Heidelberg.
- Cruz, I. F., Antonelli, F. P., & Stroe, C. (2009). AgreementMaker: efficient matching for large real-world schemas and ontologies. *Proceedings of the VLDB Endowment*, 2(2), 1586-1589.
- De Bruijn, J., Fensel, D., Kerrigan, M., Keller, U., Lausen, H., & Scicluna, J. (2006). The Web Service Modeling Ontology. *Modeling Semantic Web Services: The Web Service Modeling Language*, 23-28.
- Dietze, S., Benn, N., Domingue, J., Conconi, A., & Cattaneo, F. (2009, December). Two-fold service matchmaking—applying ontology mapping for semantic web service discovery. In *Asian Semantic Web Conference* (pp. 246-260). Springer Berlin Heidelberg.
- Dumas, M., & Fauvet, M. C. (2008). *Intergiciels et Construction d'Applications Réparties (Les Services Web)*. Queensland University of Technology, Brisbane (Australie), Université Joseph Fourier, Grenoble. <http://deptinfo.unice.fr/~baude/WS/CoursICAR-WS.pdf>.
- Ehrig, M., & Sure, Y. (2004, May). Ontology mapping—an integrated approach. In *European Semantic Web Symposium* (pp. 76-91). Springer Berlin Heidelberg.
- Ehrig, M., & Staab, S. (2004, August). Efficiency of ontology mapping approaches. In *International Workshop on Semantic Intelligent Middleware for the Web and the Grid at ECAI (Vol. 4, pp. 47-61)*.
- Euzenat J., et al..(2004). D2.2.3: State of the art on ontology alignment. Technical report, NoE Knowledge Web project deliverable, 2004. <http://knowledgeweb.semanticweb.org/>.
- Euzenat, J. & Shvaiko, P. (2007). *Ontology Matching*. Springer, Heidelberg.
- Farrell, J., & Lausen, H. (2007). Semantic annotations for WSDL and XML schema. W3C recommendation.

Bibliographie

- Fellah, A., Malki, M., & Zahaf, A. (2008). Alignement des ontologies: utilisation de WordNet et une nouvelle mesure structurelle. In *Conférence en Recherche d'Information et Applications, CORIA*, 2008. pp.401-408.
- Fellah, A., Malki, M., & Elci, A. (2016a). A Similarity Measure across Ontologies for Web Services Discovery. *International Journal of Information Technology and Web Engineering (IJITWE)*, 11(1), 22-43.
- Fellah, A., Malki, M., & Elçi, A. (2016b). Web Services Matchmaking Based on a Partial Ontology Alignment. *International Journal of Information Technology and Computer Science (IJITCS)*, 8(6), 9.
- Fensel, D., & Bussler, C. (2002). The web service modeling framework WSMF. *Electronic Commerce Research and Applications*, 1(2), 113-137.
- Gao, S., Sperberg-McQueen, C. M., Thompson, H. S., Mendelsohn, N., Beech, D., & Maloney, M. (2009). W3C XML schema definition language (XSD) 1.1 part 1: Structures. *W3C Candidate Recommendation*, 30(7.2).
- Genesereth, M. R. (1995). Knowledge interchange format specification. Available via World Wide Web URL <http://logic.stanford.edu/kif/kif.html>.
- Giunchiglia, F. & Shvaiko, P. (2003). Semantic matching. *The Knowledge Engineering Review* 18(3), 265–280.
- Giunchiglia, F., Yatskevich, M., & Shvaiko, P. (2007). Semantic matching: Algorithms and implementation. In *Journal on Data Semantics IX* (pp. 1-38). Springer Berlin Heidelberg.
- Gómez-Pérez, A., González-Cabero, R., & Lama, M. (2004). Ode sws: A framework for designing and composing semantic web services. *IEEE Intelligent Systems*, 19(4), 24-31.
- Gruber, T. (1993). A translation approach to portable ontology specifications, *Knowledge Acquisition* 5(2), pages 199-220.
- GRUBER, T., & OLSEN, G. (1994). An ontology for engineering mathematics, *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*,. Morgan-Kauffmann.
- HENDLER, J., & MCGUINNESS, D. (2001). The Darpa Agent Markup Language. http://www-ksl.stanford.edu/pub/KSL_Reports/KSL-00-10.html.

