

وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR-ANNABA



جامعة باجي مختار - عنابة

Faculté des Sciences de L'Ingéniorat Année 2015-2016

Département d'Informatique

THESE

Présentée en vue de l'obtention
du diplôme de Doctorat en Sciences

***Approche flexible pour l'intégration des
ontologies hétérogènes et évolutives***

Spécialité : Informatique
Option : Intelligence Artificielle

Préparée par

M^{me} Klai-Soukehal Sihem

Jury :

Président	M ^r Farah Nadir	Pr.	Université Badji Mokhtar-Annaba
Directeur de Thèse :	M ^r Khadir Med Tarek	Pr.	Université Badji Mokhtar-Annaba
Invité	M ^r Zimmermann Antoine	Dr.	Ecole Supérieure des Mines Saint-Etienne, France
Examinatrice	M ^{me} Boufaida Zizette	Pr.	Université Mentouri Constantine 2
Examinatrice	M ^{me} Seridi Hassina	Pr.	Université Badji Mokhtar-Annaba
Examineur	M ^r Zaidi Sahnoun	Pr.	Université Mentouri Constantine 2
Examineur	M ^r Zaarour Naceredinne	Pr.	Université Mentouri Constantine 2

UNIVERSITÉ BADJI MOKHTAR ANNABA

LABGED

THÈSE

présentée en première version en vue d'obtenir le grade de Docteur,
spécialité « Intelligence Artificielle »

par

Klai Sihem

APPROCHES FLEXIBLES POUR L'INTÉGRATION DES ONTOLOGIES HÉTÉROGÈNES ET ÉVOLUTIVES

Thèse soutenue le la date de soutenance devant le jury composé de :

M. KHADIR MED TREK Université d'Annaba (Encadreur)

À mes défunts parents et beaux parents

À mon mari

À mes sœurs, mes enfants Souha et Haithem

Et à mes amies.

REMERCIEMENTS

JE tiens à remercier mon encadreur, Pr. Khadir Med Tarek, de m'avoir encadré, encouragé et soutenu durant toutes ces années.

Mes vifs remerciements vont aussi à mon co-encadreur, Dr. Zimmermann Antoine, de m'avoir aidé. Ma thèse s'est concrétisée grâce à son expérience dans le domaine et ses orientations.

Mon profond respect et gratitude aux membres de jury qui ont accepté d'examiner et d'évaluer ce travail.

Mes remerciements vont aussi au Pr. Boissier Olivier de m'avoir accueilli dans son laboratoire, sans oublier toute personne qui a contribué à m'apporter de l'aide pendant mes séjours à Saint-Etienne.

Je ne sais comment remercier ma collègue et ma très chère amie, Mecheri Karima, pour tout ce qu'elle a fait pour me soutenir et m'aider à surmonter toutes les difficultés que j'ai rencontré durant ces années.

Je remercie aussi tous mes collègues du Département Informatique pour leur gentillesse et leurs encouragements.

Je suis reconnaissante face à l'aide et le soutien absolu et concret de mon mari Med Cherif et ma sœur Chadia qui ont été patients et disponibles pendant cette période délicate. Je tiens à les remercier du fond du cœur pour tous les efforts qu'ils ont fourni.

Annaba, le 5 novembre 2016.

Titre Approches flexibles pour l'intégration des ontologies hétérogènes et évolutives

Résumé L'intégration des ontologies pré-existantes qui sont hétérogènes, complémentaires mais alignées et le raisonnement sur l'ensemble dans un environnement contextuel reste un défi majeur dans le domaine de la recherche. Dans cette thèse, nous considérons un réseau d'ontologies alignées par deux types de correspondances à savoir les *mappings* et les *links*. Les *mappings* représentent un ensemble de relations prédéfinies tel que la subsomption, l'équivalence, la disjonction, etc. Leur sémantique est fixe dans toutes les interprétations. Tandis que les *links* servent à lier des ontologies complémentaires. Ce type de correspondances est défini par des experts et leur sémantique dépend du contexte utilisé. On suppose que ces alignements par paire d'ontologies peuvent être produits indépendamment les uns des autres dans des contextes différents et par conséquent, ils sont sémantiquement hétérogènes. Il peut être nécessaire d'aligner les terminologies d'alignements différents et donc de former des alignements de niveau supérieur. Ceci engendre un réseau d'ontologies qui à son tour peut être aligné avec d'autres réseaux d'ontologies. On peut alors parler de réseau de connaissances multi-niveaux, un concept que nous formalisons et pour lequel nous proposons deux sémantiques possibles permettant l'automatisation du raisonnement. La première approche consiste à interpréter le réseau en une ontologie unique en logique de description tout en distinguant entre les entités provenant de sources de connaissances différentes et de niveaux différents. Cette approche permet de réduire le raisonnement sur le réseau de connaissances multi-niveaux à un raisonnement en logique de description. Elle a été implémentée et testée afin de comparer les résultats pour différentes typologies du réseau. La seconde approche consiste à interpréter le réseau de connaissances multi-niveaux comme un système distribué en DDL. Ainsi le raisonnement sur le réseau de connaissances multi-niveaux se

réduit aux mécanismes d'inférence en DDL.

Mots-clés Réseau de connaissances, ontologies, alignement d'ontologies, raisonnement en DL, raisonnement en DDL.

Title Flexible approaches for integrating heterogeneous and scalable ontologies

Abstract Integration of pre-existing heterogeneous, complementary aligned ontologies, as well as reasoning on the entire system within a contextual environment, remains a major challenge. In this thesis an ontology network, aligned using two types of correspondence i.e., *mappings* and *links* correspondences are presented. *Mappings* represent a set of predefined relations such as subsumption, equivalence, disjunction, etc. with fixed semantics through all possible interpretations. While *links* are used to link complementary ontologies, with a type of correspondences defined by domain experts with context dependant semantics. Supposedly, the pair alignments are performed independently from each other in different contexts and consequently are semantically heterogeneous. It may be necessary to align different alignment terminologies, and therefore constituting higher level alignments, obtaining an ontology network, which can be in turn, aligned with others. We then can talk about multi level knowledge networks, a concept formalized here and for which two possible semantic representations permitting reasoning automation are presented. The first approach allows the interpretation of the network as a unique meta ontology using Description Logic, with a clear distinction between entities coming from different knowledge sources as well as different levels. The approach goal is to bring reasoning on an multi-level ontology network to a reasoning based on DL. This approach is implemented and tested in order to compare results, for different networks. The second approach focuses on the interpretation of the multi level ontology network as a DLL distributed system, bringing reasoning to inference mechanisms in DDL.

Keywords networked knowledge, ontologies, ontology alignments, DL-reasoning, DDL-reasoning

المخلص

دمج مجموعة من الانطولوجيات الموجودة مسبقا، الغير متجانسة، المكملة لبعضها، وجود محاذاة فيما بينها و الاستدلال المنطقي على المجموعة لا يزال يشكل تحديا كبيرا في المجال العلمي.

في هذه الأطروحة، نفترض وجود مجموعة من الانطولوجيات المتواصلة عن طريق نوعين من العلاقات، نذكر، التعيينات و الروابط.

التعيينات هي مجموعة من العلاقات المعرفة مسبقا مثل التكافؤ و الانفصال الخ... ذات معنى ثابت في جميع التفسيرات. في حين نستخدم الروابط لربط انطولوجيات مكملة لبعضها، هذا النوع من العلاقات يوضع من طرف خبراء مختلفون و المعنى المنسب لهذه الروابط يرتبط بمجال الاستخدام المعني.

اختلاف وجهات نظر هؤلاء الخبراء ينتج عنه عدم التجانس بين العلاقات التي تربط بين الانطولوجيات و يستوجب وضع علاقات أخرى بين هذه الروابط. و بالتالي يتم تشكيل محاذاة على مستوى اعلي مما يخلق شبكة من انطولوجيات متواصلة و قادرة على التواصل مع شبكات أخرى مماثلة.

بمكاننا تسميتها شبكة المعلومات المتعددة المستويات، مفهوم حيث نضع له قواعده و نقترح إخضاعه لنوعين من الاستدلال المنطقي.

المنهجية الأولى تستعمل المنطق الوصفي للاستدلال على شبكة المعلومات المتعددة المستويات و التي تنص على تحويل هذه الشبكة إلى انطولوجية واحدة مكتوبة بهذا المنطق.

أما المنهجية الثانية فهي تنص على تحويل شبكة المعلومات المتعددة المستويات إلى نظام موزع يستعمل المنطق الوصفي الموزع للاستدلال المنطقي.

الكلمات المفتاحية

شبكة المعلومات، انطولوجيا، المحاذاة بين الانطولوجيات، الاستدلال بالمنطق الوصفي، الاستدلال بالمنطق الوصفي الموزع.

TABLE DES MATIÈRES

TABLE DES MATIÈRES	xii
LISTE DES FIGURES	xiv
1 INTRODUCTION	1
1.1 PROBLÉMATIQUE	3
1.2 MOTIVATION	6
1.3 CONTRIBUTION	9
1.4 ORGANISATION DE LA THÈSE	11
2 PRÉSENTATION DES RÉSEAUX DE CONNAISSANCES ALIGNÉES	13
2.1 LANGAGES FORMELS DE REPRÉSENTATION DES CONNAISSANCES	15
2.1.1 Langage de représentation des connaissances et logique de description	15
2.1.2 Langages fondés sur les règles	22
2.1.3 Notion de contexte dans la représentation des connaissances	22
2.2 ALIGNEMENT D'ONTOLOGIES	25
2.2.1 Représentation syntaxique d'alignement d'ontologies	26
2.2.2 Représentation sémantique d'alignement d'ontologies	27
2.2.3 Formats des alignements d'ontologies	28
2.3 RÉSEAU D'ONTOLOGIES ALIGNÉES	33
2.3.1 Motivations	33
2.3.2 Composants d'un réseau d'ontologies alignées	34
2.3.3 Sémantique du réseau d'ontologies alignées	35
2.3.4 Raisonnement sur un réseau d'ontologies alignées	36

CONCLUSION	37
3 ÉTAT DE L'ART : FORMALISMES DE REPRÉSENTATION DES RÉ- SEaux D'ONTOLOGIES ALIGNÉES	39
3.1 DIFFÉRENTS FORMALISMES DE RÉSEaux DE CONNAISSANCES ALIGNÉES	41
3.1.1 Logiques de description distribuées (DDL)	42
3.1.2 \mathcal{E} -connection	45
3.1.3 Package Description Logic (PDL)	48
3.1.4 $E - SHIQ$	50
3.1.5 Integrated Distributed Description Logic (IDDL)	52
3.2 MODULARITÉ	56
3.3 MODÉLISATION CONTEXTUELLE DES CONNAISSANCES	57
3.4 DISCUSSION	59
3.4.1 La description et la structure des connaissances locales ne sont pas fixées	59
3.4.2 Représentation contextuelle des alignements	60
3.4.3 Plusieurs sémantiques peuvent être associées au forma- lisme proposé	60
3.4.4 Raisonnement	61
CONCLUSION	61
4 RÉSEau DE CONNAISSANCES MULTI-NIVEaux : SYNTAXE ET SÉ- MANTIQUE	65
4.1 SYNTAXE	67
4.1.1 Langage d'alignement	67
4.1.2 Nœud de connaissances	69
4.2 SÉMANTIQUE	71
4.2.1 Réseau de connaissances multi-niveaux en logique de description (DL)	72
4.2.2 Réseau de connaissances multi-niveaux en logique de description distribuée (DDL)	81

4.3	COMPARAISON DES DEUX APPROCHES SÉMANTIQUES DU RÉ- SEAU DE CONNAISSANCES MULTI-NIVEAUX	88
4.4	DISCUSSION	95
	CONCLUSION	96
5	IMPLÉMENTATION ET EXPÉRIMENTATION	97
5.1	IMPLÉMENTATION ET EXPÉRIMENTATION DE L'APPROCHE DL .	99
5.1.1	Architecture proposée	99
5.1.2	Format d'alignement adopté	100
5.1.3	Description du format proposé	100
5.1.4	Expérimentation et évaluation	105
5.1.5	Discussion	107
5.2	IMPLÉMENTATION DE L'APPROCHE DDL	112
5.2.1	Drago	112
5.2.2	Implémentation de l'approche DDL	114
	CONCLUSION GÉNÉRALE	117
	BIBLIOGRAPHIE	121

LISTE DES FIGURES

1.1	Représentation des connaissances par niveaux	9
4.1	Représentation récursive des noeuds	70
4.2	Interprétation d'un réseau de connaissances en utilisant l'approche DL	77
4.3	Exemple d'un réseau de connaissances multi-niveaux en Lo- gique de Description Distribuée	85
4.4	$\text{SystDis}(K) = \langle \{O_1, O_2, O_a\}, \{b_1, b_2, b_3, b_4\} \rangle$	86
4.5	L'interprétation du réseau K dans l'approche DDL	87
5.1	Architecture du raisonneur DLMLNKR	99
5.2	Les typologies possibles pour un réseau multi-niveaux de trois ontologies et les alignements respectifs	105
5.3	Exemple d'impact de taille du réseau d'ontologies	109
5.4	Exemple d'impact de nombre d'ontologies	110
5.5	Exemple d'impact de typologies pour le niveau0	110
5.6	Exemple d'impact de typologies pour le niveau1	110
5.7	Exemple d'impact de typologie pour le niveau2	111
5.8	Exemple d'impact de niveau pour la typologie 1	111
5.9	Exemple d'impact de niveau pour la typologie 2	111
5.10	Exemple d'impact de niveau pour la typologie 3	112
5.11	Architecture de Drago raisonnement par paire	113
5.12	Architecture du raisonneur DDLMLNKR	114

INTRODUCTION

SOMMAIRE

1.1	PROBLÉMATIQUE	3
1.2	MOTIVATION	6
1.3	CONTRIBUTION	9
1.4	ORGANISATION DE LA THÈSE	11

DANS cette thèse, nous évoquons la problématique de l'hétérogénéité sémantique des connaissances dans le cadre du Web sémantique ou dans le cadre des systèmes d'information au sein des organisations où plusieurs ontologies développées de manière indépendante peuvent être exploitées pour échanger des connaissances. Plusieurs travaux ont traité le problème de l'hétérogénéité sémantique des ontologies et ont proposé de les aligner par des connaissances supplémentaires. Dans le cadre de cette thèse, nous relevons le défi face aux techniques appliquées pour représenter et raisonner sur un ensemble d'ontologies alignées et nous proposons une contribution qui peut être considérée comme un prolongement des travaux existants sur l'intégration d'ontologies hétérogènes. Tenant compte du fait que les alignements qui permettent de lier une paire d'ontologies complémentaires possèdent leurs propres vocabulaires, ces alignements peuvent être développés indépendamment les uns par rapport aux autres et par conséquent, ils sont définis d'une manière contextuelle et peuvent générer une hétérogénéité sémantique. Dans la section 2.3.1 nous

présentons un scénario de motivation qui montre l'existence des cas réels ; où, les alignements ont leurs propres vocabulaires, qui peuvent être hétérogènes et qui nécessitent d'être alignés pour assurer un raisonnement correct sur l'ensemble du réseau d'ontologies alignées. Ce scénario va servir à donner des exemples le long de la thèse, il met en relief le problème d'hétérogénéité sémantique entre les alignements et montre le besoin de proposer un nouveau formalisme pour supporter cette hétérogénéité.

1.1 PROBLÉMATIQUE

Dans le cadre du Web sémantique ou dans le domaine des systèmes d'information, les ontologies sont utilisées pour représenter la sémantique des connaissances pour que cette dernière soit exploitable automatiquement par la machine. Plusieurs ontologies sémantiquement hétérogènes, évolutives, développées indépendamment dans des contextes différents et couvrant des domaines complémentaires peuvent être exploitées dans une même application qui a besoin de partager leurs connaissances. Le problème de la taille importante des ontologies, de l'hétérogénéité sémantique et de la non-accessibilité aux ontologies freine leur interopérabilité sémantique. Pour réduire le problème d'hétérogénéité sémantique, des connaissances complémentaires peuvent être introduites afin de décrire des correspondances entre les ontologies à exploiter. Ces correspondances représentent des relations entre les entités (termes ou formules) appartenant à des ontologies différentes et l'ensemble de ces correspondances est appelé alignement d'ontologies ou alignement inter-ontologies. Pour exploiter dans les raisonnements plusieurs ontologies hétérogènes et mises en correspondance, une solution simple consiste à traiter le système d'ontologies alignées comme une seule ontologie globale, où chaque ontologie locale et chaque alignement complètent simplement les connaissances sur un large domaine. Ceci peut être effectué par un processus de fusion réunissant toutes les ontologies et les alignements dans un même système centralisé ou bien en utilisant des algorithmes de raisonnement distribués, tout en se conformant à une logique classique (comme c'est le cas des systèmes SomeWhere Adjiman et al. (2005), SomeRDFS Adjiman et al. (2007) et SomeOWL Adjiman et al. (2006)). Une telle approche considère que les ontologies ainsi que les alignements décrivent chacun une partie d'une unique théorie globale, mais elle présente des inconvénients si les ontologies à combiner sont fortement hétérogènes, décrivant des points de vue ou des contextes différents, potentiellement incompatibles. L'autre solution envisagée, consiste à gérer l'ensemble des ontologies ainsi que

les alignements comme un réseau sémantique complexe où chaque nœud du réseau est une ontologie formalisant un domaine spécifique dont le contexte est différent des contextes des autres ontologies. Une telle approche nécessite la définition de formalismes modélisant le réseau d'ontologies alignées et offrant des techniques ainsi que des algorithmes pour un raisonnement contextuel. Plusieurs formalismes ont été proposés pour modéliser un réseau d'ontologies alignées pour un raisonnement contextuel mais ils diffèrent sur deux points essentiels : le premier point concerne la définition des alignements et l'autre point se focalise sur l'application de ces alignements. Pour définir les alignements, on distingue deux types de correspondances : on trouve les relations qu'on nomme *mapping* proposées pour réduire le problème d'hétérogénéité sémantique entre les termes des entités appartenant à des ontologies différentes (comme c'est le cas de Distributed Description Logic Borgida et Serafini (2003) et IDDL Zimmermann (2007)). Ces relations correspondent à un ensemble prédéfini de correspondances telles que la subsomption, l'équivalence, la disjonction, *etc.* et dont la sémantique est fixe dans toutes les interprétations (*e.g.*, $mt:appartient \xleftrightarrow{\perp} eq:appartient$). Le deuxième type d'alignement sert à lier des ontologies qui couvrent des domaines complémentaires (c'est le cas de \mathcal{E} -connection Kutz et al. (2004)). Ce type d'alignement est représenté par des liens entre des entités inter-ontologies qu'on appellera *links* (*e.g.*, $pr:G_1 \xleftrightarrow{\text{compose}} eq:DF_1$). Ce type de relation est défini par les experts dans le contexte de combinaison des domaines des ontologies à lier et la sémantique de ces *links* dépend de ce contexte.

Pour notre cas, nous nous intéressons à la proposition d'un formalisme qui supporte les deux types d'alignement afin de permettre la combinaison des ontologies hétérogènes associées à des contextes différents et couvrant des domaines complémentaires en même temps. En ce qui concerne le deuxième point de différence, qui est l'application ou le traitement des alignements, la majorité des formalismes proposés (tels que Distributed Description Logic Borgida et Serafini (2003), \mathcal{E} -connection Kutz

et al. (2004) et Package-based Description Logic Bao et al. (2009)) consiste à intégrer les alignements comme des connaissances externes au niveau de l'ontologie cible de la correspondance. L'alignement est alors défini et exploité selon le point de vue de cette ontologie et pour assurer le raisonnement sur ce réseau, chaque ontologie doit être dotée d'un mécanisme de raisonnement qui prend en charge des connaissances externes.

L'autre perception considère que les alignements doivent être exploités dans un niveau supérieur dit global et indépendant des ontologies locales (comme c'est le cas de IDDL Zimmermann (2007)). On entend par ontologies locales, les ontologies alignées appartenant à un réseau. L'avantage de cette approche est qu'elle permet une certaine flexibilité sur l'exploitation des alignements. On peut citer comme exemple, l'expressivité du langage de l'alignement qui peut être plus élevée que l'expressivité des langages définissant les ontologies locales. Elle permet aussi une meilleure réutilisabilité des alignements. Mais, il faut noter que seuls les *mappings* ont été traités dans cette approche, aucune proposition sur la prise en compte des *links* dans un niveau global n'a été donnée.

Nous nous intéressons à la deuxième approche pour la modélisation de la représentation des connaissances couvrant des domaines hétérogènes et complémentaires. La combinaison de ces ontologies nécessite la mise en place d'un nouveau formalisme qui supporte les *mappings* et les *links* entre une paire d'entités dont l'une appartient à une ontologie et l'autre entité appartient à une autre ontologie.

Les correspondances de type *links* sont introduites par des experts, or, on ne trouve pas d'experts couvrant la totalité des domaines des ontologies, alors, si on suppose que les alignements de type *links* par paire d'ontologies sont définis par des experts différents, avec chacun ses propres termes et son propre point de vue, le problème d'hétérogénéité sémantique va se poser de nouveau entre les *links* inter-ontologies. Notre contribution pour résoudre le problème de l'hétérogénéité sémantique des connaissances à plusieurs niveaux d'alignement consiste en la

proposition d'un nouveau formalisme appelé "Réseau de connaissances multi-niveaux" qui supporte cette hétérogénéité et propose une solution pour la réduire.