Bibliographie

- Horrocks, I. (2002). DAML+OIL: A Description Logic for the Semantic Web. *IEEE Data Eng. Bull.*, 25(1), 4-9.
- Jean-Mary, Y. R., Shironoshita, E. P., & Kabuka, M. R. (2009). Ontology matching with semantic verification. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3), 235-251. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2825706/>
- Jouanot, F. (2000). Un modèle sémantique pour l'interopérabilité de systèmes d'information. XVIIIe congrès INFORSID, 347-364.
- Kalfoglou, Y., & Schorlemmer, M. (2003). Ontology mapping: the state of the art. *The knowledge engineering review*, 18(01), 1-31.
- Kalyanpur, A., Sirin, E., Parsia, B., & Hendler, J. (2004). Hypermedia inspired ontology engineering environment: Swoop. In *Proceedings of 3rd International Semantic Web Conference (ISWC-2004)*, Japan (Poster).
- Kavantzas, N. (2004). *Web Services Choreography Description Language (WS-CDF) Version 1.0*, Retrieved from <http://www.w3.org/TR/ws-cdl-10/>.
- Keskes, N. (2014). Context of QoS in Web Service Selection: The Generalization of Web Services Selection Process. *Int. Arab J. e-Technol.*,3(4).
- Khater, M & Malki, M.(2014). "Improving the Performance of Semantic Web Services Discovery: Shortest Path based Approach", *IJITCS*, vol.6, no.7, pp.32-39, 2014. DOI: 10.5815/ijitcs.2014.07.05
- Kifer, M., Lausen, G., & Wu, J. (1995). Logical foundations of object-oriented and frame-based languages. *Journal of the ACM*, 42(4):741-843, 1995.
- Kiefer, C. (2009). *Non-Deductive Reasoning for the Semantic Web and Software Analysis* (Doctoral dissertation, University of Zurich).
- Klusch, M. (2008). *Semantic web service coordination*. In *CASCOM: Intelligent Service Coordination in the Semantic Web* (pp. 59-104). Birkhäuser Basel.
- Klusch, M., & Kaufer, F. (2009). WSMO-MX: A hybrid Semantic Web service matchmaker. *Web Intelligence and Agent Systems: An International Journal*, 7(1), 23-42.
- Klusch, M., & Kapahnke, P. (2009, October). OWLS-MX3: an adaptive hybrid semantic service matchmaker for OWL-S. In *Proceedings of 3rd International Workshop on Semantic Matchmaking and Resource Retrieval (SMR2)*, USA.

Bibliographie

- Klusch, M., Kapahnke, P., & Zinnikus, I. (2009, May). Hybrid adaptive web service selection with SAWSDL-MX and WSDL-analyzer. In European Semantic Web Conference (pp. 550-564). Springer Berlin Heidelberg.
- Klusch, M. (2012). Overview of the S3 Contest: Performance evaluation of semantic service matchmakers. In Semantic web services (pp. 17-34). Springer Berlin Heidelberg.
- Klusch, M. (2014). Service discovery. In Encyclopedia of Social Network Analysis and Mining (pp. 1707-1717). Springer New York.
- Klyne, G., Carroll, J. J., & McBride, B. (2014). Resource description framework (RDF): Concepts and abstract syntax. *W3C recommendation*.
- Kritikos, K. E. (2008). Qos-based web service description and discovery (Doctoral dissertation, University of Crete).
- Kuhn, H. W. (2010). The hungarian method for the assignment problem. In 50 Years of Integer Programming 1958-2008 (pp. 29-47). Springer Berlin Heidelberg.
- Le DuyNgan, T. M. H., & Goh, A. E. S. (2006, September). MOD-A Multi-Ontology Discovery System. In 1st International Workshop on Semantic Matchmaking and Resource Retrieval (p. 3).
- Leacock, C., & Chodorow, M. (1998). Combining local context and WordNet similarity for word sense identification. *WordNet: An electronic lexical database*, 49(2), 265-283.
- Levenshtein, V. I. (1966, February). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady* (Vol. 10, No. 8, pp. 707-710).
- Li, L., & Horrocks, I. (2004). A software framework for matchmaking based on semantic web technology. *International Journal of Electronic Commerce*, 8(4), 39-60.
- Lin, D. (1998, July). An information-theoretic definition of similarity. In *ICML* (Vol. 98, pp. 296-304).
- Martin, D., Burstein, M., Hobbs, J., Lassila, O., McDermott, D., McIlraith, S., ... & Sycara, K. (2004). OWL-S: Semantic markup for web services. *W3C member submission*, 22, 2007-04. <http://www.ai.sri.com/~daml/services/owl-s/1.2/overview/>
- McGuinness, D. L., & Van Harmelen, F. (2004). OWL Web Ontology Language Overview. *W3C recommendation*, 10(2004-03), 10. <https://www.w3.org/TR/owl-features/>
- Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11), 39-41.