1.2 MOTIVATION

De nos jours, les entreprises sont de plus en plus intéressées par l'application des stratégies reposant sur l'utilisation des technologies de l'information et de la communication (TIC) pour réduire les coûts, et répondre à des impératifs de rentabilité. C'est dans ce cadre qu'un projet national de recherche (PNR) est mené avec un partenaire socio-économique et s'intitulant "Capitalisation des connaissances expertes pour l'aide à la décision : application à la maintenance industrielle à Sonelgaz"¹. Le but de ce projet est de mettre en place un système à base de connaissances assisté par ordinateur afin de rassembler le savoir et le savoir-faire dans le domaine de compression pour les turbines à gaz, l'exploiter par la suite pour de l'aide à la décision. La turbine à gaz est un système industriel complexe utilisé dans les installations de production et de génération d'électricité. La réalisation de ce projet a orienté la recherche vers le développement de nouvelles approches de gestion des connaissances, basées sur des paradigmes de l'intelligence artificielle collective tels que les ontologies alignées et le raisonnement sur le réseau global. Quatre ontologies ont été construites à savoir : l'ontologie équipement (eq) modélisant les connaissances techniques et hiérarchiques de la turbine à gaz, ces connaissances sont fournies par le constructeur de la turbine. Elle contient 5033 concepts où chaque concept représente un équipement composant de la turbine à gaz. On cite comme exemple le concept *detecteur-flamme* avec l'instance DF_1 . L'ontologie *pr* modélisant les connaissances concernant les pièces de rechange concernées par les équipements de la turbine, par exemple, le concept *garniture* avec l'instance G_1 . L'ontologie *zone* (zn) permet de représenter les emplacements des équipements. Et la quatrième ontolo-

1. Société nationale de l'électricité et du gaz. <http://www.sonelgaz.dz/>

gie *mt* est créée d'une manière semi-automatique à partir d'une base de données déjà existante. Elle couvre les opérations de maintenance préventive et curative effectuées sur les équipements de la turbine. Et exploite les deux premières pour fournir des détails sur les équipements et les pièces de rechange concernées par les interventions. L'exploitation de ces ontologies nécessite l'intégration des alignements entre elles, c'est pour cette raison que des outils d'appariement indépendants ont été appliqués afin de déceler les correspondances de type *mapping* telle que la correspondance *mt:appartient* $\xleftrightarrow{\perp}$ *eq:appartient* entre la paire d'ontologies (*mt*, *eq*) et la correspondance *pr:garniture* $\xleftrightarrow{\sqsubseteq}$ *eq:instrumentation* entre la paire d'ontologies (*pr*, *eq*). Ces ensembles de correspondances (*mappings*) ont été enrichis d'une façon semi-automatique par des *links* et en faisant intervenir plusieurs experts, soit un expert par paire d'ontologies. Cette opération a révélé l'existence du problème d'hétérogénéité sémantique entre les alignements inter-ontologies et, plus exactement entre les *links* appartenant à ces alignements. Nous citons comme exemple, l'apparition de termes différents au niveau de la paire d'alignement A_{pr-eq} et A_{eq-zn} , il s'agit de $A_{pr-eq}:\text{compose}$ et de $A_{eq-zn}:\text{fait-partie-de}$, ces links ont une sémantique identique. Il est clair que pour raisonner sur l'ensemble des ontologies avec les alignements qui existent entre elles, il faut d'abord résoudre le problème d'hétérogénéité sémantique entre les *links*. Comme pour ce cas, l'insertion d'une relation d'équivalence entre les links $A_{pr-eq}:\text{compose}$ et $A_{eq-zn}:\text{fait-partie-de}$ devient nécessaire.

Exemple 1 *Voici un extrait des ontologies et des alignements associés.*

alignement inter-alignements forment un *réseau de connaissances multi-niveaux*. La figure 1.1 illustre l'exemple présenté ci-dessus.

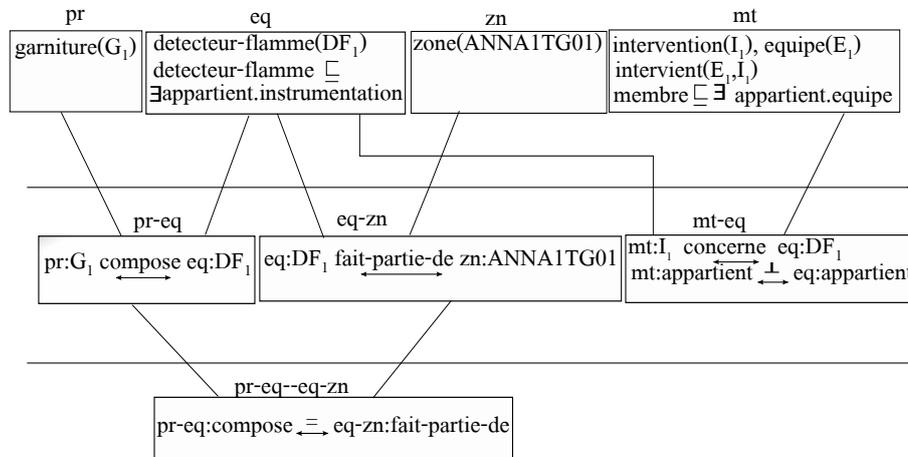


FIGURE 1.1 – Représentation des connaissances par niveaux

1.3 CONTRIBUTION

La contribution principale de ce travail consiste en la proposition d'un formalisme pour la représentation multi-niveaux de connaissances (réseau de connaissances multi-niveaux) ou (Multi-Level Networked Knowledge MLNK). Ce formalisme a été défini pour trouver des solutions au problème de l'hétérogénéité sémantique entre des alignements par paire d'ontologies. L'avantage de cette proposition réside tout d'abord dans le fait que : les alignements par paire d'ontologies sont gérés indépendamment des ontologies et des autres alignements. Chaque alignement possède un vocabulaire qui lui est propre (*links*), et une sémantique propre qui peut être distincte de celle des ontologies locales et de celle des alignements liant d'autres paires d'ontologies. Ceci permet d'une part, l'utilisation des alignements (*links*) dans des contextes différents, et d'autre part, favorise une meilleure ré-utilisabilité et évolution des alignements, notamment, la composition de ces derniers. Donc une formalisation multi-niveaux des connaissances rend une représentation contextuelle des alignements pos-

sible, et le fait d'associer une sémantique à cette formalisation, ceci permet un raisonnement contextuel à travers les niveaux des connaissances.

Les contributions apportées par ce travail sont :

1. **Une formalisation syntaxique** dépendant d'un langage d'alignement décrivant les *mappings* et les *links* entre deux vocabulaires donnés. La syntaxe de ce formalisme est définie d'une manière très générale est indépendante de tout langage, exploitant la technique de récursivité pour construire une base de connaissances hiérarchiquement structurée en niveaux ;
2. Une instanciation du formalisme générique évoqué dans le premier point. On s'intéresse au cas où les **ontologies sont en OWL**, et on propose **deux approches possibles pour l'interprétation** du réseau de connaissances multi-niveaux (MLNK). La première, utilise la logique de description, et la seconde, exploite les logiques de description distribuées. Les deux sémantiques ne peuvent être associées directement à la syntaxe du formalisme proposé, mais nécessitent une transformation du réseau en structures adaptées. Ces sémantiques ont été choisies car elles sont dotées de mécanismes de raisonnement opérationnels et peuvent être utilisés pour un raisonnement à court terme sur ce type de réseau ;
3. **Une comparaison des deux approches** a été effectuée afin de mesurer leurs pertinences et mettre en relief les cas où l'une des approches est plus adaptée que l'autre ;
4. **Une implémentation des approches DL et DDL** proposées pour interpréter le formalisme MLNK, à travers les prototypes "Description Logic Multi-Level Networked Knowledge Reasoner (DLMLNKR)" et "Distributed Description Logic Multi-Level Networked Knowledge Reasoner (DDLMLNKR)";
5. **Un ensemble d'expérimentations** réalisées grâce au prototype "Description Logic Multi-Level Networked Knowledge Reasoner ". Ces

expérimentations démontrent la faisabilité de l'approche et permettent de valider les propositions de cette thèse. Nous présentons une **comparaison et une analyse des résultats** pour **différentes typologies** du réseau de connaissances multi-niveaux.

À notre connaissance, aucune formalisation des réseaux de connaissances interconnectées à plusieurs niveaux n'a été proposée auparavant et le problème de l'alignement de vocabulaires d'alignement est une originalité de notre contribution.

1.4 ORGANISATION DE LA THÈSE

Pour le reste de la thèse, nous consacrons les deux premiers chapitres à l'état de l'art et les suivants au développement de la contribution :

- Dans le chapitre 2, nous introduisons des notions de base essentielles dans notre contexte de recherche telles que les langages de représentation des connaissances, la notion d'alignement d'ontologies, la représentation des connaissances contextuelles et les réseaux d'ontologies alignées.
- Dans le chapitre 3, nous présentons les travaux existants reliés à notre problématique : les différents formalismes élaborés pour représenter les réseaux d'ontologies alignés que nous avons classé par catégories. A travers une synthèse, nous avons aussi présenté dans ce chapitre l'apport de cette contribution par rapport aux travaux précédents.
- Dans Le chapitre 4 nous détaillons le formalisme que nous proposons en présentant une représentation syntaxique indépendante de toute sémantique. Puis nous suggérons de l'interpréter selon deux sémantiques existantes et qui ont fait leurs preuves de puissance, comme la logique de description, et de robustesse face à l'hétéro-

généité sémantique des connaissances dans un réseau d'ontologies, comme les logiques de description distribuées, et ceci, dans le but d'appliquer des raisonnements automatiques. Une comparaison des approches d'interprétation sémantiques proposées a été réalisée afin de montrer en quoi ces sémantiques diffèrent. Elle vise aussi à donner au lecteur des pistes pour choisir une sémantique plutôt que l'autre en fonction du contexte ou de l'application visée.

- Dans le chapitre 5, nous fournissons deux prototypes d'implémentation pour le système avec les approches DL et DDL montrant la faisabilité des propositions, avec une application des tests sur des connaissances alignées à plusieurs niveaux. Des expérimentations selon plusieurs cas ont donné des résultats satisfaisants.

Nous terminons avec une conclusion générale sans oublier de tracer quelques perspectives jugées utiles pour développer et améliorer le formalisme proposé.

PRÉSENTATION DES RÉSEAUX DE CONNAISSANCES ALIGNÉES

SOMMAIRE

2.1	LANGAGES FORMELS DE REPRÉSENTATION DES CONNAISSANCES	15
2.1.1	Langage de représentation des connaissances et logique de description	15
2.1.2	Langages fondés sur les règles	22
2.1.3	Notion de contexte dans la représentation des connaissances	22
2.2	ALIGNEMENT D'ONTOLOGIES	25
2.2.1	Représentation syntaxique d'alignement d'ontologies . . .	26
2.2.2	Représentation sémantique d'alignement d'ontologies . .	27
2.2.3	Formats des alignements d'ontologies	28
2.3	RÉSEAU D'ONTOLOGIES ALIGNÉES	33
2.3.1	Motivations	33
2.3.2	Composants d'un réseau d'ontologies alignées	34
2.3.3	Sémantique du réseau d'ontologies alignées	35
2.3.4	Raisonnement sur un réseau d'ontologies alignées	36
	CONCLUSION	37

CE chapitre est consacré à l'introduction des notions de base de la représentation des connaissances, des connaissances contextuelles

et des connaissances alignées ainsi que le raisonnement sur ces connaissances. Il définit les composants d'un réseau d'ontologies alignées, à savoir les ontologies et les alignements, en présentant aussi pour les ontologies, les notions de sémantique formelle pour un éventuel raisonnement. Un exemple de langage de représentation des connaissances est détaillé, il s'agit de la famille des logiques de description. Un aperçu des formats de représentation d'alignement est aussi fourni. Le réseau de connaissances alignées suppose l'existence d'ontologies multiples capables d'interagir à travers les alignements qui représentent la mise en correspondance de ces ontologies. Le chapitre s'organise de la façon suivante : Dans la section 2.1, une description générale des langages de représentation des connaissances est donnée du point de vue syntaxique et sémantique pour le raisonnement, en détaillant le langage des logiques de description, le formalisme le plus utilisé pour le Web sémantique. Ensuite, en section 2.2, on définit de façon très générale la notion d'alignement d'ontologies et on présente des formats pour leur représentation. Enfin, la section 2.3 détaille les composants, la sémantique et le raisonnement sur les réseaux de connaissances alignées, en fournissant une définition générale.

2.1 LANGAGES FORMELS DE REPRÉSENTATION DES CONNAISSANCES

La représentation formelle des connaissances (Sowa 2000) s'est développée dans le domaine de l'intelligence artificielle, elle permet d'automatiser le raisonnement sur ces connaissances. La logique est une des techniques de la représentation des connaissances où ces dernières sont représentées à l'aide d'un langage formel.

2.1.1 Langage de représentation des connaissances et logique de description

Un langage de représentation de connaissances L se caractérise par une syntaxe (la manière dont les formules sont écrites) et une sémantique (la manière dont sont interprétées ces formules). Les logiques de description forment une famille de langages de représentation de connaissances qui peuvent être utilisées pour représenter les connaissances d'un domaine d'application d'une façon structurée et formelle.

Syntaxe d'un langage de représentation des connaissances

La syntaxe est caractérisée par un ensemble de symboles et par des règles de construction qui permettent d'obtenir des formules à partir de ces symboles en appliquant des constructeurs spécifiques au langage.

On parle de signature ou de vocabulaire pour désigner un ensemble de termes structurés qui est un sous-ensemble des symboles du langage. Chaque signature permet de définir un ensemble de formules déterminées par le langage. Et le terme ontologie fait référence à un couple composé d'un vocabulaire spécifique conçu pour modéliser un domaine donné du monde réel et de l'ensemble de formules construites à partir de ce même vocabulaire.

Définition 2.1 (Ontologie) *Soit L un langage de représentation de connaissances. Une ontologie O de L est un couple $\langle \Sigma(O), A(O) \rangle$ où $\Sigma(O)$ est une signature dans L et $A(O)$*

est un ensemble d'axiomes qui est un sous-ensemble des formules que l'on peut construire à partir de $\Sigma(O)$.

Sémantique d'un langage de représentation des connaissances

La sémantique d'un langage L est caractérisée par une notion d'interprétation associant aux symboles d'une signature des éléments d'un domaine d'interprétation et des contraintes définissant la satisfaction des formules.

Définition 2.2 (Interprétation) *Une interprétation I d'une signature Σ d'un langage L est un couple $\langle \Delta^I, \cdot^I \rangle$ où Δ^I est le domaine d'interprétation et \cdot^I est une fonction d'interprétation.*

Une interprétation de l'ontologie O est une interprétation de la signature de O . La notion d'interprétation n'est pas suffisante pour raisonner sur les connaissances, on a besoin d'une autre notion qui permet de relier les formules aux interprétations qui satisfont ces formules. Il s'agit de la notion de satisfaction.

Définition 2.3 (Relation de satisfaction) *La relation de satisfaction \models_L d'un langage L définit une relation entre les interprétations d'une signature Σ et les formules sur cette signature.*

On dit qu'une interprétation satisfait une ontologie si et seulement si elle satisfait tous les axiomes de cette ontologie, elle est appelée alors, modèle de l'ontologie.

Définition 2.4 (Modèle d'une ontologie) *Soit L un langage de représentation de connaissances. Soit O une ontologie de L . Une interprétation I de O satisfait O si et seulement si I satisfait tous les axiomes de O . La notation $I \models_L O$ permet de formuler que I est un modèle de O .*

Pour résumer, la sémantique d'un langage L de représentation des connaissances est caractérisée par ses interprétations et la relation de satisfaction.

Raisonner avec un langage de représentation des connaissances

Le raisonnement sur les connaissances se traduit par la vérification de la cohérence d'une ontologie (c.à.d. pas de contradiction mais ne détermine pas si une ontologie représente convenablement un domaine de connaissances) et les conséquences logiques d'une ontologie.

Définition 2.5 (Cohérence) *Soit L un langage de représentation de connaissances. Une ontologie O de L est cohérente si et seulement s'il existe un modèle de O .*

Définition 2.6 (Conséquence logique) *Soit L un langage de représentation de connaissances. Soit O une ontologie et f est une formule de L . On dit que f est une conséquence sémantique de O lorsque tous les modèles de O satisfont f et on la note $O \models_L f$.*

Exemple de langage : la famille des Logiques de Description

Cette section est consacrée à la présentation d'une famille de langages de représentation de connaissances, il s'agit de la famille des logiques de description (Baader et al. 2003). Les connaissances dans ces logiques sont représentées sous forme de concepts, rôles et individus. Les concepts et les rôles peuvent être primitifs (atomiques ou élémentaires) et servent de base à la construction d'autres concepts et rôles définis. Ils possèdent une description structurée et une sémantique leur a été attribuée par l'intermédiaire d'une interprétation. Les concepts définissent des classes d'individus, tandis que les rôles représentent des relations entre des couples d'individus et possèdent des propriétés telles que la transitivité et l'inverse de rôle (voir la table 2.1).

Syntaxe Les connaissances en DL sont divisées selon deux niveaux :

- **T-box** (terminology box) qui représente un ensemble de formules relatives aux concepts et rôles.
- **A-box** (assertions box) qui représente un ensemble de formules relatives aux instances des concepts et des rôles.

Comme tout langage de représentation de connaissances, les logiques de description possèdent la notion de signature, cette dernière étant constituée des types concepts, rôles et individus. Le langage permet également la création des éléments complexes ainsi que des axiomes en utilisant des constructeurs autorisés par le langage.

Définition 2.7 (Signature) *Une signature en logique de description est un triplet $\Sigma = \langle C, \mathcal{R}, \mathcal{U} \rangle$ où C est un ensemble fini de concepts atomiques, \mathcal{R} un ensemble fini de rôles atomiques et \mathcal{U} un ensemble fini d'individus.*

En référence à l'exemple 1, la signature de l'ontologie eq comporte les classes `detecteur-flamme` et `instrumentation`, le rôle `appartient` et l'individu `DF1`.

Définition 2.8 (Formule) *Une formule f de DL est l'un des axiomes autorisés par le langage DL et qui se trouve dans la table 2.1.*

La table 2.1 donne en exemple les syntaxes et les sémantiques de chaque terme atomique, ainsi que différents constructeurs de concepts, constructeurs de rôle et axiomes où :

- Le concept \top dénote le concept le plus général, il inclut tous les individus possibles ;
- le concept \perp dénote le concept le plus spécifique et il est vide ;
- le constructeur \sqcap permet de définir une conjonction entre concepts ou entre rôles ;
- la quantification universelle : $\forall R.C$ est un concept défini où tous les éléments qui sont en relation avec les éléments du concept C par le rôle R appartiennent à $\forall R.C$;
- la quantification existentielle $\exists R.C$ est un concept défini où parmi les éléments qui sont en relation avec les éléments de C par le rôle R appartiennent à $\exists R.C$;
- etc...

Le langage $\mathcal{AL} = \{\top, \perp, C \sqcap D, \forall R.C, \exists R.C\}$ peut être enrichi des constructeurs pour constituer d'autres langages. Prenons par exemple l'enrichissement de \mathcal{AL} avec la négation de concept primitif ou défini, qui est noté $(\neg C)$ donne naissance à un autre langage de la famille des logiques de description dit *Attribute Language with Complement* (\mathcal{ALC}).

Terme atomique	Syntaxe	Sémantique	
individu	a	$(a)^I \in \Delta$	(\mathcal{AL})
concept atomique	A	$(A)^I \subseteq \Delta$	(\mathcal{AL})
role atomique	R	$(R)^I \subseteq \Delta \times \Delta$	(\mathcal{AL})
Constructeur de concepts	Syntaxe	Sémantique	
concept universel	\top	$(\top)^I = \Delta$	(\mathcal{AL})
concept vide	\perp	$(\perp)^I = \emptyset$	(\mathcal{AL})
conjonction	$C \sqcap D$	$(C \sqcap D)^I = (C)^I \cap (D)^I$	(\mathcal{AL})
disjonction	$C \sqcup D$	$(C \sqcup D)^I = (C)^I \cup (D)^I$	\mathcal{U}
négation	$\neg C$	$\Delta \setminus (C)^I$	\mathcal{C}
restriction existentielle	$\exists R.C$	$\{x \mid \exists y, \langle x, y \rangle \in (R)^I \wedge y \in (C)^I\}$	(\mathcal{AL})
restriction de valeur	$\forall R.C$	$\{x \mid \forall y, \langle x, y \rangle \in (R)^I \Rightarrow y \in (C)^I\}$	(\mathcal{AL})
restriction numérique	$\leq nR$	$\{x \mid \{y, \langle x, y \rangle \in I(R)\} \leq n\}$	\mathcal{N}
restriction numérique	$\geq nR$	$\{x \mid \{y, \langle x, y \rangle \in I(R)\} \geq n\}$	\mathcal{N}
nominaux	$\{a_1, \dots, a_n\}$	$\{(a_1)^I, \dots, (a_n)^I\}$	\mathcal{O}
Constructeur de roles	Syntaxe	Sémantique	
conjonction de roles	$R \sqcap S$	$(R \sqcap S)^I = (R)^I \cap (S)^I$	$\cdot \sqcap$
disjonction de roles	$R \sqcup S$	$(R \sqcup S)^I = (R)^I \cup (S)^I$	$\cdot \sqcup$
complément de roles	$\neg R$	$(\Delta \times \Delta) \setminus (R)^I$	$\cdot \neg$
clôture transitive	R^+	clôture transitive de $(R)^I$	$\cdot +$
role inverse	R^-	$\{\langle y, x \rangle \mid \langle x, y \rangle \in (R)^I\}$	$\cdot -$
Axiomes de la T-Box	Syntaxe	Contrainte d'interprétation	
subsumption	$C \sqsubseteq D$	$(C)^I \subseteq (D)^I$	(\mathcal{AL})
inclusion de roles	$R \sqsubseteq S$	$(R)^I \subseteq (S)^I$	\mathcal{H}
transitivité de roles	$\text{Trans}(R)$	$(R)^I = (R^+)^I$	\mathcal{S}
Axiomes de la A-Box	Syntaxe	Contrainte d'interprétation	
appartenance à un concept	$C(a)$	$(a)^I \in (C)^I$	(\mathcal{AL})
appartenance à un role	$R(a_1, a_2)$	$\langle (a_1)^I, (a_2)^I \rangle \in (R)^I$	(\mathcal{AL})
identité	$a_1 = a_2$	$(a_1)^I = (a_2)^I$	

TABLE 2.1 – logique de description

On peut actuellement introduire la notion d'ontologie en logique de description à partir de ces définitions.