Bibliographie

- Mohebbi, K., Ibrahim, S., Khezrian, M., Munusamy, K., & Tabatabaei, S. G. H. (2010). A comparative evaluation of semantic web service discovery approaches. In Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services (pp. 33-39). ACM.
- Mrissa, M. (2007). *Médiation Sémantique Orientée contexte pour la composition de services Web* (Doctoral dissertation, Université Lyon 1).
- Nadalin, A., AmberPoint, G. T., BEA, P. D., BEA, H. L., CommerceOne, S. C., ContentGuard, T. D., ... & Hitachi, Y. K. (2002). *Web Services Security*.
- Niepert, M., Buckner, C., & Allen, C. (2008, May). Answer Set Programming on Expert Feedback to Populate and Extend Dynamic Ontologies. In FLAIRS Conference (pp. 500-505).
- NOY, N., FERGERSON, R. W., & MUSEN, M. A. (2000). The knowledge model of Protégé2000 : combining interoperability and flexibility, in Proceedings of the International Conference on Knowledge Engineering and Knowledge Management (EKAW'00).
- Noy, N. F., & Musen, M. A. (2000). Prompt: algorithm and tool for automated ontology merging and alignment. In Proceeding of Seventeenth National Conference on Artificial Intelligence AAAI.
- Noy, N. F. & Musen, M. A. (2003). The prompt suite: interactive tools for ontology merging and mapping. *Int. J. Hum.-Comput. Stud.* 59(6), 983–1024.
- OAEI: Ontology Alignment Evaluation Initiative Benchmark Test Library 2014. Retrieved July 7, 2015 from <http://oaei.ontologymatching.org/2014/benchmarks>.
- ONTOEDIT. (2004). Ontology Editor Home Page, <http://www.ontoprise.de/com/>.
- Oundhakar, S., Verma, K., Sivashanmugam, K., Sheth, A. P., & Miller, J. A. (2005). Discovery of web services in a multi-ontology and federated registry environment. *International Journal of Web Services Research*, 2(3), 1-32.
- Paolucci, M., Kawamura, T., Payne, T. R., & Sycara, K. (2002). Semantic matching of web services capabilities. In *The Semantic Web—ISWC 2002* (pp. 333-347). Springer Berlin Heidelberg.
- Plebani, P., & Pernici, B. (2009). URBE: Web service retrieval based on similarity evaluation. *Knowledge and Data Engineering, IEEE Transactions on*, 21(11), 1629-1642.
- PRO. (2002). PROTEGE2000, Protege2000 Ontology Editor Home Page, <http://protege.stanford.edu/>.