Définition 2.9 (Ontologie) *Une ontologie en logique de description est une paire $\langle \Sigma, F \rangle$ telle que Σ est une signature et $F = \langle T, A \rangle$ contient des axiomes T de la **T-box**, et des assertions A de la **A-box**, tel que c'est défini en bas de la table 2.1.*

Exemple 2 Si on reprend l'exemple 1, la signature de l'ontologie eq, $\Sigma(\text{eq}) = \{\text{detecteur-flamme}, \text{instrumentation}, \text{DF}_1, \text{appartient}\}$ et l'ensemble des ses axiomes $F(\text{eq}) = \{\text{detecteur-flamme}(\text{DF}_1), \text{detecteur-flamme} \sqsubseteq \exists \text{appartient.instrumentation}\}$

La syntaxe des logiques de description étant définie et il est possible actuellement de présenter leur sémantique suivant la théorie des modèles.

Sémantique Une sémantique est attribué aux concepts et aux roles en logiques de description. Les concepts qui contiennent un ensemble d'individus sont donc interprétés comme des sous-ensembles du domaine d'interprétation. Les couples d'individus sont liés par des roles et sont interprétés comme des sous-ensembles du produit $\Delta^I \times \Delta^I$.

Définition 2.10 (Interprétation) Soit $\Sigma = \langle \mathcal{C}, \mathcal{R}, \mathcal{U} \rangle$ une signature en logique de description. Une interprétation I de Σ est une paire $\langle \Delta^I, \cdot^I \rangle$, où le domaine d'interprétation Δ^I est un ensemble non vide, et \cdot^I est une fonction qui associe à chaque concept un sous-ensemble de Δ^I , pour chaque role un sous-ensemble de $\Delta^I \times \Delta^I$ et chaque individu est un élément de Δ^I .

La relation de satisfaction en logique de description est une opération qui permet de vérifier un axiome et ceci moyennant la contrainte d'interprétation qui lui a été associée. On peut citer à titre d'exemple qu'un concept C est subsumé par un concept D (noté par $C \sqsubseteq D$) si et seulement si $C^I \subseteq D^I$ pour toute interprétation I .

Définition 2.11 (Relation de satisfaction) À chaque axiome de logique de description est associée une contrainte d'interprétation (voir la troisième colonne de la table 2.1). On dit qu'une interprétation I satisfait une formule f si et seulement si la contrainte d'interprétation associée à l'axiome de f est satisfaite. On note cette relation comme suit : $I \models f$.

La notion de modèle est définie pour désigner l'interprétation qui satisfait toutes les formules ou les axiomes d'une ontologie donnée.

Définition 2.12 (Modèle) *Soit $O = \langle \Sigma, F \rangle$ une ontologie de logique de description. Une interprétation de Σ est un modèle de O si et seulement si I satisfait l'ensemble des formules de F , on note cette relation $I \models F$. On note également $\mathcal{M}(O)$ l'ensemble des modèles de O .*

Raisonnement Les deux notions essentielles de raisonnement logique sont la consistance d'une ontologie et la conséquence sémantique ou l'inférence.

Définition 2.13 (Consistance) *Une ontologie O en logique de description est consistante si et seulement s'il existe au moins un modèle de O .*

À partir de la notion de modèle, nous pouvons définir les notions de conséquence sémantique, ainsi que celles relatives au raisonnement.

Définition 2.14 (Conséquence sémantique) *Soient O une ontologie et f une formule de logique de description. On dit que f est une conséquence sémantique de O lorsque tous les modèles de O satisfont f qu'on note $O \models f$.*

Les logiques de description sont applicables dans le Web Sémantique et sont utilisées dans le Standard OWL (Dean et Schreiber 2004). Les langages OWL-DL et OWL-Lite (Web Ontology Languages) de W3C sont basés sur les logiques de description. Et une des techniques utilisée pour le raisonnement en DL est l'algorithme de tableau, qui a été implémenté dans de nombreux raisonneurs, on peut citer à titre d'exemple Fact++¹, Pellet², Hermit³. Pour la preuve de tableau, on passe par les étapes suivantes :

- éliminer la TBox en remplaçant les termes par leur définition ;
- transformer en Negative Normal Form (NNF) ;
- montrer que la négation de ce qu'on veut prouver est insatisfaisable en appliquant les règles de tableau pour DL (Baader et Sattler 2000).

1. <http://OWL.man.ac.uk/factplusplus/>

2. <http://pellet.OWLdl.com/>

3. <http://hermit-reasoner.com/>

2.1.2 Langages fondés sur les règles

Parmi les connaissances acquises auprès des experts dans un domaine donné, on trouve les connaissances de type contraintes ou critères de décision. Pour les formaliser sous forme de faits et de règles, on a besoin de langage à base de règles. C'est ce type de formalisation qui permet d'automatiser les décisions et de développer des systèmes experts. Pour le Web sémantique, le langage OWL est étendu avec les règles, c'est SWRL (Horrocks et al. 2004) et il a été implémenté dans les raisonneurs Pellet⁴ et Hermit⁵.

Semantic Web Rule Language

(SWRL) est issu d'une intégration des ontologies et des règles. C'est un langage pour le Web sémantique, il est le résultat de la combinaison de OWL et de RuleML (Boley et al. 2001). Il étend l'ensemble des axiomes de OWL en y intégrant les clauses de Horn. Les règles définies par SWRL sont composées d'un antécédent (corps) et de la conséquence (en-tête) et si les conditions définies dans l'antécédent sont vérifiées alors les conditions définies dans la conséquence sont exécutées. Exemple de règle : $\text{parent}(?x, ?y) \wedge \text{frère}(?y, ?z) \Rightarrow \text{oncle}(?x, ?z)$. Les alignements aussi peuvent être formalisés sous forme de règles (voir 2.2.3).

2.1.3 Notion de contexte dans la représentation des connaissances

Plusieurs ontologies peuvent être construites différemment même pour un même domaine ou pour des domaines complémentaires. Cette différence peut concerner la terminologie où un expert peut utiliser le terme voiture tandis qu'un autre préfère l'utilisation du terme automobile. Elle peut concerner aussi la sémantique des connaissances. Un expert considère le terme "java" pour définir un langage de programmation du point de vue informatique, un autre expert l'utilise pour définir un animal. Cette

4. <http://pellet.OWLdl.com/>

5. <http://hermit-reasoner.com/>

différence peut provenir aussi de la granularité dans la spécialisation des concepts. L'un peut considérer que le concept "employé" se divise en sous-concepts "Cadre", "Maitrise" et "exécutant" et un autre expert peut le diviser en "administratif" et "enseignant". Toutes ces différences sont dûes au fait que les experts sont différents dans leurs subjectivités, leurs préférences, leurs domaines, leurs objectifs et leurs expériences, etc. donc ils sont différents dans leurs points de vue et ceci dépend du contexte d'utilisation de la connaissance. La notion de contexte est apparu dans le domaine de représentation des connaissances pour représenter la notion de point de vue. Le travail de Benerecetti et al. (2001) montre trois formes du contexte dans la représentation des connaissances :

1. Un contexte est une portion, il représente un sous-ensemble des connaissances d'un domaine donné ;
2. un contexte est une approximation, il représente des connaissances dont le niveau de granularité ou d'abstraction est différent de celui des connaissances des autres contextes ;
3. un contexte correspond à une perspective, il représente des connaissances liées à des évènements extérieurs tels que le lieu, le temps, etc. et qui diffèrent d'un contexte à un autre.

En intelligence artificielle, les travaux sur les contextes ont été initiés par McCarthy (1987) et McCarthy (1993) où il présente un modèle permettant la représentation et le raisonnement multi-contextes ; toute assertion A de la logique est associée à un contexte sous forme $c:A$ (où c correspond à un contexte). Les propositions sont aussi associées à un contexte, notées $is-true(c,p)$ indiquant que la proposition p est considérée vrai par le contexte c , et les raisonnements sont réalisés dans ce contexte. Par exemple, le concept $phd-etudiant$ est inclu dans le concept $etudiant$ dans un contexte c est formulé de la façon suivante : $is-true(c, phd-etudiant \sqsubseteq etudiant)$. Le contexte peut être décrit dans un méta-niveau où les combinaisons entre contextes sont aussi possibles,

exemple : $is - true(c, ist - true(d, p))$ où c et d sont des contextes et la proposition p est aussi valide dans c si elle est valide dans d . Ils peuvent avoir des propriétés, on cite à titre d'exemple : $\forall x(c(x)) \Rightarrow is - true(x, p)$ veut dire que la proposition p est valide pour tout contexte de type c . Après cela, plusieurs études sur les logiques contextuelles spécialement dans le domaine de la représentation des connaissances ont été apportées à commencer par le travail de Guha (1992). Giunchiglia (1993) définit le contexte du point de vue raisonnement, il considère que le contexte représente un ensemble de connaissances modélisées pour un but donné, où, le raisonnement sur cet ensemble est performant ou cohérent. La base de connaissances générale, dans ce cas est une interaction entre les contextes (c_1, c_2, \dots, c_n) où n est un nombre fini de contextes, donc le raisonnement sur la base de connaissances est un raisonnement à travers les contextes. Les travaux de Giunchiglia et Serafini (1994) et de Ghidini et Giunchiglia (2001) introduisent des règles d'inférence pour faire des raisonnements au sein des contextes et aussi entre des contextes via des règles de passerelles connectant les contextes locaux. Distributed First Order Logics est proposé par Ghidini et Serafini (2000) et qui permet de formaliser la sémantique de réseaux d'ontologies exprimées en logique de premier ordre et reliées par des passerelles. Bao et al. (2010) ont étendu la théorie de McCarthy (McCarthy 1987), (McCarthy 1993) en fournissant des formalisations concrètes avec le prédicat $isin$ et un nombre de constructeurs ont été introduits pour combiner des contextes $(c1 \cap c2, c1 \cup c2, \neg c)$ et pour relier les contextes $(c1 \Rightarrow c2)$. Ces travaux ont été formalisés comme des théories propositionnelles de premier ordre et des bridges rule sont fournies pour faire intéropérer les différentes théories contextuelles. D'autres travaux plus adaptés pour le Web sémantique car ils se basent sur les logiques de description ont été développés (Borgida et Serafini 2003), (Kutz et al. 2004), (Bao et al. 2009), (Zimmermann 2007), (Vouros et Santipantakis 2012). Parmi les travaux existants, on trouve deux grand axe de recherche

sur la représentation et le raisonnement sur des connaissances contextualisées :

1. les travaux qui modélisent le contexte comme étant une méta-connaissance dans un niveau supérieur où les contextes peuvent être reliés comme c'est le cas de CKR (Homola et Serafini 2012), (Joseph et Serafini 2011) et (Joseph 2015) ;
2. les travaux qui contextualisent des ontologies déjà existantes et développées indépendamment les unes des autres. Ceci permet d'intégrer la notion de point de vue dans les ontologies du Web sémantique ou même dans le Web d'entreprise. On considère que chaque ontologie représente un contexte et les ontologies contextualisées deviennent des ontologies locales d'un réseau d'ontologies connectées par des connaissances supplémentaires dites alignements (Borgida et Serafini 2003), (Kutz et al. 2004), (Bao et al. 2009), (Zimmermann 2007), (Vouros et Santipantakis 2012).

Il existe aussi des travaux qui utilisent la notion de contexte dans des ontologies multi-point de vue. On cite comme exemple le travail de Bach (2006) et de Djakhdjakha et al. (2014). Le formalisme que nous proposons entre dans le cadre de la contextualisation des ontologies indépendantes et alignées. La suite du chapitre est focalisée sur la description d'un réseau d'ontologies alignées. Nous commençons par présenter les alignements d'ontologies qui constitue un composant important dans ce type de réseau.

2.2 ALIGNEMENT D'ONTOLOGIES

L'alignement d'ontologie est le résultat de l'appariement entre des ontologies, il représente les correspondances établies entre ces ontologies pour les connecter en vue de les exploiter dans un raisonnement sur le système global ou des transformations d'ontologies. Le livre de Euzenat et Shvaiko (2007) détaille les techniques ainsi que les outils existants pour

l'appariement entre ontologies. Dans cette section, nous allons décrire la représentation syntaxique d'alignement d'ontologies section 2.2.1, la représentation sémantique section 2.2.2 et un aperçu sur les formats disponibles pour le stockage des alignements est donné dans la section 2.2.3

2.2.1 Représentation syntaxique d'alignement d'ontologies

D'une façon la plus générale, un alignement d'ontologies peut être défini comme un ensemble de correspondances. Et chaque correspondance est une expression de relation existante entre des entités appartenant à des ontologies différentes. L'alignement est donc composé des entités et des relations où les entités ne peuvent être que des entités alignables appartenant aux vocabulaires déjà existants des ontologies à appairer. Les correspondances sont des connaissances inter-ontologies et pour les représenter on a besoin d'un langage doté de syntaxe et d'une sémantique (section 2.2.2). Pour définir ce langage, Euzenat et Shvaiko (2007), ont séparé le langage d'entités du langage d'ontologie afin d'exprimer les entités susceptibles d'être mises en relation pour appartenir à des correspondances.

Définition 2.15 (Langage d'entités) *Soit V un vocabulaire d'ontologie, Le langage d'entités d'un vocabulaire est donné par $E(V)$, une fonction qui permet de déterminer les entités alignables du vocabulaire.*

Exemple 3 *En logique de description, les entités alignables sont les concepts, les rôles et les individus.*

Afin d'exprimer les correspondances entre les entités, un ensemble de relations peut être défini pour lier les entités alignables, et l'ensemble des entités alignables reliées entre elles constitue le langage d'alignement.

Définition 2.16 (Langage d'alignement) *Un langage d'alignement L_A est un couple $\langle E, \mathcal{R} \rangle$ où E est un langage d'entité et \mathcal{R} est un ensemble de symboles.*

Définition 2.17 (Correspondance) *soient O_1, O_2 deux ontologies. Soit $L_A = \langle E, \mathcal{R} \rangle$ un langage d'alignement. Une correspondance est un 5-uplet $\langle id, e_1, e_2, r, n \rangle$ où :*

- id est un identifiant unique de la correspondance ;
- $e_1 \in E(O_1)$ et $e_2 \in E(O_2)$ sont des entités alignables des ontologies O_1 et O_2 respectivement ;
- $r \in \mathcal{R}$ est une relation entre e_1 et e_2 ;
- $n \in [0..1]$ est une mesure de confiance.

Dans la suite de la thèse, nous utilisons une définition réduite de la correspondance où l'identifiant et la mesure de confiance sont ignorés car ce qui nous intéresse sont les entités et les relations.

- Exemple 4**
- Un exemple de langage possible où l'ensemble des symboles de la relation $\mathcal{R} = \{\sqsubseteq, \equiv, \perp, \in, =\}$, et le langage d'alignement se réduit aux termes de vocabulaires déjà existants (e.g., $mt:appartient \xleftrightarrow{\perp} eq:appartient$) ;
 - un autre exemple consiste à considérer les relations entre les entités de vocabulaires différents comme des rôles qu'on appelle des links (e.g., $pr:G_1 \xleftrightarrow{\text{compose}} eq:DF_1$).

2.2.2 Représentation sémantique d'alignement d'ontologies

La sémantique d'un langage d'alignement définit comment les correspondances sont interprétées. Pour une famille d'ontologies liées par un ensemble de correspondances, il existe plusieurs façons d'interprétation :

- l'interprétation de l'alignement est une portion d'un domaine d'interprétation unique comportant aussi les interprétations des ontologies, elle est adoptée dans la sémantique non contextuelle du réseau de connaissances alignées (voir section 2.3.3).
- l'interprétation de l'alignement est définie par une relation dite "relation de domaine" entre les interprétations locales qui sont forcément distinctes. Cette relation fait le lien entre un élément d'un domaine d'interprétation et l'élément correspondant dans l'autre domaine d'interprétation du point de vue de ce dernier. Elle est utilisée par la sémantique contextuelle (voir section 2.3.3).
- les alignements sont interprétés dans un domaine global indépen-

dant de tous les domaines d'interprétation des ontologies locales. Une fonction d'égalisation qui permet de projeter un élément d'un domaine d'interprétation local à son image dans le domaine universel ou global : $\epsilon_o : D_o \rightarrow U$ où D_o est un domaine d'interprétation local et U est un domaine universel. Elle est utilisée par la sémantique contextuelle et intégrée (voir section 2.3.3).

2.2.3 Formats des alignements d'ontologies

Cette section décrit divers formats proposés pour représenter des alignements dans le but de les stocker afin d'effectuer des opérations telles que la comparaison, la combinaison ou autres.

OWL peut être considéré comme un langage pour représenter des correspondances entre ontologies OWL uniquement, en utilisant les primitives *equivalentClass*, *equivalentProperty*, *subClass* et autres qui permettent de relier des éléments provenant d'ontologies différentes.

```
<OWL:Property rdf:about="onto1\#author">
<OWL:equivalentProperty rdf:resource="onto2\#author"/>
</OWL:Property>
<OWL:Class rdf:about="onto1\#Book">
<OWL:equivalentClass rdf:resource="onto2\#Volume"/>
</OWL:Class>
<OWL:Class rdf:about="onto2\#titre">
<OWL:subClass="onto1\#name"/>
<OWL:Class>
```

onto1 et *onto2* sont deux ontologies différentes, le concept *author* de *onto1* est équivalent à *author* de *onto2*, le concept *Book* de *onto1* est équivalent au concept *Volume* de *onto2* et *titre* de *l'onto2* est un sous-concept de *name* de l'ontologie *onto1*.

L'inconvénient de ce format est que son utilisation est limitée aux onto-

logies en OWL où la sémantique adoptant le domaine d'interprétation unique est exigée et ne tolère pas une gestion ni une réutilisation autonome et indépendante des alignements puisque ces derniers sont intégrés dans l'ontologie globale.

SWRL (Semantic Web Rule Language) (Horrocks et al. 2004) SWRL est utilisé pour représenter les alignements sous forme de règles, c'est le cas où une règle est constituée à partir des entités appartenant à des ontologies différentes. Nous présentons un exemple de correspondance exprimée en SWRL et qui permet de déterminer si une équipe de O_3 est disponible c'est-à-dire qui n'est pas en cours d'intervention ou n'est pas programmée pour une autre intervention. (Voir l'aperçu des ontologies dans l'article Klai et Khadir (2013))

$$\begin{aligned} & \text{SWRLb:dateTime}(?presentTime) \wedge \text{Inetrvient}(?O3:Equipe1, ?O2:Int) \wedge \\ & \text{endAt}(?O2:Int, ?EndTime1) \wedge \text{SWRLb:greaterhand}(?EndTime1, ?presentTime) \\ & \vee (\text{Planifier}(?O3:Equipe1, ?O3:Int2)) \wedge \text{endAt}(?O3:Int2, ?Endtime2) \wedge \\ & \text{SWRLb:greaterThand}(?EndTime2, ?presentTime) \rightarrow \text{availability}(?O3:Equipe1, "False") \end{aligned}$$

C-OWL (Contextualized OWL) C-OWL est une extension du langage (Bouquet et al. 2003) OWL utilisé pour exprimer des mappings entre des concepts ou roles d'ontologies décrites en OWL. Les constructeurs de C-OWL sont appelés "passerelles" (bridge rules). Les passerelles sont directionnelles puisque la relation est exprimée du point de vue de l'ontologie cible. Autrement dit, une passerelle d'une ontologie O_1 vers une autre ontologie O_2 n'est pas équivalente à une passerelle inverse de O_2 vers O_1 . Il est une application du formalisme des logiques de description distribuées (DDL), et il permet d'exprimer des alignements relativement simples car il n'y a pas de classes construites dans une passerelle, seulement des classes nommées. Voici un exemple de mapping écrit sous le format C-OWL :

```
<cowl:Mapping>
```

```

<owl:sourceOntology>
<OWL:Ontology rdf:about="http://source.owl"/>
<owl:sourceOntology>
<owl:targetOntology>
<OWL:Ontology rdf:about="http://target.owl"/>
<owl:targetOntology>
<owl:bridgeRule>
<owl:Equivalent>
<owl:source>
<OWL:Class rdf:about="http://swrc\#Article" />
</owl:source>
<owl:target>
<OWL:Class_ rdf:about="http://shoe\#Paper" />
</owl:target>
<owl:Equivalent>
<owl:bridgeRule>
<owl:Mapping>

```

Cet extrait décrit le *mapping* qui exprime que le concept *Article* de l'ontologie source *swrc* est équivalent au concept *Paper* de l'ontologie cible *shoe*.

C-OWL a été étendu par Vouros et Santipantakis (2012) pour stocker des *links*.

```

<owl:linking>
<owl:LinkProperty
  rdf:resource="http://unit4.OWL\#presentedAt"/>
</owl:Linking>
<OWL:Class rdf:about="http://unit1.OWL\#Article">
<OWL:equivalentClass>
<OWL:Restriction>
<OWL:onProperty

```

```
    rdf:resource="http://unit4.OWL#presentedAt"/>
<OWL:allValuesFrom
    rdf:resource="http://unit4.OWL#Event"/>
<OWL:Restriction>
<OWL:equivalentClass>
</OWL:Class>
```

Cet extrait décrit une correspondance de type *link* et qui exprime que le concept *Article* de l'ontologie *unit1* est présenté par *presentedAt* le concept *Event* de l'ontologie *unit4*.

MAFRA signifie Mapping Framework (Maedche et al. 2002). C'est un système complet pour extraire des alignements d'ontologies et les exécuter en tant que transformation de données d'une ontologie à une autre. Dans MAFRA, le format d'alignement est spécifié par une ontologie appelée *Semantic Bridging Ontology* (SBO). Ce système est destiné à traduire des instances d'une ontologie vers une autre. Lorsqu'une instance d'une entité identifiée dans une correspondance est repérée et qu'elle vérifie les conditions, elle est traduite en une instance de l'entité identifiée dans la seconde ontologie, tout en lui appliquant des conversions par le biais des services. Ce format est expressif, mais il manque de polyvalence car il est fortement lié à l'architecture de MAFRA.

Le format de l'API d'alignement Le format associé à l'API d'alignement Euzenat (2004) permet le stockage, la manipulation et l'utilisation d'alignement dans diverses applications, de façon indépendante du langage utilisé pour décrire les ontologies. Un alignement peut être converti en axiomes OWL ou bien en règles SWRL. Les éléments de l'alignement sont les suivants :

- **xml**

Cet attribut prend pour valeur ("yes"/"no"), ce qui indique si l'ali-

gnement peut être lu en tant que fichier XML et compilé avec une DTD;

– **onto1**

indique l'ontologie source du *mapping* ;

– **onto2**

indique l'ontologie cible du *mapping* ;

– **map**

Cet attribut prend pour valeur (Cell) et qui signifie une correspondance de type *mapping* ;

```
<?xml version='1.0' encoding='utf-8' standalone='no'?>
  <rdf:RDF xmlns="http://alignment#"
    xmlns:align="http://heterogeneity/alignment#"
    xmlns:rdf="http://-rdf-syntax-ns#"
    xmlns:xsd="http://XMLSchema#">
    <Alignment>
    <alignment IRI = "http://....\#"/>
      <xml>yes</xml>
      <onto1>...</onto1>
      <onto2>...</onto2>
    </Alignment>
  </rdf:RDF>
```

L'élément *ontology* contient trois attributs :

– **rdf:about** contient l'URI qui identifie l'ontologie ;

– **location** contient l'URL correspondant à la localisation de l'ontologie ;

– **formalism** décrit le nom et l'URI du langage dans lequel l'ontologie

est exprimée.

Et enfin, l'élément **Cell** identifie une correspondance de type *mapping* et il est défini par les attributs suivants :

- **entity1** contient l'URI de l'entité située à gauche de la correspondance ;
- **entity2** contient l'URI de l'entité située à droite de la correspondance ;
- **relation** contient la relation qui existe entre les deux entités. Cette relation peut être définie par les symboles suivants : > pour une relation (subsumes), < pour la relation (is subsumed), = pour (équivalent), % pour (incompatible), HasInstance et InstanceOf.

2.3 RÉSEAU D'ONTOLOGIES ALIGNÉES

Les formalismes proposés pour la représentation et le raisonnement des réseaux de connaissances alignées sont tout d'abord développés pour répondre à une motivation. Dans la prochaine section, nous présentons les plus importantes motivations résolues par les formalismes existants.

2.3.1 Motivations

- **Combinaison d'ontologies** : On parle de combinaison d'ontologies si les ontologies locales décrivent des portions séparées, complémentaires dans un domaine donné.
- **Import d'ontologies** : L'import d'ontologies est surtout utilisé pour favoriser la réutilisation des concepts, relations ou individus définis dans d'autres ontologies.
- **Mapping entre des ontologies** : Dans ce cas, les ontologies sont alignées par des correspondances sémantiques motivées par la résolution du problème d'hétérogénéité sémantique entre les entités appartenant à des ontologies différentes.

- **La décomposition d'ontologies complexes** : Il s'agit de la décomposition d'ontologie complexe en une famille d'ontologies O_1, \dots, O_n telle que leur combinaison est équivalente à l'ontologie initiale. Cette approche est différente des trois premières, et nous n'en tenons pas compte dans cette thèse.
- **La médiation entre ontologies** : Le principe de médiation consiste à ne pas relier directement les ontologies, elle est assurée par une gestion indépendante et séparée des alignements dans un niveau global, et ceci permet la composition des alignements pour un échange et une meilleure réutilisation de ces derniers entre les différentes applications.

Une fois la motivation définie, pour chaque formalisme sont présentés les composants du réseau. Les réseaux d'ontologies alignées sont composés d'ontologies locales et des alignements.

2.3.2 Composants d'un réseau d'ontologies alignées

- **Ontologie locale** : Une ontologie locale dont la représentation syntaxique et sémantique a été détaillée dans la section 2.1, est appelée aussi module ou base de connaissances source, c'est une ontologie appartenant à une famille d'ontologies à intégrer ou à combiner. C'est une famille d'ontologies $\{O_i\}$ indexées par un ensemble fini d'indices I . Chaque ontologie O_i dans le réseau est formulée dans un langage de représentation de connaissances donné et peut avoir sa propre sémantique, on peut parler alors de sémantique locale.
- **Alignement** : L'alignement représente donc la connection entre les ontologies locales, c'est une famille de correspondances A_{ij} entre une paire d'ontologies O_i et O_j (voir la section 2.2) pour le détail.

La représentation et l'exploitation des alignements représentent des points de différence des formalismes existants. Certains formalismes expriment des *mappings* d'autres des *links*. Ils sont définis selon le point de vue de l'ontologie cible ou selon un point de vue global. Ils peuvent

concerner uniquement des concepts, ou intégrer des rôles et des individus.

De même que pour une ontologie, pour raisonner sur un réseau d'ontologies alignées, il faut que ce dernier soit doté d'une sémantique formelle, la seule différence c'est que la sémantique d'un réseau dépend des sémantiques des ontologies locales et de la sémantique des alignements.

2.3.3 Sémantique du réseau d'ontologies alignées

Une sémantique formelle doit être définie pour le réseau, on distingue trois types de sémantique.

- **Sémantique non contextuelle ou centralisée** : Cette sémantique est formalisée par une logique classique et ne nécessite pas la définition d'un nouveau formalisme. Il y a un domaine d'interprétation unique pour tout le réseau qui est l'union des domaines d'interprétation locaux et de l'interprétation d'alignement. L'interprétation est un modèle si elle satisfait tous les axiomes des ontologies locales et des alignements. Cette approche est utilisée dans le cas où chaque ontologie locale et chaque alignement complètent simplement les connaissances sur un large domaine, décrivant chacun une partie d'une unique théorie globale, mais elle présente des inconvénients si les ontologies à combiner sont fortement hétérogènes avec des points de vue ou des contextes différents et incompatibles ;
- **Sémantique contextuelle ou distribuée** : Cette sémantique est formalisée par des logiques distinctes et séparées. L'interprétation distribuée du réseau est composée de l'ensemble des interprétations locales et des relations de domaine qui permettent de lier une interprétation locale à une autre interprétation du point de vue de l'interprétation cible. La sémantique contextuelle a rendu la cohabitation de plusieurs interprétations incompatibles possibles, autrement dit, le raisonnement sur un réseau d'ontologies alignées peut être cohé-

rent malgré la contradiction des points de vue des ontologies locales.

Une interprétation distribuée satisfait le réseau d'ontologies si :

- Chaque interprétation locale est un modèle de son ontologie d'origine ;
- les combinaisons d'interprétations satisfont les contraintes posées par les alignements.
- **Sémantique contextuelle et intégrée** : On peut dire que c'est la combinaison de la sémantique centralisée (du point de vue alignement) et la sémantique distribuée (du point de vue des ontologies locales). Dans ce cas les interprétations locales sont distinctes et ne sont pas reliées directement, elles sont reliées moyennant un domaine d'interprétation supplémentaire, servant à interpréter les connaissances inter-ontologies selon un point de vue global.

2.3.4 Raisonnement sur un réseau d'ontologies alignées

Le réseau doit être doté d'un algorithme qui permet de raisonner à travers l'ensemble des connaissances alignées. Il existe plusieurs disciplines de raisonnement, on trouve :

- Raisonnement centralisé non contextuel : un raisonneur unique pour un contexte unique ;
- Raisonnement centralisé contextuel : un seul raisonneur pour un raisonnement contextuel ;
- Raisonnement distribué non contextuel : ce raisonnement est caractérisé par la mise en place de plusieurs raisonneurs séparés au niveau local communiquant par des messages, soit un raisonneur par ontologie ou par une paire d'ontologies ;
- Raisonnement distribué contextuel : le raisonnement contextuel permet de raisonner sur plusieurs interprétations distinctes, soit une interprétation par ontologie locale, ce type de raisonnement est introduit par Giunchiglia (1993). Le raisonnement contextuel est appliqué dans un environnement distribué si à chaque ontologie lo-

cale, on associe un raisonneur et l'ensemble des raisonneurs peuvent communiquer par des messages.

Après avoir pris connaissances des composants, caractéristiques des réseaux de connaissances alignées, nous présentons actuellement la définition générale de ce dernier.

Définition 2.18 (Réseau d'ontologies alignées) *Un réseau d'ontologies alignées est un couple $(O_i, (A_{ij})_{i,j \in I})$ où O_i désigne un ensemble fini d'ontologies locales et A_{ij} est une famille d'alignement entre toute paire d'ontologies O_i et O_j , doté d'une sémantique formelle pour un raisonnement à travers les connaissances locales et les alignements.*

CONCLUSION DU CHAPITRE

Dans ce chapitre, nous avons présenté les réseaux d'ontologies alignées, leurs motivations, leurs composants et leurs sémantiques pour le raisonnement, tout en détaillant les notions de base, tels que les langages de représentation de connaissances, la notion de contexte, la représentation des alignements et la notion de sémantique pour le raisonnement.

ÉTAT DE L'ART : FORMALISMES DE REPRÉSENTATION DES RÉSEAUX D'ONTOLOGIES ALIGNÉES

SOMMAIRE

3.1	DIFFÉRENTS FORMALISMES DE RÉSEAUX DE CONNAISSANCES ALIGNÉES	41
3.1.1	Logiques de description distribuées (DDL)	42
3.1.2	\mathcal{E} -connection	45
3.1.3	Package Description Logic (PDL)	48
3.1.4	$E - SHIQ$	50
3.1.5	Integrated Distributed Description Logic (IDDL)	52
3.2	MODULARITÉ	56
3.3	MODÉLISATION CONTEXTUELLE DES CONNAISSANCES	57
3.4	DISCUSSION	59
3.4.1	La description et la structure des connaissances locales ne sont pas fixées	59
3.4.2	Représentation contextuelle des alignements	60
3.4.3	Plusieurs sémantiques peuvent être associées au forma- lisme proposé	60

3.4.4 Raisonement	61
CONCLUSION	61

DANS ce chapitre, nous présentons les différents travaux réalisés et liés au contexte de notre thèse. Nous nous intéressons essentiellement aux formalismes élaborés pour représenter et raisonner sur les réseaux d'ontologies alignées 3.1 tels que Distributed Description Logic (DDL), \mathcal{E} -connection, Package Description Logic (PDL), Integrated Distributed Description Logic (IDDL) et $E - SHIQ$ en présentant leurs syntaxes et leurs sémantiques pour le raisonnement. Nous nous penchons vers l'aspect de la modularité 3.2 puisque nous pouvons considérer que la structure des réseaux de connaissances multi-niveaux permet de représenter un ensemble de connaissances de façon modulaire. Chaque ontologie alignée forme un module, les alignements permettent de rassembler plusieurs modules pour former un module plus large, lui-même pouvant être réutilisé. Il nous paraît donc intéressant de comparer notre approche aux travaux existants dans le domaine des ontologies modulaires. Nous nous intéressons aussi à l'aspect de la modélisation contextuelle des connaissances 3.3. Contrairement au réseau d'ontologies alignées où les ontologies sont indépendantes, la modélisation contextuelle des connaissances considère que les contextes sont des classes ou des instances reliées dans le niveau des méta-connaissances, où, les relations entre les contextes peuvent être propagées pour relier les connaissances locales.

3.1 DIFFÉRENTS FORMALISMES DE RÉSEAUX DE CONNAISSANCES ALIGNÉES

Les réseaux de connaissances alignées sont constitués d'ontologies hétérogènes et autonomes, c'est-à-dire, développées indépendamment les unes des autres mais reliées par des alignements d'ontologies dans le but de les intégrer afin de faciliter l'échange des connaissances. Il existe plusieurs formalismes basés sur la logique spécifiant leur syntaxe et leur sémantique, pour représenter et raisonner sur un ensemble d'ontologies reliées par des correspondances. Les correspondances de type *mappings* se retrouvent dans les systèmes SomeWhere (Adjiman et al. 2005), SomeRDFS (Adjiman et al. 2007), SomeOWL (Adjiman et al. 2006)), les formalismes *Distributed Description Logic* (Borgida et Serafini 2003) et *Integrated Distributed Description Logics* (Zimmermann 2007). Le formalisme *Package-based Description Logics* (Bao et al. 2009), bien que n'utilisant pas explicitement cette notion de correspondance, peut être exprimé syntaxiquement en terme de réseau d'ontologies alignées. Les *links* apparaissent dans le formalisme \mathcal{E} -connection (Kutz et al. 2004) mais peuvent être utilisés dans des formules qui impliquent des termes de plus de deux ontologies. La combinaison des relations de type *mappings* avec des relations de type *links* dans un même formalisme a été étudiée par Vouros et Santipantakis (2012).

Certains travaux sur l'intégration des ontologies existantes et autonomes dans un environnement distribué portent sur des systèmes, dits systèmes d'inférences pair-à-pair, (SomeWhere (Adjiman et al. 2005), SomeRDFS (Adjiman et al. 2007) et SomeOWL (Adjiman et al. 2006)), constituant des réseaux de paires entre lesquelles des communications s'établissent via des correspondances de type *mappings*. La sémantique du réseau est équivalente à la sémantique non contextuelle c'est-à-dire l'interprétation du réseau est unique, comportant toutes les théories locales et celles des correspondances. Le raisonnement est distribué et s'effectue

par échange de messages entre les paires. L'objectif de ce raisonnement est d'obtenir de nouvelles conséquences sémantiques moyennant des inférences sur l'ensemble du réseau. Le principe consiste à réécrire les requêtes et propager la recherche vers les autres paires concernées par les *mappings* existants.

Les prochains travaux concernent les formalismes qui représentent des ontologies développées dans des contextes distincts et incompatibles mais alignés.

3.1.1 Logiques de description distribuées (DDL)

La famille des logiques de description distribuées (*Distributed Description Logics DDL*) est introduite par Borgida et Serafini (2002) puis développée par Serafini et Taminin (2004). La première version de DDL supporte des ontologies locales en (*SHIQ*) et des mappings entre des concepts uniquement. Serafini et Taminin (2007) ont étendu DDL avec des correspondances entre les individus ce qui permet la propagation des individus entre les différentes ontologies alignées. Ghidini et al. (2007), Ghidini et al. (2008) proposent une axiomatisation des mappings entre les rôles pour supporter la propagation de la hiérarchie de ces dernières entre les différentes ontologies. Homola (2010) dans sa thèse, a proposé l'extension de la sémantique de DDL afin de supporter la propagation de la hiérarchie des concepts complexes concernés par les mappings et prendre en charge la composition des mappings à travers la chaîne des ontologies appartenant à un réseau. Le formalisme DDL est développé dans le but de réduire l'hétérogénéité sémantique entre les ontologies d'un réseau. Il permet de relier et de raisonner sur un ensemble d'ontologies en logique de description exprimant des points de vues différents ou existant dans des contextes différents. Les ontologies sont reliées entre elles par le biais des règles passerelles (*Bridge rules*). Les passerelles en logique de description distribuée représentent des relations entre concepts, rôles et

individus d'ontologies différentes. Elles sont orientées puisqu'elles sont établies entre une ontologie source et une ontologie cible, du point de vue de l'ontologie cible.

Afin de représenter un système de connaissances distribué en DDL, une syntaxe concrète fondée sur OWL a été définie.

Syntaxe Les ontologies d'un réseau en DDL sont en logique de description. La syntaxe des passerelles qui les relie est définie comme suit :

Définition 3.1 (Passerelles) *Soient O_i et O_j deux ontologies. Une passerelle de O_i vers O_j ($i \neq j$) est une expression de l'une des formes suivantes :*

- $i:X \xrightarrow{\sqsubseteq} j:Y$ est une règle "intra" (into bridge rule) ;
- $i:X \xrightarrow{\sqsupseteq} j:Y$ est une règle "extra" (onto bridge rule) ;
- $i:a \xrightarrow{\equiv} j:b$ est une règle d'individu (individual correspondence)

où X (respectivement Y) est soit un concept, soit un rôle de O_i (respectivement O_j) et a (respectivement b) est un individu de O_i (respectivement O_j). La combinaison de deux passerelles $i:X \xrightarrow{\sqsubseteq} j:Y$ et $i:X \xrightarrow{\sqsupseteq} j:Y$ dénote une correspondance d'équivalence $i:X \xrightarrow{\equiv} j:Y$

Un système de connaissances en DDL est composé de plusieurs ontologies locales indicées et des passerelles qui expriment les *mappings* entre les entités appartenant à ces ontologies.

Définition 3.2 (Système distribué en DDL) *Un système distribué en DDL est une paire $\langle (O_i)_{i \in [1,n]}, (\mathcal{B}_{ij})_{i,j \in [1,n], i \neq j} \rangle$ où (O_i) est une famille d'ontologies locales et \mathcal{B}_{ij} est un ensemble de passerelles entre O_i et O_j pour tout $i, j \in [1, n]$ et $i \neq j$.¹*

Sémantique et raisonnement On interprète un système distribué en DDL en affectant à chaque ontologie une interprétation locale en logique de description, puis en définissant des relations entre domaines d'interprétation d'ontologies différentes. Une relation de domaine r_{ij} représente une relation ensembliste entre un domaine d'interprétation de l'ontolo-

1. Dans la suite de la thèse, pour plus de lisibilité, nous omettrons l'ensemble dans lequel évolue les indices en écrivant par exemple (O_i) à la place de $(O_i)_{i \in [1,n]}$.

gie O_i et un domaine d'interprétation de l'ontologie O_j . Un élément du domaine d'interprétation de O_i peut être relié à plusieurs éléments du domaine de O_j et vice versa. L'ensemble des interprétations locales et les relations de domaine constituent l'interprétation distribuée.

Définition 3.3 (Interprétation distribuée) *Soit un système distribué en DDL $S = \langle \{O_i\}, \{B_{ij}\} \rangle$.*

Une interprétation distribuée de S est une paire $\mathcal{I} = \langle \{I_i\}, \{r_{ij}\} \rangle$, avec pour tout $i \in [1, n]$, $I_i = \langle \Delta^{I_i}, \cdot^{I_i} \rangle$ est une interprétation de l'ontologie O_i et pour tout $i, j \in [1, n]$ et $i \neq j$, $r_{ij} \subseteq \Delta^{I_i} \times \Delta^{I_j}$ est une relation de domaine.

L'interprétation distribuée est un modèle du réseau d'ontologies en DDL, si chaque interprétation locale est un modèle de son ontologie associée, et si la relation de domaine satisfait les passerelles.

Définition 3.4 (Relation de satisfaction) *Soit l'interprétation distribuée $\mathcal{I} = \langle \{I_i\}, \{r_{ij}\} \rangle$. On*

définit la relation de satisfaction \models_d en DDL de la façon suivante :

- *si α est un axiome ou une assertion en logique de description alors $\mathcal{I} \models_d i : \alpha$ si et seulement si $I_i \models \alpha$;*
- *si X et Y sont deux concepts ou deux rôles alors $I \models_d i : X \sqsubseteq j : Y$ si et seulement si $r_{ij}(X^{I_i}) \subseteq Y^{I_j}$;*
- *si X et Y sont deux concepts ou deux rôles alors $I \models_d i : X \sqsupseteq j : Y$ si et seulement si $r_{ij}(X^{I_i}) \supseteq Y^{I_j}$;*
- *si a et b sont deux individus alors $I \models_d i : a \bar{\rightarrow} j : b$ si et seulement si $r_{ij}(a^{I_i}) = \{b^{I_j}\}$;*
- *si B_{ij} est un ensemble de règles passerelles alors $\mathcal{I} \models_d B_{ij}$ si et seulement si pour tout $b \in B_{ij}$, $\mathcal{I} \models_d b$.*

Cette définition nous permet de définir la notion de modèle d'un système distribué.

Définition 3.5 (Modèle distribué) *Soit l'interprétation distribuée $\mathcal{I} = \langle \{I_i\}, \{r_{ij}\} \rangle$ et un système*

distribué $S = \langle \{O_i\}, \{B_{ij}\} \rangle$. On dit que \mathcal{I} est un modèle distribué de S si et seulement si :

- *pour tout $i \in [1, n]$ et tout axiome α dans O_i , $\mathcal{I} \models_d i : \alpha$;*

- pour tout $i, j \in [1, n], i \neq j, \mathcal{I} \models B_{ij}$.

Ces formalismes sont dotés d'un algorithme introduit par Taminin (2007) pour un raisonnement distribué sur un ensemble d'ontologies exprimées en \mathcal{SHIQ} DL qui a été implémenté en un prototype intitulé Drago (Serafini et Taminin 2005) et les passerelles (bridges rule) sont exprimées dans le format d'alignement C-OWL 2.2.3.

3.1.2 \mathcal{E} -connection

\mathcal{E} -connection (Kutz et al. 2004), est un formalisme qui combine plusieurs logiques où chaque logique représente une portion d'un système complexe. Il est défini à travers un système de description abstrait (Abstract Description System) qui est une généralisation commune des différentes logiques telles que la logique de description, la logique modale, la logique de temps et d'espace et d'autres formalismes logiques. Les ontologies locales sont combinées en utilisant des *links* qui peuvent être assimilés à des inter-ontologies roles. Pour définir sa syntaxe, on va s'intéresser au formalisme \mathcal{E} -connection dans les logiques de description.

Syntaxe La syntaxe de \mathcal{E} -connection dans les logiques $\mathcal{C}_{\mathcal{HIQ}+}^{\mathcal{E}}(\mathcal{SHOIQ}, \mathcal{SHOQ}, \mathcal{SHIO})$ considère qu'il existe un ensemble de bases de connaissances à connecter où chacune d'elle est formalisée dans une des logiques \mathcal{SHOIQ} ou \mathcal{SHOQ} ou \mathcal{SHIO} . Le vocabulaire de \mathcal{E} -connection est un quadruplet $\langle N_{C_i}^{mi}, N_{R_i}^{mi}, N_{L_{ij}}^{mi}, N_{I_i}^{mi} \rangle$ avec $N_{C_i}^{mi}$ est un ensemble de i-concepts (i-concepts est une notation donnée aux concepts appartenant à une même ontologie O_i), $N_{R_i}^{mi}$ est un ensemble de i-roles (i-roles est une notation donnée aux roles appartenant à la même ontologie O_i), $N_{I_i}^{mi}$ est un ensemble d'individus appartenant à l'ontologie O_i et $N_{L_{ij}}^{mi}$ est un ensemble de ij-roles (ij-roles est une notation donnée aux roles liant une entité appartenant à une ontologie O_i à une entité appartenant à une ontologie O_j) où $i, j \in I$ et

I est un ensemble non fini d'index et la variable m_i représente la logique de l'ontologie locale où $m_i = SHOIQ/SHOQ/SHIO$.