Bibliographie

- Quix, C., Gal, A., Sagi, T., Kensch, D. (2010). An Integrated Matching System: GeRoMeSuite and SMB-Results for OAEI 2010. In *Proceeding of 5th International Workshop on Ontology Matching (OM 2010)*, Shanghai, China, 7–11 November 2010.
- Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *IEEE transactions on systems, man, and cybernetics*, 19(1), 17-30.
- Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. *arXiv preprint cmp-lg/9511007*.
- Rodriguez, M. A., & Egenhofer, M. J. (2003). Determining semantic similarity among entity classes from different ontologies. *Knowledge and Data Engineering, IEEE Transactions on*, 15(2), 442-456.
- SAWSDL-TC: SAWSDL Service Retrieval Test Collection. Latest version SAWSDL-TC3.0, Retrieved from: <http://projects.semwebcentral.org/projects/sawSDL-tc/>.
- Schulte, S. (2010). "Web Service Discovery Based on Semantic Information-Query Formulation and Adaptive Matchmaking." PhD diss., TU Darmstadt.
- Sivashanmugam, K., Verma, K., Sheth, A., & Miller, J. A. (2003). Adding Semantics to Web Services Standards. In *International Conference on Web Services (ICWS2003)*, pages 395–401. CSREA Press.
- Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: principles and methods. *Data & knowledge engineering*, 25(1-2), 161-197.
- Sycara, K., Klusch, M., Widoff, S., & Lu, J. (1999). Dynamic service matchmaking among agents in open information environments.
- Sycara, K., Widoff, S., Klusch, M., & Lu, J. (2002). Larks: Dynamic matchmaking among heterogeneous software agents in cyberspace. *Autonomous agents and multi-agent systems*, 5(2), 173-203.
- Syeda-Mahmood, T., Shah, G., Akkiraju, R., Ivan, A. A., & Goodwin, R. (2005, July). Searching service repositories by combining semantic and ontological matching. In *Web Services, 2005. ICWS 2005. Proceedings. 2005 IEEE International Conference on* (pp. 13-20). IEEE.
- SyncRO Soft, S. R. L. (2015). Oxygen XML Editor Version 17.1. Craiova: Syncro Soft srl. Retrieved from <https://www.oxygenxml.com/whatisnew17.1.html>
- Tversky, A., & Shafir, E. (2004). *Preference, belief, and similarity: selected writings*. MIT Press.

Bibliographie

- Usanavasin, S., Takada, S., & Doi, N. (2005). Semantic web services discovery in multi-ontology environment. In *On the Move to Meaningful Internet Systems 2005: OTM 2005 Workshops* (pp. 59-68). Springer Berlin Heidelberg.
- USCHOLD, M., & KING, M. (1995). Towards a methodology for building ontologies. (pp. 15-30). Edinburgh: Artificial Intelligence Applications Institute, University of Edinburgh.
- Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The knowledge engineering review*, 11(02), 93-136.
- Valtchev, P. Euzenat, J., Loup, D. & Touzani, M. (2004). Ontology alignment with OLA. In *Proc. 3rd ISWC2004 workshop on Evaluation of Ontology- based tools (EON)* (pp. 59-68).
- Vedamuthu, A. S., Orchard, D., Hirsch, F., Hondo, M., Yendluri, P., Boubez, T., & Yalçinalp, U. (2007). Web services policy 1.5-framework. *W3C Recommendation*, 4, 1-41.
- Verma, K., Sheth, A., Oundhakar, S., Sivashanmugam, K., & Miller, J. (2007). Allowing the use of Multiple Ontologies for Discovery of Web Services in Federated Registry Environment. Department of Computer Science, University of Georgia, Athens, Georgia. Technical Report# UGA-CS-LSDIS-TR-07-011, 1-27.
- Weise, T., Blake, M. B., & Bleul, S. (2014). Semantic Web Service Composition: The Web Service Challenge Perspective. In *Web Services Foundations* (pp. 161-187). Springer New York.
- Wood, D. (1995). Standard generalized markup language: Mathematical and philosophical issues. In *Computer Science Today* (pp. 344-365). Springer Berlin Heidelberg.
- Wu Z. & Palmer M. (1994). Verb semantics and lexical selection. in *proc. of the 32nd Annual Meeting of Computational Linguistics*, Las Cruces, 1994, pp. 133-138.
- Xu, P., Wang, Y., Cheng, L., & Zang, T. (2010, November). Alignment Results of SOBOM for OAEI 2010. In *Proceedings of the 5th International Conference on Ontology Matching-Volume 689* (pp. 203-211). CEUR-WS. Org.
- Zinnikus, I., Rupp, H. J., & Fischer, K. (2006). Detecting similarities between web service interfaces: The wsdl analyzer. In *Interoperability for Enterprise Software and Applications: Proceedings of the Workshops and the Doctorial Symposium of the Second IFAC/IFIP I-ESA International Conference: EI2N, WSI, IS-TSPQ 2006* (pp. 145-156). ISTE.