La base de connaissance combinée $\Sigma = \langle \mathcal{K}, \mathcal{R}, \mathcal{L}, \mathcal{A} \rangle$ où :

$\mathcal{K} = \{\mathcal{K}_i\}_{i \in I}$ est l'ensemble des TBox locaux appelés TBox combinés, contenant tous les axiomes sur les concepts tels que les axiomes de subsumption de concepts ($C \sqsubseteq D$)

$\mathcal{R} = \{\mathcal{R}_i\}_{i \in I}$ est l'ensemble des RBox locaux, contenant tous les axiomes de i -role pour $i \in I$ tel que i -role-subsumption ($P \sqsubseteq S$) et les axiomes de transitivité définis sur i -role.

$\mathcal{L} = \{\mathcal{L}_{ij}\}_{i,j \in I}$ est LBox, contenant tous les axiomes de ij -role pour $i, j \in I$ tel que ij -role-subsumption ($Q \sqsubseteq T$) et les axiomes de transitivité définis sur ij -role.

$\mathcal{A} = \{\mathcal{A}_i\}_{i \in I} \cup \{\mathcal{A}_{ij}\}_{ij}$ est le ABox combiné, $\{\mathcal{A}_i\}$ contient les assertions locales $C(a)$ avec $a \in N_{C_i}^{m_i}$ est $\{\mathcal{A}_{ij}\}_{ij}$ contient l'assertion de ij -role dont la forme est la suivante :

- $R(a, b)$ si $R \in N_{R_i}^{m_i}$ et $a, b \in N_{I_i}$;
- $a.P.b$ si $i \neq j$ avec $a \in N_{I_i}, b \in N_{I_j}$ et $P \in N_{L_{ij}}^{m_i}$.

Sémantique et raisonnement La sémantique de \mathcal{E} -connection dans $\mathcal{C}_{\mathcal{HIQ}^+}^{\mathcal{E}}(SHOIQ, SHOQ, SHIO)$ se traduit par une interprétation combinée \mathcal{I} pour une base de connaissances combinées $\Sigma = \langle \mathcal{K}, \mathcal{R}, \mathcal{L}, \mathcal{A} \rangle$. $\mathcal{I} = \langle \{\Delta^{\mathcal{I}_i}\}_{i \in I}, \{\mathcal{I}_j\}_{i \in I}, \{\mathcal{I}_{ij}\}_{i,j \in I} \rangle$ pour tout $i, j \in I$, et si $i \neq j$ on a $\Delta^{\mathcal{I}_i} \cap \Delta^{\mathcal{I}_j} = \emptyset$ où $\Delta^{\mathcal{I}_i}$ est une famille des domaines d'interprétation locaux, \mathcal{I}_i représente la fonction d'interprétation locale et \mathcal{I}_{ij} est la fonction inter-interprétations où chaque ij -role $P \in \mathcal{L}_{ij}$ est interprété par $P^{\mathcal{I}_{ij}}$ qui est un sous ensemble de $\Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$.

- L'interprétation combinée \mathcal{I} satisfait le TBox combiné noté par ($\mathcal{I} \models \mathcal{K}$) si \mathcal{I} satisfait chaque TBox local noté ($\mathcal{I} \models \{\mathcal{K}_i\}$) et $\mathcal{I} \models \{\mathcal{K}_i\}$ si \mathcal{I} satisfait chaque axiome dans $\{\mathcal{K}_i\}$ comme $\mathcal{I} \models C \sqsubseteq D$ et l'interprétation $\mathcal{I} \models C \sqsubseteq D$ si $C^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$. En \mathcal{E} -connection, la définition

des i -concepts complexe doit satisfaire les contraintes définies dans le tableau 3.1.

- L'interprétation \mathcal{I} satisfait i -role Box combiné noté par $(\mathcal{I} \models \mathcal{R})$ si $\mathcal{I} \models \{\mathcal{R}_i\}$ et $\mathcal{I} \models \{\mathcal{R}_i\}$ si \mathcal{I} satisfait chaque i -role-axiome comme $\mathcal{I} \models P \sqsubseteq S$ et l'interprétation $\mathcal{I} \models P \sqsubseteq S$ si $P^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$. L'interprétation \mathcal{I} doit satisfaire aussi les axiomes de transitivité définis sur les i -roles.
- L'interprétation \mathcal{I} satisfait ij -role Box noté par $(\mathcal{I} \models \mathcal{L})$ si $\mathcal{I} \models \{\mathcal{L}_{ij}\}$ et $\mathcal{I} \models \{\mathcal{L}_{ij}\}$ si \mathcal{I} satisfait chaque ij -role-axiome par exemple $\mathcal{I} \models P \sqsubseteq S$ et $\mathcal{I} \models P \sqsubseteq S$ si $P^{\mathcal{I}_{ij}} \subseteq S^{\mathcal{I}_{ij}}$ où P, S sont des ij -roles ou *links*. L'interprétation \mathcal{I} doit satisfaire aussi les axiomes de transitivité définis sur les *links*.
- L'interprétation combinée \mathcal{I} satisfait ABox combiné noté par $(\mathcal{I} \models \mathcal{A})$ dans les cas suivants :
 - $\mathcal{I} \models C(a)$ si $a^{\mathcal{I}_i} \in C^{\mathcal{I}_i}$;
 - $\mathcal{I} \models R(a, b)$ si $(a^{\mathcal{I}_i}, b^{\mathcal{I}_i}) \in R^{\mathcal{I}_i}$;
 - $\mathcal{I} \models a.P.b$ si $(a^{\mathcal{I}_i}, b^{\mathcal{I}_j}) \in P^{\mathcal{I}_{ij}}$;
 où C est un concept, R est un i -role et P est un *link*.

Syntaxe	Sémantique
$\exists P.Z$	$(\exists P.Z)^{\mathcal{I}_i} = \{x \in \Delta^{\mathcal{I}_i} \mid \exists y \in \Delta^{\mathcal{I}_j}(x, y) \in P^{\mathcal{I}_{ij}} \text{ et } y \in Z^{\mathcal{I}_j}\}$
$\forall P.Z$	$(\forall P.Z)^{\mathcal{I}_i} = \{x \in \Delta^{\mathcal{I}_i} \mid \forall y \in \Delta^{\mathcal{I}_j}(x, y) \in P^{\mathcal{I}_{ij}} \Rightarrow y \in Z^{\mathcal{I}_j}\}$
$\geq nS.Z$	$(\geq S.Z)^{\mathcal{I}_i} = \{x \in \Delta^{\mathcal{I}_i} \mid \#\{\exists y \in \Delta^{\mathcal{I}_j}(x, y) \in S^{\mathcal{I}_{ij}}\} \geq n\}$
$\leq nS.Z$	$(\leq S.Z)^{\mathcal{I}_i} = \{x \in \Delta^{\mathcal{I}_i} \mid \#\{\exists y \in \Delta^{\mathcal{I}_j}(x, y) \in S^{\mathcal{I}_{ij}}\} \leq n\}$

TABLE 3.1 – Sémantique des concepts complexes définis à partir *links* en \mathcal{E} -connection

L'algorithme de tableaux pour le raisonnement sur \mathcal{E} -connection adoptant les langages $\mathcal{C}_{\mathcal{H}Q}^{\mathcal{E}}(\mathcal{SHOIQ}, \mathcal{SHOQ}, \mathcal{SHIO})$, $\mathcal{C}_{\mathcal{H}I}^{\mathcal{E}}(\mathcal{SHOIQ}, \mathcal{SHOQ}, \mathcal{SHIO})$, $\mathcal{C}_{\mathcal{H}Q+}^{\mathcal{E}}(\mathcal{SHOIQ}, \mathcal{SHOQ}, \mathcal{SHIO})$ et $\mathcal{C}_{\mathcal{H}I+}^{\mathcal{E}}(\mathcal{SHOIQ}, \mathcal{SHOQ}, \mathcal{SHIO})$ est introduit dans Parsia et Cuenca-Grau (2005) et l'extension du raisonneur Pellet supporte \mathcal{E} -connection dans les logiques de description.

3.1.3 Package Description Logic (PDL)

Le formalisme P-DL (Bao et al. 2009) motivé par l'inconvénient de l'importation des ontologies en totalité car OWL `:import` permet la réutilisation des connaissances à travers l'importation des ontologies. Chaque ontologie est capable de réutiliser une autre ontologie par une importation totale en référençant son URL, l'importation d'une partie de l'ontologie ou quelques concepts n'est pas tolérée. P-DL *SHOIQP* permet l'import des concepts, roles ou individus entre des ontologies formalisées en *SHOIQ*. Ce formalisme supporte une sémantique contextualisée.

Syntaxe Un package en *SHOIQP* regroupe *SHOIQ* TBox et RBox. La signature d'un package P_i notée $Sig(P_i)$ est un ensemble de concept NC_i , un ensemble de roles NR_i et un ensemble d'individus NI_i utilisés dans le package P_i . La signature d'un package $Sig(P_i)$ inclut aussi l'ensemble de role inverse $\bar{NR}_i = NR_i \cup \{R^- | R \in NR_i\}$ où R^- est le role inverse du role R . Elle est divisée en deux parties, on trouve : la signature locale $Loc(P_i)$ et la signature externe $Ext(P_i)$. $Trans_i(R)$ dénote que $R \in NR_i$ est transitive dans P_i . L'expression sous la forme $R \sqsubseteq S$, où R et S sont des i-roles dénote l'axiome de subsumption entre les roles dans P_i . L'ensemble des concepts en *SHOIQP* dans P_i est défini comme suit : $C := A | o | \neg_k C | C \cap C | C \cup C | \forall R.C | \exists R.C | (\leq nS.C) | (\geq nS.C)$ où $A \in NC_i$, $o \in NI_i$, n est un entier non nul, $R, S \in \bar{NR}_i$, $\neg_k C$ dénote la négation d'un concept contextualisé C dans P_k .

Sémantique et raisonnement Chaque package a son propre domaine d'interprétation. Pour une base de connaissances basée sur les packages, on a $\Sigma = \{P_i\}$, l'interprétation distribuée est un tuple $\mathcal{I} = \langle \{\mathcal{I}_i\}, \{r_{ij}\}_{P_i \in \mathcal{P}^+} \rangle$, où \mathcal{I}_i est l'interprétation locale du package P_i , avec le domaine $\Delta^{\mathcal{I}_i}$, $r_{ij} \subset \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$ est la relation de domaine qui représente l'image du terme importé dans le package cible P_j tel que :

- l'image d'un concept A est définie par $r_{ij}(A) = \{y \in \Delta^{\mathcal{I}_j} \mid \exists x \in A \text{ et le couple } (x, y) \in r_{ij}\}$;
- l'image d'un role S est définie par $r_{ij}(S) = \{(z, w) \in \Delta^{\mathcal{I}_j} \times \Delta^{\mathcal{I}_j} \mid \exists (x, y) \in S, (x, z) \in r_{ij} \wedge (y, w) \in r_{ij}\}$;

Chaque interprétation locale $\mathcal{I}_i = \langle \Delta^{\mathcal{I}_i}, \mathcal{I}_i \rangle$ consiste en un domaine $\Delta^{\mathcal{I}_i}$ et une fonction d'interprétation \mathcal{I}_i , qui permet d'assigner à chaque concept un sous ensemble de $\Delta^{\mathcal{I}_i}$, et à chaque role un sous ensemble $\Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_i}$ et à chaque individu un élément dans $\Delta^{\mathcal{I}_i}$. La fonction d'interprétation \mathcal{I} satisfait les équations suivantes avec R,S sont des j-roles, C,D sont des concepts :

$$\begin{aligned}
R^{\mathcal{I}_i} &= (R^{\mathcal{I}_i})^+, \text{ si } \text{Trans}_i(R) \in \mathcal{R}_i \\
(R^-)^{\mathcal{I}_i} &= \{(x, y) \mid (y, x) \in R^{\mathcal{I}_i}\} \\
(C \cap D)^{\mathcal{I}_i} &= C^{\mathcal{I}_i} \cap D^{\mathcal{I}_i} \\
(C \sqcup D)^{\mathcal{I}_i} &= C^{\mathcal{I}_i} \cup D^{\mathcal{I}_i} \\
(\neg_j C)^{\mathcal{I}_i} &= r_{ji}(\Delta^{\mathcal{I}_j}) \setminus C^{\mathcal{I}_i} \\
(\exists R.C)^{\mathcal{I}_i} &= \{x \in r_{ji}(\Delta^{\mathcal{I}_j}) \mid \exists y \in \Delta^{\mathcal{I}_i}, (x, y) \in R^{\mathcal{I}_i} \wedge y \in C^{\mathcal{I}_i}\} \\
(\forall R.C)^{\mathcal{I}_i} &= \{x \in r_{ji}(\Delta^{\mathcal{I}_j}) \mid \forall y \in \Delta^{\mathcal{I}_i}, (x, y) \in R^{\mathcal{I}_i} \longrightarrow y \in C^{\mathcal{I}_i}\} \\
(\geq nS.C)^{\mathcal{I}_i} &= \{x \in r_{ji}(\Delta^{\mathcal{I}_j}) \parallel y \in \Delta^{\mathcal{I}_i} \mid (x, y) \in S^{\mathcal{I}_i} \wedge y \in C^{\mathcal{I}_i}\} \mid \geq n\} \\
(\leq nS.C)^{\mathcal{I}_i} &= \{x \in r_{ji}(\Delta^{\mathcal{I}_j}) \parallel y \in \Delta^{\mathcal{I}_i} \mid (x, y) \in S^{\mathcal{I}_i} \wedge y \in C^{\mathcal{I}_i}\} \mid \leq n\}
\end{aligned}$$

L'interprétation $\mathcal{I} = \langle \{\mathcal{I}_i\}, \{r_{ij}\}_{p_i \in P_i^+} \rangle$ est un modèle de la base de connaissances en \mathcal{SHOIQP} ($\Sigma = \{P_i\}$), notée par $\mathcal{I} \models \Sigma$, si $\cup_i \Delta^{\mathcal{I}_i} \neq \emptyset$, i.e., le domaine d'interprétation local est non vide et les conditions suivantes sont satisfaites :

1. Consistance : pour tout i, j, k appartenant à un ensemble d'indices non vide on a $P_i \in P_k^*$ et $P_k \in P_j^*$, on a $r_{ij} = r_{kj} \circ r_{ik}$;
2. pour chaque i-concept C qui apparait dans P_j , on a $r_{ij}(C^{\mathcal{I}_i}) = C^{\mathcal{I}_j}$.
3. pour chaque i-role R qui apparait dans P_j , on a $R^{\mathcal{I}_i} = r_{ij}(R^{\mathcal{I}_j})$.
4. préservation de cardinalité pour les roles : pour chaque i-role R qui apparait dans P_j et chaque $(x, x') \in r_{ij}, y \in R^{\mathcal{I}_i}(x)$ si et seulement si $r_{ij}(y) \in R^{\mathcal{I}_j}(x')$.

5. pour chaque i -nominal o qui apparaît dans P_j , $(o^{\mathcal{I}_i}, o^{\mathcal{I}_j}) \in r_{ij}$.
6. $\mathcal{I}_i \models P_i$, pour chaque i .

3.1.4 $E - SHIQ$

Le formalisme de (Vouros et Santipantakis 2012) est développé dans le but de combiner et de raisonner sur un ensemble d'ontologies hétérogènes et complémentaires, ceci en combinant les correspondances de type *mappings* et les correspondances de type *links* au niveau des alignements entre ontologies, mais ces dernières sont toujours exprimées selon le point de vue de l'ontologie cible. Une extension du format d'alignement C-owl a été effectué pour stocker les links. Côté sémantique, les auteurs combinent entre la sémantique de \mathcal{E} -connection et celle de DDL et ils adoptent une sémantique distribuée avec un raisonnement distribué en combinant les raisonnements locaux avec la possibilité de propagation des subsumptions entre les différentes ontologies. Mais ils ne tiennent pas compte de l'hétérogénéité sémantique des links au niveau des alignements.

Syntaxe Soient un ensemble d'ontologies indicées par un ensemble non vide d'indices noté I . On considère N_{C_i} , N_{R_i} et N_{o_i} les ensembles de concepts, rôles et individus respectivement et appartenant à l'ontologie O_i . La base de connaissances distribuée dans le formalisme $E - SHIQ$ est définie comme suit :

Définition 3.6 (Base de connaissances distribuée) Une base de connaissances distribuée Σ est un triplet $\langle T, R, C \rangle$ où :

- T est un TBox distribué, $T = (\mathcal{T}_i)_{i \in I}$ où T est l'ensemble des axiomes des i -concepts (les concepts appartenant à l'ontologie O_i) tel que :
 $i:C \sqsubseteq i:D$ ou $(i:C \sqsubseteq D)$;
- R est le RBox distribué, $R = (\mathcal{R}_{ij})_{i,j \in I}$ où R est l'ensemble des axiomes des i -rôles (les rôles appartenant à l'ontologie O_i) dans le cas où $i=j$ et des axiomes sur ij -property (links) dans le cas où $i \neq j$;

- C est l'ensemble des correspondances (uniquement les mappings) dont la forme suivante : $i:C \xrightarrow{\Xi} j:D$;
- (DAB) Abox distribué inclut \mathcal{A}_i (les ABox locaux) et les \mathcal{A}_{ij} contenant les assertions concernant les rôles locaux, les links et les correspondances.

On voit bien que pour les *links*, les auteurs de ce formalisme optent pour la syntaxe adoptée pour \mathcal{E} -connection, contrairement au réseau de connaissances multi-niveaux où les *links* représentent le vocabulaire des alignements qui est indépendant des ontologies locales.

Sémantique

- Chaque TBox local \mathcal{T}_i est localement interprété par une interprétation locale \mathcal{I}_i qui est composée du domaine d'interprétation Δ_i et de la fonction $\cdot^{\mathcal{I}_i}$ qui assigne à chaque concept un sous-ensemble de Δ_i ;
- \mathcal{R}_{ij} est interprété par la fonction $\cdot^{\mathcal{I}_{ij}}$ qui assigne à chaque ij-property un sous-ensembl $\Delta_i \times \Delta_j$ avec $\mathcal{I}_{ij} = (\Delta_i, \Delta_j, \cdot^{\mathcal{I}_{ij}})$ pour $i \neq j \in I$ et dans le cas où $i = j$ on a $\mathcal{I}_{ii} = (\Delta_i, \Delta_{ii}, \cdot^{\mathcal{I}_{ii}}) = (\Delta_i, \Delta_i, \cdot^{\mathcal{I}_i}) = \mathcal{I}_i$, c'est le cas de l'interprétation des i-roles.

Pour l'interprétation des mappings, comme en DDL, on utilise la relation de domaine pour lier une entité d'une ontologie à l'entité correspondante de l'autre ontologie.

Définition 3.7 (relation de domaine) *la relation de domaine r_{ij} de Δ_i à Δ_j pour $i \neq j \in I$, représente le lien entre les individus du point de vue j et c'est un sous-ensemble de $\Delta_i \times \Delta_j$. Soit D un sous-ensemble de Δ_i , on a $r_{ij}(D)$ dénote $\bigcup_{d \in D} r_{ij}(d)$.*

L'interprétation distribuée de la base de connaissances dans $E - \mathcal{SHIQ}$ est donc un ensemble formé par une famille d'interprétations combinées et des relations de domaines liant ces interprétations.

Définition 3.8 (Interprétation distribuée) *Soit I un ensemble d'indices et $i, j \in I$, l'interprétation distribuée \mathcal{I} de la base de connaissances Σ est $\mathcal{I} = \langle (\mathcal{I}_{ij})_{i,j \in I}, (r_{ij})_{i \neq j \in I} \rangle$ où :*

- $\mathcal{I}_{ij} = \langle \Delta_i, \Delta_j, \cdot^{\mathcal{I}_{ij}} \rangle$ $i, j \in I$;
- r_{ij} une relation de domaine pour $i \neq j$.

On voit bien que c'est la combinaison des interprétations d'alignement déjà utilisées par \mathcal{E} -connection pour les links et DDL pour les mapping.

L'interprétation distribuée \mathcal{I} satisfait (\models_d) les éléments de la base de connaissances distribuée comme suit :

1. $\mathcal{I} \models_d i : C \sqsubseteq D$, si $\mathcal{I}_i \models C \sqsubseteq D$;
2. $\mathcal{I} \models_d \mathcal{T}_i$ si $\mathcal{I} \models i : C \sqsubseteq D$ pour tout axiome $C \sqsubseteq D$ dans \mathcal{T}_i ;
3. $\mathcal{I} \models_d i:C \stackrel{\sqsubseteq}{\mapsto} j:D$ si $r_{ij}(C^{\mathcal{I}_i}) \subseteq D^{\mathcal{I}_j}$;
4. $\mathcal{I} \models_d i:C \stackrel{\supseteq}{\mapsto} j:D$, si $r_{ij}(C^{\mathcal{I}_i}) \supseteq D^{\mathcal{I}_j}$;
5. $\mathcal{I} \models_d C_{ij}$, si \mathcal{I} satisfait toutes les correspondances de C_{ij} ;
6. $\mathcal{I} \models_d R \sqsubseteq S$ si $\mathcal{I}_{ij} \models R \sqsubseteq S$ où $R \sqsubseteq S$ appartient à \mathcal{R}_{ij} ;
7. $\mathcal{I} \models_d \mathcal{R}_{ij}$ si $\mathcal{I} \models_d R \sqsubseteq S$ et si \mathcal{I} satisfait tous les axiomes de transitivité des i-roles et des links appartenant à \mathcal{R}_{ij} ;
8. $\mathcal{I} \models_d \Sigma$ si pour tout $i, j \in I$, $\mathcal{I} \models_d \mathcal{T}_i$, $\mathcal{I} \models_d \mathcal{R}_{ij}$, $\mathcal{I} \models_d \mathcal{A}_i$ et $\mathcal{I} \models_d C_{ij}$;
9. $\mathcal{I} \models_d i:a \stackrel{\sqsubseteq}{\mapsto} j:b$ si $b^j \in r_{ij}(a^i)$, $i \neq j \in I$.

$E - SHIQ$ étend l'algorithme de tableau de la logique de description qui a été implémenté dans le raisonneur Pellet.

3.1.5 Integrated Distributed Description Logic (IDDL)

Les logiques de description distribuées intégrées ont été présentées par Zimmermann (2007). Dans IDDL, les ontologies locales sont représentées en logique de description, et les alignements sont des correspondances bien particulières reliant soit des concepts, soit des roles, soit des individus de deux ontologies dans un point de vue global.

Syntaxe Les correspondances en IDDL sont similaires aux passerelles (bridge rule) en DDL sauf que celles-ci sont définies selon le point de vue général, ceci peut être expliqué de la façon suivante : si on a une correspondance de type $i:A \stackrel{\sqsubseteq}{\mapsto} j:B$ où A est un concept ou role de l'ontologie

O_i et B un concept ou rôle de l'ontologie O_j alors on peut affirmer la correspondance suivante $j:B \xleftrightarrow{\exists} i:A$. Ce n'est pas le cas pour DDL car elles sont définies selon le point de vue de l'ontologie cible.

Définition 3.9 (Correspondance) *Une correspondance entre les ontologies i et j est l'une des formules suivantes :*

- $i:C \xleftrightarrow{\sqsubseteq} j:D$ est une subsumption de concepts inter-ontologies ;
- $i:R \xleftrightarrow{\sqsubseteq} j:S$ est une subsumption de rôles inter-ontologies ;
- $i:C \xleftrightarrow{\perp} j:D$ est une exclusion de concepts inter-ontologies ;
- $i:R \xleftrightarrow{\perp} j:S$ est une exclusion de rôles inter-ontologies ;
- $i:a \xleftrightarrow{\in} j:C$ est une appartenance d'individus inter-ontologies ;
- $i:a \xleftrightarrow{=} j:b$ identité d'individus inter-ontologies.

Définition 3.10 (Réseau d'ontologies alignées) *Un réseau d'ontologies alignées est un couple (O_i, A_{ij}) telle que O_i est un ensemble d'ontologies et A_{ij} est une famille de correspondances reliant les ontologies O_i pour $i \in [1..n]$.*

Sémantique et raisonnement La sémantique des réseaux d'ontologies alignées en IDDL dépend des sémantiques locales. Un réseau en IDDL peut comprendre des ontologies en *ALC*, *SHIQ* ou *SROIQ*. Informellement, interpréter un réseau IDDL consiste à corrélérer les domaines d'interprétation grâce à la fonction d'égalisation. Cette dernière permet de projeter chaque entité appartenant à une correspondance dans un domaine d'interprétation global.

Définition 3.11 (Fonction d'égalisation) *Etant donnée une famille d'interprétations locales \mathcal{I}_i , une famille de fonctions d'égalisation ϵ_i telle que pour tout $\mathcal{I}_i, \epsilon_i : \Delta^{\mathcal{I}_i} \rightarrow \Delta$ où Δ_i est un domaine d'interprétation local, Δ est appelé le domaine global d'interprétation et $i \in [1..n]$.*

Dans IDDL, une interprétation distribuée assigne une interprétation de logique de description à chaque ontologie du réseau, ainsi qu'une fonction d'égalisation qui met en corrélation les connaissances locales dans un domaine d'interprétation global.

Définition 3.12 (Interprétation distribuée) Soit $S = \langle O_i, A_{ij} \rangle$ un réseau d'ontologies alignées.

Une interprétation distribuée de S est une paire $\langle \mathcal{I}_i, \epsilon_i \rangle$ où \mathcal{I}_i est une famille d'interprétation locales, ϵ_i est une famille de fonctions d'égalisation, tels que pour tout $i \in [1..n]$.

L'image d'un concept ou d'un role dans un domaine global à travers la fonction d'égalisation est définie de la façon suivante.

Définition 3.13 (Fonction d'égalisation de concepts et de roles) Etant donnée une interprétation distribuée $\langle \mathcal{I}_i, \epsilon_i \rangle$ d'un réseau d'ontologies S , la fonction d'égalisation est étendue pour définir les images des concepts et des roles :

- $\epsilon_i(C_i^{I_i}) = \bigcup_{x \in C_i^{I_i}} \{\epsilon_i(x)\}$ et
- $\epsilon_i(R_i^{I_i}) = \bigcup_{(x,y) \in R_i^{I_i}} \{\epsilon_i(x), \epsilon_i(y)\}$.

La notion de satisfiabilité locale est identique à celle des logiques de description et la satisfiabilité des correspondances met en jeu les fonctions d'égalisation.

Définition 3.14 (Satisfaction d'une correspondance) Soit S un réseau d'ontologies alignées et O_i, O_j deux ontologies de S . Soit $\mathcal{I} = \langle \mathcal{I}_i, \epsilon_i \rangle$ une interprétation distribuée. On définit la satisfaction d'une correspondance c (ce que l'on note $\mathcal{I} \models_d c$) comme suit :

$$\mathcal{I} \models_d i:C \xrightarrow{\sqsubseteq} j:D \text{ si et seulement si } \epsilon_i(C^{I_i}) \subseteq \epsilon_j(D^{I_j})$$

$$\mathcal{I} \models_d i:R \xrightarrow{\sqsubseteq} j:S \text{ si et seulement si } \epsilon_i(R^{I_i}) \subseteq \epsilon_j(S^{I_j})$$

$$\mathcal{I} \models_d i:C \xrightarrow{\perp} j:D \text{ si et seulement si } \epsilon_i(C^{I_i}) \cap \epsilon_j(D^{I_j}) = \emptyset$$

$$\mathcal{I} \models_d i:R \xrightarrow{\perp} j:S \text{ si et seulement si } \epsilon_i(R^{I_i}) \cap \epsilon_j(S^{I_j}) = \emptyset$$

$$\mathcal{I} \models_d i:a \xrightarrow{\in} j:C \text{ si et seulement si } \epsilon_i(a^{I_i}) \in \epsilon_j(C^{I_j})$$

$$\mathcal{I} \models_d i:a \xrightarrow{=} j:b \text{ si et seulement si } \epsilon_i(a^{I_i}) = \epsilon_j(b^{I_j})$$

Et pour tout axiome local $i:\alpha$, $\mathcal{I} \models_d i:\alpha$ si et seulement si $\mathcal{I}_i \models \alpha$. Autrement dit, la satisfaction locale implique la satisfaction globale. Une interprétation distribuée \mathcal{I} satisfait un alignement A_{ij} si et seulement si elle satisfait toutes les correspondances de A_{ij} et elle satisfait une ontolo-

gie O_i si et seulement si elle satisfait tous les axiomes de O_i . Quand une interprétation distribuée satisfait toutes les ontologies et les alignements, on l'appelle modèle du réseau.

Définition 3.15 (Modèle d'un réseau d'ontologies alignées) Soit $S = \langle O_i, A_{ij} \rangle$ un réseau d'ontologies alignées. Une interprétation distribuée \mathcal{I} est un modèle de S (ce que l'on note $\mathcal{I} \models_d S$), si et seulement si :

- pour tout $O_i, \mathcal{I} \models_d O_i$;
- pour tout $A_{ij}, \mathcal{I} \models_d A_{ij}$.

La procédure de raisonnement dans IDDL est développée pour vérifier si un réseau d'ontologies $S = \langle O_i, A_{ij} \rangle$ est cohérent ou non, mais les correspondances sont limitées à la subsomption et l'exclusion de concepts et la subsomption de role. Elle permet d'exploiter des systèmes de raisonnement locaux associés aux ontologies locales tels que Racer ou Pellet. Il existe un système de raisonnement supplémentaire qui accède à l'ensemble des alignements, il n'accède pas au contenu des ontologies, par contre, il communique avec les systèmes de raisonnement locaux en leur demandant si certains axiomes sont cohérents vis-à-vis de leur ontologie. Le test de cohérence repose sur la propagation des axiomes des niveaux locaux au niveau global et vice versa.

DRAOn (Duc et al. 2013) est le raisonneur développé pour le raisonnement sur une base de connaissances distribuées en IDDL.

Tous les formalismes décrits auparavant sont fondés sur les logiques de description. D'autres logiques locales sont utilisés pour formaliser des réseaux de connaissances alignées. Nous citons comme exemple le travail de Ghidini et Serafini (2000) qui utilise *Distributed First Order Logic* et la logique du première ordre comme logique locale, complétée de règles passerelles généralisant celles de DDL. Joseph et al. (2014) proposent le formalisme RDF (Cyganiak et al. 2014) qui représente un réseau de connais-

sances comme un ensemble de quadruplets (c, s, p, o) où c est un identifiant du contexte du triplet (s, p, o) (c'est-à-dire, de l'ontologie dans laquelle il apparait). Des passerelles sont définies entre les quadruplets de contextes différents afin de pouvoir raisonner sur les connaissances à travers les contextes. Ce formalisme est également accompagné d'une procédure de raisonnement utilisant une technique de chaînage avant ne fonctionnant que si les passerelles entre contextes sont acycliques. Encore une fois, ces formalismes ne tiennent compte que d'un seul niveau d'intégration de connaissances hétérogènes.

3.2 MODULARITÉ

Les formalismes évoqués précédemment (DDL, \mathcal{E} -connection, P-DL, IDDL) ont été décrits comme des langages d'ontologies modulaires et comparés comme tels (Cuenca-Grau et Kutz 2007).

Une partie des travaux sur la modularisation aborde la détection et la décomposition en modules de grandes ontologies monolithiques (d'Aquin 2012) ainsi que les propriétés permettant un raisonnement modulaire (Konev et al. 2009). Nous ne nous intéressons pas à ces aspects. En revanche, nos travaux se rapprochent de la construction d'ontologies modulaires, en particulier à partir de modules hétérogènes. Dans ce domaine, le langage d'ontologies distribuées (*Distributed Ontology Language* (Mossakowski et al. 2013a)) permet de combiner plusieurs ontologies et de décrire des *mappings* entre elles, ainsi que de fournir des métadonnées sur l'ensemble des ontologies. Ce langage a été soumis pour une normalisation au sein du *Object Management Group*². Il est intéressant de noter que la sémantique formelle de ce langage n'est pas définie de façon unique, tout comme celle des réseaux de connaissances multi-niveaux. En effet, les auteurs proposent trois sémantiques possibles (Mossakowski et al. 2013b), se conformant à l'analyse des sémantiques distribuées de Zimmermann et

2. <http://ontoiop.org/>

Euzenat (2006). Encore une fois, le langage ne traite pas de l'hétérogénéité à plusieurs niveaux.

La construction d'ontologies modulaires est également abordée selon une approche ingénierie logiciel (Euzenat et al. 2007) où les auteurs proposent de définir un module par (1) son ontologie interne, (2) des interfaces permettant d'exposer certains termes et axiomes pour les voir réutilisés par d'autres modules, (3) des modules importés et (4) des alignements entre modules importés. Ceci permet alors une construction hiérarchique similaire à notre notion de réseau de connaissances multi-niveaux car un module composé peut être à son tour importé avec d'autres alignements. En revanche, ce travail ne traite pas de l'alignement de termes d'alignements (c'est-à-dire, des *links*). De même que dans le présent travail, les auteurs ne spécifient pas une sémantique unique, mais, caractérisent plutôt un ensemble de sémantiques possibles pouvant se conformer aux logiques de description, DDL, ou encore P-DL. Cette approche a été implémentée dans un outil appelé ModOnto (Bezerra et al. 2008).

Dans le même ordre d'idée, Ensan et Du (2011) proposent un formalisme d'ontologies modulaires basé sur les interfaces (*Interface-Based modular Ontology Formalism* ou IBF). Dans ce cas, les relations de type *mapping* et de type *links* sont exprimés sous forme d'axiomes entre les entités d'ontologies modulaires différentes mais à travers les interfaces, afin d'encapsuler les axiomes des ontologies.

3.3 MODÉLISATION CONTEXTUELLE DES CONNAISSANCES

Certains travaux sur le raisonnement contextuel n'utilisent pas des correspondances telles que nous les définissons ici. Au contraire, les connaissances de multiples contextes sont exploitées conjointement par l'intermédiaire d'une meta-description des contextes eux-mêmes. Les relations établies entre les contextes jouent un rôle similaire aux alignements inter-ontologies, mais, n'expriment pas des correspondances entre entités de

différentes ontologies. Klarman (2013) propose un tel formalisme structuré en deux niveaux. Le premier niveau concerne la méta-ontologie exprimée en DL, où, les contextes sont considérés comme des instances. Le deuxième niveau concerne les connaissances locales détaillant les contextes. Ce formalisme correspond au langage $\mathcal{C}_{\mathcal{L}_O}^{\mathcal{L}_C}$ qui est la composition du langage objet \mathcal{L}_O (celui des connaissances locales) et le langage contextuel \mathcal{L}_C . L'implémentation de ce travail a fourni un système workflow qui permet de raisonner sur les deux niveaux et répondre à des requêtes SPARQL. Par ailleurs, le formalisme permet également d'intégrer la notion de temps et de provenance dans les connaissances contextuelles. Nous n'aborderons pas ces dimensions du contexte dans les réseaux de connaissances multi-niveaux.

Contextualized Knowledge Repository (CKR (Homola et Serafini 2012)) introduit la notion de classe contexte et généralise la propagation des connaissances en introduisant un opérateur, *eval*, pour représenter l'extension d'un concept ou d'un rôle donné dans un autre contexte. CKR permet de formaliser la relation de couverture entre les contextes qui établissent la relation d'inclusion entre leurs domaines correspondants ce qui rend l'opération d'héritage entre les objets appartenant aux contextes concernés par cette relation possible. CKR est une structure à deux niveaux : le niveau supérieur décrit les méta-connaissances sur les contextes selon plusieurs dimensions (temporelle, spatiale, etc.) et le niveau inférieur constitué par les contextes locaux contient les instances valides dans ce contexte. Et pour favoriser la réutilisation des contextes, les connaissances de chaque contexte sont organisées en un ensemble de modules. CKR repose sur trois langages : le méta-langage \mathcal{L}_Γ , le langage-objet \mathcal{L}_Σ et le langage \mathcal{L}_Σ^e qui est l'extension de \mathcal{L}_Σ avec les expressions *eval* dans \mathcal{L}_Σ se basant tous sur le langage DL.

Joseph et Serafini (2011), Joseph (2015) proposent un formalisme à deux niveaux (niveau méta-connaissances et niveau objet) similaire à CKR mais fondé sur RDF au lieu de DL. Ils se sont intéressés aux dépôts de

connaissances RDF(S) constitués d'un ensemble de triplets. La base de connaissances contextualisées (Contextualised Knowledge Base CKB) est une paire $\langle \mathcal{M}, \mathcal{G} \rangle$ où \mathcal{M} est le niveau méta-connaissance qui comporte le graphe des dimensions (temps, localisation, ect...), les définitions des contextes ainsi que les relations entre les contextes. \mathcal{G} est un ensemble de graphes RDF nommés, soit un graphe par contexte et qui représente l'ensemble des triplets valides dans ce contexte. Cet ensemble de graphes représente le niveau objet de CKB. Le raisonnement s'effectue sur les deux niveaux en application des règles d'inférence qu'ils ont définis pour permettre la propagation des connaissances à travers les niveaux ainsi qu'à travers les contextes.

Encore une fois, ces formalismes ne résolvent pas le problème de l'hétérogénéité à un niveau supérieur. En effet, les connaissances sur les contextes font introduire des termes spécifiques au méta-langage qui pourraient varier d'un réseau de connaissances à un autre.

3.4 DISCUSSION

Dans cette section, nous exposons certaines caractéristiques qui peuvent être considérées comme des apports de notre contribution par rapport aux travaux décrits dans les sections précédentes que nous classons dans les titres suivants :

3.4.1 La description et la structure des connaissances locales ne sont pas fixées

Notre réseau de connaissances multi-niveaux est composé d'un ensemble de nœuds alignés, ces mêmes nœuds sont composés des sous-nœuds alignés et ainsi de suite, où les nœuds les plus élémentaires sont des ontologies. L'alignement des nœuds composés de sous-nœuds et des alignements entre eux permet d'aligner les alignements et grâce à cette structure les alignements inter-alignements peuvent être formalisés. Au-

cun formalisme cité ci-dessus ne tolère une représentation dynamique des connaissances locales et alignées. En plus, la formalisation syntaxique des connaissances locales (ontologie et nœud) dans notre formalisme est décrite d'une manière abstraite et indépendante de tout langage et par conséquent, elle peut être adaptée à toute logique. Les formalismes DDL, P-DL, $E - SHIQ$ et IDDL sont développés pour un réseau d'ontologies en logique de description. Les ontologies dans le formalisme DFOL peuvent être exprimées en logique de premier ordre. Dans \mathcal{E} -connection, les ontologies locales d'un même réseau peuvent être représentées en diverses logiques accompagnées d'un système de description abstrait.

3.4.2 Représentation contextuelle des alignements

Les alignements dans les réseaux de connaissances multi-niveaux sont exprimés dans un langage dit "langage d'alignement" indépendant des langages des ontologies concernées. Ils ont leurs propres vocabulaires et peuvent être de type *mapping* ou de type *link* et ils sont exprimés entre deux vocabulaires, selon le point de vue de la combinaison de la paire d'ontologies concernées. A la différence de DDL et IDDL qui ne définissent et n'interprètent que des *mappings*, \mathcal{E} -connection et $E - SHIQ$ expriment des *links* mais ne prennent pas en considération le conflit de l'hétérogénéité des alignements car ils sont orientés et interprétés selon le point de vue de l'ontologie cible de la correspondance. La définition des correspondances dans un point de vue global a été déjà présentée dans le formalisme IDDL, mais vu l'absence des *links*, donc de vocabulaire d'alignement, ne nécessite pas d'alignement de niveaux supérieurs.

3.4.3 Plusieurs sémantiques peuvent être associées au formalisme proposé

Pour un réseau de connaissances multi-niveaux composé d'une famille d'ontologies en logique de description et un ensemble d'alignements correspondant, nous avons associé deux sémantiques pour l'interpréter :

- L’approche DL qui adopte la sémantique de la famille des logiques de description, la sémantique du réseau dans cette approche est centralisée, similaire à la sémantique de SomeWhere et SomeRDF et même la sémantique de l’importation de OWL ;
- L’approche DDL qui adopte la sémantique des logiques de description distribuée avec une sémantique contextuelle et distribuée comme c’est le cas de DDL, PDL, \mathcal{E} -connection et $E - SHIQ$ sauf que dans notre cas les alignements ne sont pas interprétés selon le point de vue d’une ontologie cible de la correspondance, mais ils sont interprétés dans un niveau externe et indépendant des ontologies locales, ce niveau externe est représenté par un domaine d’interprétation associé à l’ontologie alignement générée.

3.4.4 Raisonnement

plusieurs raisonnements peuvent être associés, le prototype DLMLKBR permet de raisonner sur le formalisme proposé adoptant l’approche DL où le raisonnement est centralisé et non contextuel comme c’est le cas du raisonnement pour l’importation dans OWL. Les algorithmes de SomeWhere et SomeRDF peuvent aussi être exploités, mais dans le cas où les *links* sont ignorés afin d’assurer un raisonnement distribué et non contextuel pour notre formalisme. L’implémentation de l’approche proposée DDL en exploitant le raisonneur DRAGO permet d’avoir un raisonnement distribué et contextuel sur le réseau de connaissances multi-niveaux.

CONCLUSION DU CHAPITRE

Dans ce chapitre, nous avons présenté les formalismes existants pour représenter et raisonner sur un ensemble d’ontologies alignées. Ces formalismes ont été développés pour résoudre un problème d’hétérogénéité et répondre à une motivation tracée auparavant (voir la table 3.2). Mais au-

Formalisme	Motivation	Connaissances locales	Alignement	Sémantique	Raisonnement
SomeWhere SomeRDF	réécriture des requêtes	ontologie OWL ontologie RDF	mappings, point de vue de l'ontologie cible	centralisée	distribué non contextuel
DFOL	réduire l'hétérogénéité sémantique des ontologies	ontologie en logique en premier ordre	passerelles directionnelles	distribuée	contextuel
DDL	réduire l'hétérogénéité sémantique des ontologies	ontologie DL	mappings, point de vue de l'ontologie cible	distribuée	distribué contextuel
\mathcal{E} -connection	combinaison des ontologies	ontologie en logique	<i>link</i> point de vue de l'ontologie cible	distribuée	combiné contextuel
P-DL	import des entités des ontologies	ontologie DL	terme étranger point de vue de l'ontologie cible	distribuée	distribué contextuel
$E - SHIQ$	combinaison des ontologies ; réduire l'hétérogénéité sémantique des ontologies	ontologie DL	<i>mapping, link</i> point de vue de l'ontologie cible	distribuée	distribué contextuel
IDDL	réduire l'hétérogénéité sémantique ; des ontologies ; médiation	ontologie DL	<i>mapping</i> point de vue global	intégrée	distribué contextuel
MLNK	réduire l'hétérogénéité sémantique des ontologies et des alignements ; combinaison des ontologies	nœuds composés de sous nœuds alignés indépendant de langage ; le nœud le plus élémentaire est une ontologie	<i>mapping, link</i> point de vue global	DL : centralisée DDL : distribuée	centralisé, \neg contextuel distribué, contextuel

TABLE 3.2 – Synthèse

cun de ces formalismes n'a soulevé le problème de l'hétérogénéité sémantique entre les alignements inter-ontologies qui peut surgir dans le cas où le langage d'alignement possède son propre vocabulaire. Dans le chapitre suivant, nous détaillons le formalisme proposé, et pour garder le principe d'une gestion indépendante des alignements, les correspondances sont exprimées selon un point de vue général par rapport à une paire d'ontologies. Le formalisme proposé apporte une solution pour résoudre la problématique de l'hétérogénéité à plusieurs niveaux des connaissances où l'hétérogénéité dans le niveau 0 concerne l'hétérogénéité dans le niveau local des connaissances c.à.d le niveau des ontologies et les niveaux supérieurs concernent l'hétérogénéité inter-ontologies, donc au niveau des alignements puis les niveaux inter-alignements.

LE FORMALISME RÉSEAU DE CONNAISSANCES MULTI-NIVEAUX : SYNTAXE ET SÉMANTIQUE

SOMMAIRE

4.1	SYNTAXE	67
4.1.1	Langage d'alignement	67
4.1.2	Nœud de connaissances	69
4.2	SÉMANTIQUE	71
4.2.1	Réseau de connaissances multi-niveaux en logique de description (DL)	72
4.2.2	Réseau de connaissances multi-niveaux en logique de description distribuée (DDL)	81
4.3	COMPARAISON DES DEUX APPROCHES SÉMANTIQUES DU RÉSEAU DE CONNAISSANCES MULTI-NIVEAUX	88
4.4	DISCUSSION	95
	CONCLUSION	96

DANS ce chapitre, nous proposons un formalisme de représentation d'un réseau de connaissances multi-niveaux (Klai et al. 2016b), (Klai

et al. 2016a). Nous commençons par détailler la syntaxe de ses composants puis nous proposons deux approches d'interprétation de ce réseau à savoir les approches DL et DDL. Une comparaison de ces deux approches sémantiques selon certains critères sera donnée par la suite.

La représentation syntaxique et la formalisation de la sémantique de notre réseau de connaissances multi-niveaux suppose d'une part, que l'on sache représenter chaque composant de ce réseau et en formaliser la sémantique, et d'autre part que l'on puisse représenter les relations entre ces composants. Le réseau de connaissances multi-niveaux proposé est composé des ontologies locales dont nous maintenons la représentation syntaxique et sémantique présentée dans la section 2.1, des alignements et des nœuds de connaissances.

4.1 SYNTAXE

4.1.1 Langage d'alignement

Un langage d'alignement L_A permet de décrire des correspondances entre deux vocabulaires. Comme tout langage, il se caractérise aussi par une syntaxe (la manière dont les correspondances sont écrites) et une sémantique (la manière dont sont interprétées ces correspondances). La syntaxe d'un langage d'alignements est définie par :

- un ensemble de termes, appelés *links*, spécifiques au langage d'alignement, noté $V(L_A)$;
- une fonction $E(L_A)$ associant à toute signature Σ d'un langage de représentation de connaissances L un ensemble d'entités alignables ;
- un ensemble de symboles de relations $R(L_A)$.

Ainsi, la syntaxe du langage d'alignements L_A est caractérisée par un triplet $\langle V(L_A), E(L_A), R(L_A) \rangle$, qu'on notera, lorsqu'il n'y a pas d'ambiguïté $\langle V, E, R \rangle$.

On distingue deux types de correspondances : les correspondances de type *mappings* et les correspondances de type *links*.

Définition 4.1 (Correspondance de type *mapping*) Soient V_1 et V_2 deux vocabulaires à aligner. Soit $L_A = \langle V, E, R \rangle$ un langage d'alignements. Une correspondance de type *mapping* est un triple $\langle e_1, e_2, r \rangle$ que l'on note $e_1 \xrightarrow{r} e_2$ où :

- $e_1 \in E(V_1)$ et $e_2 \in E(V_2)$ sont des entités alignables relativement à V_1 et V_2 ;
- $r \in R$ dénote une relation existante entre e_1 et e_2 .

Dans l'exemple 1, $\text{eq:appartient} \xleftrightarrow{\perp} \text{mt:appartient}$ est une correspondance de type *mapping*. Le terme *appartient* figure dans les vocabulaires des deux ontologies *eq* et *mt* avec des sens différents. Les deux ontologies *eq* et *mt* sont formalisées en logique de description et les *mappings* sont construits en utilisant l'ensemble d'opérateurs $R = \{\sqsubseteq, \equiv, \perp, \in, =\}$. Ces opérateurs ont une sémantique fixe dont une interprétation mathématique est associée :

- \sqsubseteq correspond à un sous ensemble (i.e., \subseteq) ;
- \equiv correspond à une équivalence d'ensembles ;
- \perp correspond à des ensembles disjoints ;
- \in correspond à l'appartenance d'un élément à un ensemble ;
- $=$ correspond à l'égalité d'individus.

Définition 4.2 (Correspondance de type *link*) Soient V_1 et V_2 deux vocabulaires à aligner.

Soit $L_A = \langle V, E, R \rangle$ un langage d'alignements. Une correspondance de type *link* est une formule $e_1 \xleftrightarrow{l} e_2$ où :

- $e_1 \in E(V_1)$ et $e_2 \in E(V_2)$ sont des entités alignables relativement aux vocabulaires V_1 et V_2 ;
- $l \in V$ dénote une relation existante entre e_1 et e_2 .

Toujours en se référant à l'exemple 1, $\text{eq:DF}_1 \xleftrightarrow{\text{fait-partie-de}} \text{zn:ANNA1TG01}$ et $\text{mt:I}_1 \xleftrightarrow{\text{concerne}} \text{eq:DF}_1$ sont des correspondances de type *links*. Les termes **fait-partie-de** et **concerne** ne figuraient pas dans les vocabulaires des ontologies, on les a introduit au niveau des alignements pour lier des entités de vocabulaires différents. Ainsi, l'alignement possède son propre vocabulaire et, comme tout vocabulaire, il peut être aligné avec d'autres vocabulaires pour éviter le problème d'hétérogénéité sémantique entre eux.

Définition 4.3 (Alignement) Soient V_1 et V_2 deux vocabulaires à aligner. Un alignement est un triplet $\langle V, \kappa, \lambda \rangle$ où :

- V est le vocabulaire de l'alignement ;
- κ est l'ensemble de correspondances de type mappings, $e_1 \xleftrightarrow{r} e_2$ où $e_1 \in E(V_1)$, $e_2 \in E(V_2)$ et $r \in R$;
- λ est l'ensemble de correspondances de type links, $e_1 \xleftrightarrow{l} e_2$ où $e_1 \in E(V_1)$, $e_2 \in E(V_2)$ et $l \in V$;

Exemple 5 En DDL et en IDDL, l'alignement est entre les signatures des ontologies, ils n'expriment pas de termes au niveau d'alignement, donc les ensembles V et λ sont vides.

Exemple 6 Si on considère l'alignement $A_{\text{mt-eq}}$ de l'exemple 1, le vocabulaire de l'alignement $V = \{\text{concerne}\}$, $\kappa = \{\text{eq:appartient} \xleftrightarrow{\perp} \text{mt:appartient}\}$ et l'ensemble de correspondances de type $\lambda = \{\text{mt:I}_1 \xleftrightarrow{\text{concerne}} \text{eq:DF}_1\}$.

Exemple 7 En \mathcal{E} -connections, les connaissances inter-ontologies peuvent évoquer des termes entre deux ou plusieurs ontologies. Cependant, si on réduit les axiomes de \mathcal{E} -connection en cette forme $\langle E_i \rangle^j(a_i, b_j)$, où E_i est un link, a_i est un individu de l'ontologie O_i et b_j est un individu de l'ontologie O_j , alors, on peut les assimiler à l'ensemble λ , avec l est un terme du vocabulaire d'alignement (mais l'ensemble κ est vide).

4.1.2 Nœud de connaissances

Les composants de base d'un réseau de connaissances ont été introduits. Il est maintenant possible d'introduire la notion de nœud de connaissances qui généralise la notion d'ontologie. Informellement, une ontologie est un nœud de connaissances de niveau 0, tandis que tout nœud de connaissances de niveaux $m > 0$ est construit à partir d'un ou plusieurs nœuds de niveau inférieur, liés entre eux par des alignements (voir la figure 4.1).

Formellement, le nœud est défini comme suit :

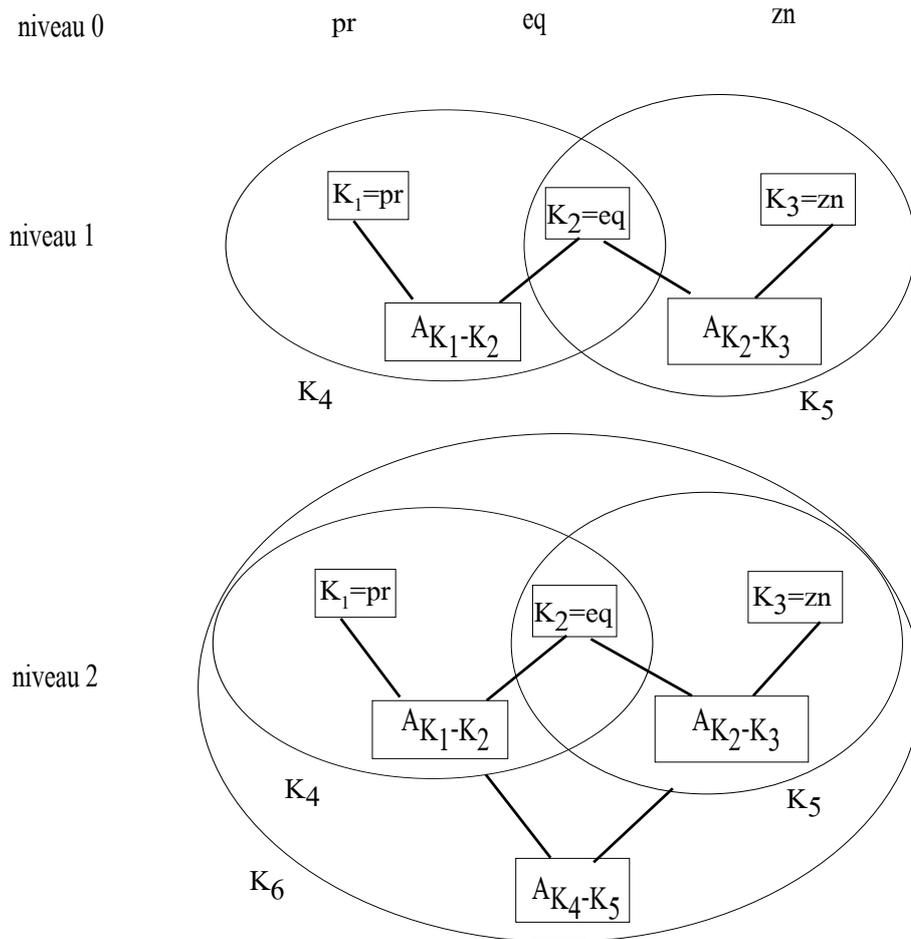


FIGURE 4.1 – Représentation récursive des noeuds

Définition 4.4 (Nœud de connaissances) *Un nœud de connaissances est un couple $K = \langle V_K, A_K \rangle$ où V_K est le vocabulaire, noté aussi $\text{Voc}(K)$, V_K et A_K sont définis d'une manière récursive :*

- une ontologie O est un nœud de connaissances avec le vocabulaire $\text{Voc}(O) = \Sigma(O)$ et $A_K = A(O)$ l'ensemble des axiomes de O ;
- pour $n \geq 1$, si K_1, \dots, K_n sont des nœuds de connaissances avec les vocabulaires $\text{Voc}(K_1), \dots, \text{Voc}(K_n)$, et pour tout $i, j \in [1, n]$, Λ_{ij} est un alignement de $\text{Voc}(K_i)$ et $\text{Voc}(K_j)$, alors $K = \langle V_K, A_K \rangle$ est un nœud de connaissances avec :
 - le vocabulaire

$$V_K = \bigcup_{i,j \in [1,n]} \text{Voc}(\Lambda_{ij}) \cup \bigcup_{i \in [1,n]} \{i:e \mid e \in \text{Voc}(K_i)\}$$

$$\text{– et } A_K = \langle (K_i)_{i \in [1,n]}, (\Lambda_{ij})_{i,j \in [1,n]} \rangle.$$

Si le réseau de connaissances ne contient que des ontologies et des alignements d'ontologies alors on l'appelle un réseau d'ontologies alignées. Dans le cas où il contient des ontologies, des alignements inter-ontologies et des alignements inter-nœuds alors le nœud formé à partir de la définition précédente est appelé *réseau de connaissances multi-niveaux*.

4.2 SÉMANTIQUE

Nous avons choisi d'exploiter des paradigmes existants disposant des procédures de raisonnement complètes ainsi que des outils opérationnels tels que les logiques de description et les logiques de description distribuées. Les deux manières d'interpréter ont été étudiées pour pouvoir raisonner sur le réseau de connaissances multi-niveaux.

4.2.1 Réseau de connaissances multi-niveaux en logique de description (DL)

Dans cette interprétation, nous présentons un réseau de connaissances multi-niveaux comme une ontologie unique en DL, cette ontologie est construite à partir d'un ou plusieurs nœuds liés entre eux par des alignements (Klai et al. 2016b). La particularité de cette ontologie est que ses éléments sont préfixés par l'indice des sources de connaissances correspondantes afin de distinguer entre les entités provenant de sources de connaissances différentes. Donc la représentation de cette ontologie nous incite à introduire de nouvelles notions tels que la signature combinée et l'ensemble combiné de formules. Dans cette section, nous allons montrer que le raisonnement sur le réseau de connaissances multi-niveaux se limite à l'application de l'algorithme de raisonnement en logique de description.

Avant de définir la signature combinée et l'ensemble combiné de formules et les constituants de l'ontologie DL générée à partir d'un nœud de connaissances multi-niveaux, nous allons donner une fonction supplémentaire qui permet d'indicer les éléments d'une ontologie donnée.

Définition 4.5 (indicer les éléments d'une ontologie) *Soit i un indice, on définit la fonction prefix sur les termes, les axiomes et les ontologies de la façon suivante :*

- $\text{prefix}(X, i) = \{i:X\}$ où X est un concept atomique, un role atomique ou un individu ;
- $\text{prefix}(\Sigma, i) = \{\text{prefix}(X, i) \mid X \in \Sigma\}$ où Σ est une signature d'ontologie et X correspond à un concept atomique, un role atomique ou un individu ;
- $\text{prefix}(\bullet, i) = \bullet$ où $\bullet \in \{\perp, \top\}$;
- $\text{prefix}(X \bullet Y, i) = \text{prefix}(X, i) \bullet \text{prefix}(Y, i)$ où $\bullet \in \{\sqcup, \sqcap\}$ et X, Y désignent, soit deux concepts, soit deux roles ;
- $\text{prefix}(\bullet R.C, i) = \bullet \text{prefix}(R, i). \text{prefix}(C, i)$ où $\bullet \in \{\exists, \forall\}$ et R est un role et C un concept ;
- $\text{prefix}(\bullet n.R, i) = \bullet n \text{prefix}(R, i)$ où $\bullet \in \{\leq, \geq\}$ et R est un role ;
- $\text{prefix}(\neg X, i) = \neg \text{prefix}(X, i)$ où X est un concept ou un role ;

- $\text{prefix}(R^\bullet, i) = \text{prefix}(R, i)^\bullet$ où $\bullet \in \{+, -\}$ et R est un role ;
- $\text{prefix}(X \sqsubseteq Y, i) = \text{prefix}(X, i) \sqsubseteq \text{prefix}(Y, i)$ où X et Y sont soit deux concepts, soit deux roles ;
- $\text{prefix}(\text{Trans}(R), i) = \text{Trans}(\text{prefix}(R, i))$ où R est un role ;
- $\text{prefix}(C(a), i) = \text{prefix}(C, i)(i:a)$ où C est un concept et a un individu ;
- $\text{prefix}(R(a, b), i) = \text{prefix}(R, i)(i:a, i:b)$ où R est un role, a, b des individus ;
- $\text{prefix}(a = b, i) = \{i:a = i:b\}$ où a, b sont des individus ;
- $\text{prefix}(F, i) = \{\text{prefix}(f, i) \mid f \in F\}$ où F est un ensemble d'axiomes en logique de description ;
- $\text{prefix}(O, i) = \langle \text{prefix}(\Sigma(O), i), \text{prefix}(F(O), i) \rangle$ où O est une ontologie en logique de description.

La signature combinée est la signature d'une ontologie qui résulte de la transformation du réseau de connaissances multi-niveaux. Il s'agit de l'union des signatures des nœuds locaux et des vocabulaires des alignements qui les lient.

Définition 4.6 (Signature combinée) *Soit N un nœud de connaissances multi-niveaux, la signature combinée Σ_{comb} est définie de la manière récursive suivante :*

- si N est une ontologie alors $\Sigma_{\text{comb}}(N) = \Sigma(N) = \langle C, \mathcal{R}, \mathcal{U} \rangle$ en référence à la définition 2.7 ;
- si N est un nœud de connaissances composé des sous-nœuds et d'alignements entre eux, pour $n \geq 1$, on a $\Sigma_{\text{comb}}(N_1), \dots, \Sigma_{\text{comb}}(N_n)$ les signatures combinées des nœuds locaux et pour tout $i, j \in [1, n]$, V_{ij} est le vocabulaire d'alignements entre $\Sigma_{\text{comb}}(N_i)$ et $\Sigma_{\text{comb}}(N_j)$ et prefix est la fonction donnée dans la définition 4.5. Nous définissons role une fonction qui associe à tout link de V_{ij} un role spécifique en logique de description. Alors $\Sigma_{\text{comb}}(N) = \bigcup_{i \in [1, n]} \{\text{prefix}(\Sigma_{\text{comb}}(N_i), i)\} \cup \{\text{role}(l) \mid l \in V_{ij}\}$.

L'exemple 8 détaille l'application de cette définition.

L'ensemble combiné de formules est l'ensemble des formules de l'ontologie résultante de la transformation du réseau. C'est l'union des formules

des nœuds locaux et des formules générées à partir des correspondances. Nous allons commencer par définir la fonction qui associe à toute correspondance un axiome. Rappelons que les correspondances de A_{ij} sont de deux types : les *mappings* représentés par l'ensemble κ et les *links* représentés par l'ensemble λ (voir la définition 4.3).

Définition 4.7 (Transformation des correspondances en DL-axiomes) *Soit A_{ij} pour $i, j \in [1, n]$ un alignement entre un nœud i et un nœud j . Nous définissons trans une fonction qui associe à toute correspondance de A_{ij} un axiome en logique de description. Nous définissons cette transformation pour chaque type de correspondance de la façon suivante :*

- $\text{trans}(\{i:A \xleftarrow{\sqsubseteq} j:B\}) = \{\text{prefix}(A, i) \sqsubseteq \text{prefix}(B, j)\};$
- $\text{trans}(\{i:A \xleftarrow{\equiv} j:B\}) = \{\text{prefix}(A, i) \equiv \text{prefix}(B, j)\};$
- $\text{trans}(\{i:A \xleftarrow{\perp} j:B\}) = \{\text{prefix}(A, i) \sqsubseteq \neg \text{prefix}(B, j)\};$
- $\text{trans}(\{i:u \xleftarrow{\in} j:A\}) = \{\text{prefix}(A, j)(i:u)\};$
- $\text{trans}(\{i:u \xleftarrow{=} j:u'\}) = \{i:u = j:u'\};$
- $\text{trans}(\{i:u \xleftarrow{l} j:u'\}) = \{\text{role}(l)(i:u, j:u')\};$
- $\text{trans}(\{i:A \xleftarrow{l} j:B\}) = \{\text{prefix}(A, i) \sqsubseteq \exists \text{role}(l).\text{prefix}(B, j)\}$

où A, B, u et u' sont des entités alignables et l est un symbole de link.

Définition 4.8 (Ensemble combiné de formules) *Soit N un nœud de connaissances multi-niveaux. L'ensemble combiné de formules F_{comb} pour une signature combinée Σ_{comb} est défini récursivement comme suit :*

- si N est une ontologie alors $F_{\text{comb}}(N) = A(N)$ l'ensemble des axiomes de N ;
- si N est un nœud de connaissances composé de sous nœuds N_1, \dots, N_n et d'alignement A_{ij} , pour tout $i, j \in [1, n]$ $F_{\text{comb}}(N_1), \dots, F_{\text{comb}}(N_n)$ sont des ensembles d'axiomes des nœuds locaux, $\text{trans}(A_{ij})$ la fonction qui associe à toute correspondance de A_{ij} un axiome en logique de description (voir la définition 4.7) et prefix la fonction qui attribue des indices aux éléments des sous nœuds (voir la définition 4.5). Alors l'ensemble combiné de

formules $F_{comb}(N) = \bigcup_{i \in [1,n]} \{\text{prefix}(F_{comb}(N_i), i)\} \cup \bigcup_{i,j \in [1,n]} \{\text{trans}(c) \mid c \in A_{ij}\}$.

On peut actuellement introduire la notion de réseau de connaissances multi-niveaux en logique de description à partir de ces définitions.

Définition 4.9 (Réseau de connaissances multi niveaux en DL) *Soit N un nœud de connaissances multi-niveaux, nous définissons F_{DL} une fonction qui permet de transformer le nœud N en une ontologie DL d'une manière récursive suivante :*

- si N est une ontologie alors $F_{DL}(N) = N$;
- si N est un nœud composé de sous nœuds N_i pour $i \in [1, n]$ et d'alignements A_{ij} pour $i, j \in [1, n]$ alors $F_{DL}(N) = F_{DL}(\{N_i\}, \{A_{ij}\}) = \langle \Sigma_{comb}(N), F_{comb}(N) \rangle$ tels que $\Sigma_{comb}(N)$ est une signature combinée et $F_{comb}(N)$ est un ensemble combiné de formules.

Cette définition nous permet d'étendre la définition 4.5 pour préfixer les termes d'un nœud composé de sous nœuds.

Définition 4.10 (Nœud préfixé) *Pour un nœud de connaissances X , le nœud préfixé $\text{prefix}(X, i)$ est l'ontologie $\langle \text{prefix}(\Sigma_{comb}(X), i), \text{prefix}(F_{comb}(X), i) \rangle$.*

Exemple 8 $K_1 = \text{pr}$, $K_2 = \text{eq}$ et $K_3 = \text{zn}$. L'alignement entre K_1 et K_2 donne lieu à un nœud $K_4 = \{K_1, K_2, A_{K_1-K_2}\}$, la transformation de ce nœud en une ontologie DL est donnée par $F_{DL}(K_4) = F_{DL}(K_1, K_2, A_{K_1-K_2}) = \langle \Sigma_{comb}(K_4), F_{comb}(K_4) \rangle$. En application de la définition 4.6, on obtient la signature combinée de K_4 et en application de la définition 4.8, on obtient l'ensemble des axiomes combinés de K_4 et qui figurent dans la table 4.1.

On procède de la même façon pour le nœud $K_5 = \{K_2, K_3, A_{K_2-K_3}\}$ et le nœud $K_6 = \{K_4, K_5, A_{K_4-K_5}\}$.

À travers l'exemple 8, nous allons appliquer les définitions précédentes pour générer une ontologie DL à partir du réseau de connaissances multi-niveaux dont les ontologies, alignements et inter-alignements sont présentés dans l'exemple 1.

Remarque : le nœud K_6 est composé des nœuds K_4 et K_5 , chacun com-

Noeuds	Signatures combinées	Axiomes combinés
$K_1 = \text{pr}$ $K_2 = \text{eq}$	$\Sigma_{comb}(K_1) = \{\text{garniture}, G_1\}$ $\Sigma_{comb}(K_2) = \{\text{detecteur-flamme},$ $\text{instrumentation}, DF_1, \text{appartient}\}$	$F_{comb}(K_1) = \{\text{garniture}(G_1)\}$ $F_{comb}(K_2) = \{\text{detecteur-flamme}(DF_1),$ $\text{detecteur-flamme} \sqsubseteq$ $\exists \text{appartient.instrumentation}\}$ $F_{comb}(K_3) = \{\text{zone}(ANNA1TG01)\}$
$K_3 = \text{zn}$	$\Sigma_{comb}(K_3) = \{\text{zone}, ANNA1TG01\}$	$F_{comb}(K_3) = \{\text{zone}(ANNA1TG01)\}$
$K_4 =$ $\{K_1, K_2,$ $A_{K_1-K_2}\}$	$\Sigma_{comb}(K_4) = \{k_1:\text{garniture}, k_1:G_1,$ $k_2:\text{detecteur-flamme},$ $k_2:\text{instrumentation},$ $k_2:DF_1, k_2:\text{appartient},$ $\text{compose}\}$	$F_{comb}(K_4) = \{k_1:\text{garniture}(k_1:G_1),$ $k_2:\text{detecteur-flamme}(k_2:DF_1),$ $k_2:\text{detecteur-flamme} \sqsubseteq$ $\exists(k_2:\text{appartient}).(k_2:\text{instrumentation}),$ $\text{compose}(k_1:G_1, k_2:DF_1)\}$
$K_5 =$ $\{K_2, K_3,$ $A_{K_2-K_3}\}$	$\Sigma_{comb}(K_5) = \{k_2:\text{detecteur-flamme},$ $k_2:\text{instrumentation}, k_2:DF_1,$ $k_2:\text{appartient}, k_3:\text{zone},$ $k_3:ANNA1TG01,$ $\text{fait-partie-de}\}$	$F_{comb}(K_5) = \{k_2:\text{detecteur-flamme}(k_2:DF_1),$ $k_2:\text{detecteur-flamme} \sqsubseteq \exists(k_2:\text{appartient})$ $(k_2:\text{instrumentation}),$ $k_3:\text{zone}(k_3:ANNA1TG01),$ $\text{fait-partie-de}(k_2:DF_1, k_3:ANNA1TG01)\}$
$K_6 =$ $\{k_4, K_5,$ $A_{K_1-K_2}\}$	$\Sigma_{comb}(K_6) = \{k_4:k_1:\text{garniture},$ $k_4:k_1:G_1, k_4:k_2:\text{detecteur-flamme},$ $k_4:k_2:\text{appartient}, k_4:k_2:DF_1,$ $k_4:k_2:\text{instrumentation},$ $k_5:k_3:\text{zone}, k_5:k_3:ANNA1TG01,$ $k_5:k_2:\text{appartient}, k_5:k_2:\text{instrumentation},$ $k_5:k_2:\text{detecteur-flamme}, k_5:k_2:DF_1$ $k_4:\text{compose}, k_5:\text{fait-partie-de}\}$	$F_{comb}(K_6) = \{k_4:k_1:\text{garniture}(k_4:k_1:G_1)$ $k_4:k_2:\text{detecteur-flamme} \sqsubseteq \exists(k_4:k_2:\text{appartient}).$ $(k_4:k_2:\text{instrumentation}),$ $k_4:k_2:\text{detecteur-flamme}(k_4:k_2:DF_1),$ $k_5:k_3:\text{zone}(k_5:k_3:ANNA1TG01),$ $k_4:\text{compose}(k_1:G_1, k_2:DF_1),$ $k_5:\text{fait-partie-de}(k_2:DF_1, k_3:ANNA1TG01),$ $k_5:k_2:\text{detecteur-flamme}(k_5:k_2:DF_1),$ $k_5:k_2:\text{detecteur-flamme} \sqsubseteq \exists(k_5:k_2:\text{appartient}).$ $(k_5:k_2:\text{instrumentation}),$ $k_4:\text{compose} \equiv k_5:\text{fait-partie-de}\}$

TABLE 4.1 – Signatures et axiomes combinés

prend le nœud K_2 , donc les termes de K_2 sont dupliqués car ils peuvent être représentés dans des points de vue différents selon K_4 et K_5 .

Une interprétation \mathcal{I} de logique de description est assignée à l'ontologie DL générée à partir de la transformation d'un nœud de connaissances multi-niveaux. La relation de satisfaction au niveau d'un nœud dépend de la relation de satisfaction en logique de description. Pour un nœud X composé de $\{O_i\}, \{A_{ij}\}$ pour $i, j \in [1..n]$, l'interprétation \mathcal{I} de l'ontologie en DL résultante de la transformation englobe les interprétations locales \mathcal{I}_i et \mathcal{I}_j . Autrement dit, les interprétations locales sont des sous-ensembles de l'interprétation globale \mathcal{I} , et le domaine d'interprétation global $\Delta^{\mathcal{I}} = \Delta^{\mathcal{I}_i} \cup \Delta^{\mathcal{I}_j}$ (ceci est illustré par la figure 4.2) où l est un *link*

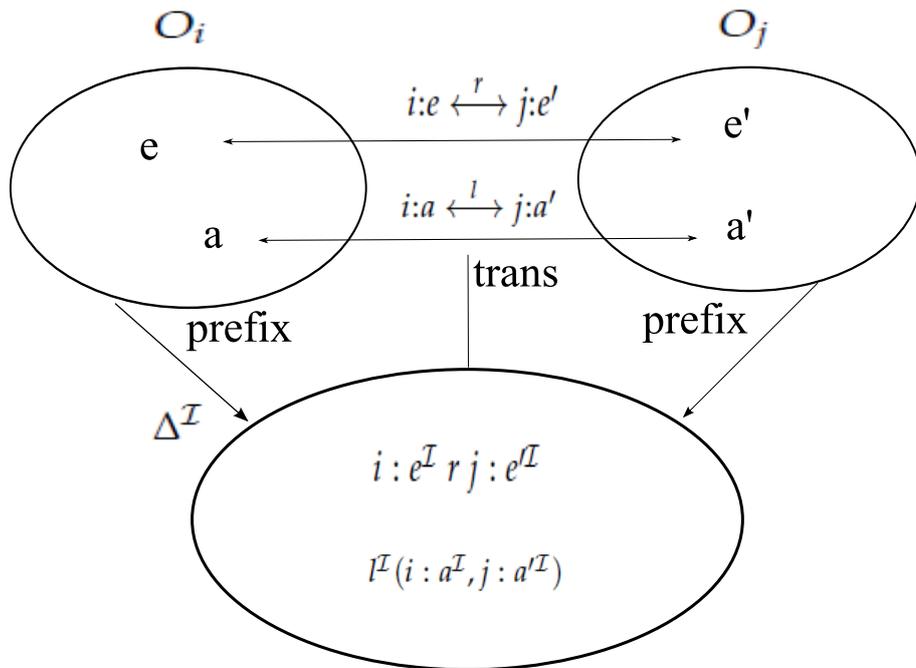


FIGURE 4.2 – Interprétation d'un réseau de connaissances en utilisant l'approche DL

avec $r \in \{\sqsubseteq, \equiv, \perp, \in, =\}$, $i:e, j:e'$ sont des concepts, rôles ou individus provenant respectivement des ontologies O_i et O_j et $i:a, j:a'$ sont des individus provenant respectivement des ontologies O_i et O_j .

Définition 4.11 (Interprétation d'un nœud de connaissances) Une DL-interprétation d'un nœud de connaissances multi-niveaux X est une interprétation en logique de description de la signature combinée de X .

Définition 4.12 (Relation de satisfaction pour un nœud de connaissances) Soient X un nœud de connaissances multi-niveaux, I une DL-interprétation de X et $F_{DL}(X)$ est la fonction donnée dans la définition 4.9. Alors l'interprétation I DL-satisfait X (noté $I \models_{N-DL} X$) si et seulement si I satisfait $F_{DL}(X)$ au sens des logiques de description. Dans ce cas, on dit que I est un DL-modèle de X .

Exemple 9 Soit un nœud de connaissances $X = \langle \{O_1, O_2\}, \{A_{12}\} \rangle$ avec $O_1 = \{C \sqsubseteq D\}$, $O_2 = \{C \equiv B\}$ et $A_{12} = \{1:C \overset{\perp}{\leftarrow} 2:C\}$. Le modèle de O_1 est une interprétation de la signature $\{C, D\}$ et le modèle de O_2 est une interprétation de la signature $\{C, B\}$. La transformation de X en une ontologie DL est donnée comme suit :

- $\text{prefix}(O_1, 1) = \{1:C \sqsubseteq 1:D\}$;
- $\text{prefix}(O_2, 2) = \{2:C \equiv 2:B\}$;
- $\text{trans}(\{1:C \overset{\perp}{\leftarrow} 2:C\}) = \{1:C \sqsubseteq \neg 2:C\}$;
- $F_{DL}(X) = \{1:C \sqsubseteq 1:D, 2:C \equiv 2:B, 1:C \sqsubseteq \neg 2:C\}$;

Un modèle de X est une interprétation I de la signature combinée $\{1:C, 1:D, 2:C, 2:B\}$ alors $I \models_{N-DL} X$ si $I \models F_{DL}(X)$ et I satisfait $F_{DL}(X)$ si $I \models 1:C \sqsubseteq 1:D$, $I \models 2:C \equiv 2:D$ et $I \models 1:C \sqsubseteq \neg 2:C$ avec :

- $I \models 1:C \sqsubseteq 1:D$ si et seulement si $(1:C)^I \subseteq (1:D)^I$;
- $I \models 2:C \equiv 2:D$ si et seulement si $(2:C)^I \equiv (2:B)^I$;
- $I \models 1:C \sqsubseteq \neg 2:C$ si et seulement si $(1:C)^I \subseteq \neg(2:C)^I$.

Le raisonnement automatique sur un nœud X se limite à un raisonnement sur l'ontologie générée $F_{DL}(X)$ où la cohérence de X est déduite à partir de la cohérence de $F_{DL}(X)$. Et la notion de structuration des connaissances en niveaux nous permet d'utiliser la relation d'implication entre les nœuds de niveaux différents.

Définition 4.13 (Implication de nœuds) Soient X, Y deux nœuds de connaissances multi-niveaux, on dit que X DL-implique Y si et seulement si tous les DL-modèles de X sont des DL-modèles de Y .

Propriété 4.1 Soient X un nœud de connaissances global comprenant des nœuds locaux

N_1, \dots, N_k , prefix la fonction qui affecte le préfixe du nœud à tous ses termes alors X implique tous ses nœuds locaux préfixés noté ($X \models_{N-DL} \text{prefix}(N_i, i)$ pour tout $i \in [1, k]$) mais l'inverse n'est pas vrai. On peut dire qu'un nœud X implique directement ses nœuds locaux dans le cas particulier qui présente le fait que les termes des nœuds locaux sont strictement disjoints car, ce cas ne nous oblige pas à ajouter les préfixes aux termes afin de les distinguer.

Preuve: Soit un nœud $X = \langle \{N_i\}, \{A_{ij}\} \rangle$. Démontrer la propriété 4.1 revient à montrer que si une interprétation I est un DL-modèle de X alors I est aussi un DL-modèle de $\text{prefix}(N_i, i)$.

On suppose que $I \models_{N-DL} X$. Ceci implique que $I \models_{FDL} \langle \{N_i\}, \{A_{ij}\} \rangle$ d'après la définition 4.12 d'où $I \models \langle \Sigma_{comb}(N_i), F_{comb}(N_i) \rangle$ d'après la définition 4.9. Il s'en suit que $I \models \bigcup_{i \in [1, n]} \{ \text{prefix}(F_{comb}(N_i), i) \} \cup \bigcup_{i, j \in [1, n]} \{ \text{trans}(c) \mid c \in A_{ij} \}$ d'après la définition 4.6 et la définition 4.8 donc en particulier, $I \models \text{prefix}(F_{comb}(N_i), i)$. Cette dernière assertion équivaut à $I \models \text{prefix}(N_i, i)$ d'après la définition 4.10. Or, $\text{prefix}(N_i, i)$ est une ontologie en DL donc ce qui précède implique $I \models_{N-DL} \text{prefix}(N_i, i)$ par la définition 4.12. CQFD

Exemple 10 Soit un nœud $X = \langle \{O_1, O_2\}, \{A_{12}\} \rangle$ avec $O_1 = \{C \sqsubseteq D\}$ et $O_2 = \{C \equiv B\}$ Un modèle de O_1 est une interprétation de la signature $\{C, D\}$, un modèle de O_2 est une interprétation de la signature $\{C, B\}$ et un modèle du nœud X est une interprétation de la signature combinée $\{1:C, 1:D, 2:C, 2:B\}$ $\text{prefix}(O_1, 1) = \{1:C \sqsubseteq 1:D\}$ $\text{prefix}(O_2, 2) = \{2:C \equiv 2:B\}$ En appliquant la propriété, le nœud X implique $\text{prefix}(O_1, 1)$ et implique $\text{prefix}(O_2, 2)$, car $X \models_{N-DL} 1:C \sqsubseteq 1:D$ mais X ne satisfait pas $C \sqsubseteq D$ et $X \models_{N-DL} 2:C \equiv 2:B$ mais X ne satisfait pas $C \equiv B$.

La transformation d'un nœud de connaissances multi-niveaux nous permet de rattacher ce dernier aux logiques de description qui disposent de plusieurs outils de raisonnement opérationnels, tels que Fact++¹, Pel-

1. <http://owl.man.ac.uk/factplusplus/>

let², Hermit³ et qui peuvent être exploités pour le raisonnement sur un nœud.

On peut assimiler cette approche à celle de *l'importation de OWL* car OWL adopte un mécanisme d'import d'ontologies pour supporter l'intégration des ontologies. C'est une sémantique centralisée avec une interprétation unique \mathcal{I} qui satisfait les ontologies importées et les axiomes inter-ontologies, mais l'approche *owl:import* ne supporte pas l'existence d'un module indépendant des alignements, ces derniers sont insérés en tant que axiomes. Néanmoins, l'application de la sémantique des logiques de description pour la représentation et le raisonnement sur un réseau d'ontologies hétérogènes et alignées selon plusieurs niveaux permet de réduire le problème du conflit de l'hétérogénéité sémantique entre les termes provenant des ontologies et des alignements différents. Ceci, grâce à la fonction *prefix* qui permet à son tour de différencier entre ces termes, et les alignements intégrés à l'ensemble global sous forme d'axiomes. Elle peut être appliquée pour des contextes différents mais compatibles car, elle présente l'inconvénient de ne pas permettre la cohabitation des points de vues distincts et incompatibles et peut conduire à une inconsistance ou fournir des inférences inadaptées. Nous démontrons ceci dans la section 4.3. D'autres sémantiques ont été proposées pour représenter et raisonner sur un ensemble d'ontologies alignées tout en tolérant l'incompatibilité de leurs points de vue. Certaines de ces sémantiques ont été présentées dans le chapitre 3.

Dans la section suivante, nous adoptons la famille des logiques de description distribuées qui représente l'une des sémantiques contextuelles les plus connues, afin d'interpréter le réseau de connaissances multi-niveaux.

2. <http://pellet.owldl.com/>

3. <http://hermit-reasoner.com/>

4.2.2 Réseau de connaissances multi-niveaux en logique de description distribuée (DDL)

Cette approche consiste à transformer le réseau de connaissances multi-niveaux en un système distribué DDL (Klai et al. 2016a). Toujours pour garder le principe de perception global des alignements, nous allons générer une ontologie en logique de description (dite ontologie-alignement) à partir de ces derniers puis créer des correspondances d'équivalence qui vont jouer le rôle de passerelles entre les ontologies locales et les ontologies-alignements. L'ensemble des ontologies locales, ontologies-alignements et passerelles constitue le réseau de connaissances multi-niveaux en DDL. Nous allons commencer par détailler la procédure de construction de l'ontologie-alignement.

Génération de l'ontologie-alignement

L'ontologie-alignement est une ontologie en logique de description générée à partir de la transformation de l'alignement entre une paire de nœuds. Comme toute ontologie en DL, elle dispose d'une signature et d'un ensemble d'axiomes. La signature de cette ontologie est constituée à partir des termes (concepts, rôles ou individus) figurant dans les correspondances de type *mappings* et qui appartiennent aux vocabulaires locaux alignés. Elle contient aussi les termes figurant dans les correspondances de type *links* et qui appartiennent aux vocabulaires de l'alignement.

La signature de l'ontologie-alignement est constituée de l'ensemble des termes figurant à gauche et à droite des *mappings* et des termes *links* appartenant aux alignements.

Définition 4.14 (Signature de l'ontologie-alignement) *Soit K un nœud de connaissance multi-niveaux, la signature-alignement Σ_A est définie comme suit selon les cas :*

- si K est une ontologie alors $\Sigma_A = \emptyset$;
- si K est un nœud de connaissances multi-niveaux composé de sous nœuds K_1, \dots, K_n et de A_{ij} l'alignement entre K_i et K_j pour $i, j \in [1, n]$, alors :

$$\Sigma_A(K) = \bigcup_{i,j \in [1,n]} \{i:X, j:Y \mid i:X \xleftarrow{r} j:Y \in A_{ij}\} \cup \bigcup_{i,j \in [1,n]} \text{Voc}(A_{ij})$$

où X et Y sont des concepts, roles ou individus et $r \in \{\sqsubseteq, \equiv, \perp, \in, =\}$, et $\text{Voc}(A_{ij})$ désigne le vocabulaire d'alignement, donc les links de A_{ij} .

Toutes les correspondances de l'alignement A_{ij} deviennent des formules de l'ontologie-alignement.

Définition 4.15 (Formules de l'ontologie-alignement) *Soit K un nœud de connaissances multi-niveaux, l'ensemble des formules de l'ontologie-alignement F_A peut être défini selon les cas suivants :*

- si K est une ontologie alors $F_A = \emptyset$;
- si K est un nœud de connaissances multi-niveaux composé de sous nœuds K_1, \dots, K_n et de A_{ij} l'alignement entre K_i et K_j pour $i, j \in [1, n]$ et trans est la fonction qui associe à toute correspondance de A_{ij} un axiome en logique de description (voir la définition 4.7). L'ensemble des formules de l'ontologie-alignement $F_A = \{f \mid f \in \text{trans}(A_{ij})\}$.

À partir des définitions présentées ci-dessus, nous pouvons actuellement définir l'ontologie-alignement.

Définition 4.16 (Ontologie-alignement) *Soit un nœud $K = \langle \{K_i\}, \{A_{ij}\} \rangle$ pour $i, j \in [1, n]$, K_i sont des nœuds locaux et A_{ij} est l'alignement qui lie K_i et K_j . On peut définir OntoAlign l'ontologie-alignement générée à partir de l'alignement A_{ij} de K , $\text{OntoAlign}(K) = \langle \Sigma_A, F_A \rangle$.*

Les passerelles d'un nœud de connaissances multi-niveaux sont des correspondances d'équivalence établies entre les termes de l'ontologie-alignement et les termes appartenant aux ontologies sources correspondants.

Définition 4.17 (Passerelles vers l'ontologie-alignement) *Soit K un nœud de connaissances. On*

définit les passerelles orientées vers l'ontologie alignement (noté $B(K)$) comme suit selon les cas :

- si K est une ontologie alors $B(K) = \emptyset$;
- si K est un nœud de connaissances composés des sous nœuds K_i et de l'alignement A_{ij} l'alignement entre K_i et K_j avec $i, j \in [1, n]$, alors $B(K)$ contient les passerelles définies comme suit, pour $i \in [1, n]$:
 - si K_i est une ontologie et X est un concept ou un rôle de K_i alors $i:X \xrightarrow{\equiv} \text{OntoAlign}(K):i:X \in B(K)$;
 - si K_i est une ontologie et a est un individu de K_i alors $i:a \xrightarrow{\equiv} \text{OntoAlign}(K):i:a \in B(K)$;
 - si K_i est un nœud composé et X est un concept ou rôle de $\text{OntoAlign}(K_i)$ alors $\text{OntoAlign}(K_i):X \xrightarrow{\equiv} \text{OntoAlign}(K):k_i:X \in B(K)$;
 - si K_i est un nœud composé et a est un individu de $\text{OntoAlign}(K_i)$ alors $\text{OntoAlign}(K_i):a \xrightarrow{\equiv} \text{OntoAlign}(K):k_i:a \in B(K)$.

Un réseau de connaissances multi-niveaux en DDL est composé de plusieurs nœuds locaux reliés à leurs ontologies d'alignement par le biais des passerelles.

Définition 4.18 (Réseau de connaissances multi-niveaux en DDL) *Soit K un nœud de connaissances. On définit le système en logique de description distribuée SystDis associé à K comme suit : $\text{SystDis}(K) = \langle \text{Onto}(K), \text{Bridge}(K) \rangle$ avec $\text{Onto}(K)$ est une famille des ontologies du système distribué DDL formé à partir du nœud K est qui est définie récursivement comme suit :*

- $\text{Onto}(K) = \{K\}$, si K est une ontologie DL ;
- $\text{Onto}(K) = \text{Onto}(K_1) \cup \text{Onto}(K_2) \cup \dots \cup \text{Onto}(K_n) \cup \text{OntoAlign}(K)$ si K est un nœud avec K_i nœuds locaux.

$\text{Bridge}(K)$ est l'ensemble des passerelles du système distribué DDL formé à partir de K et qui est définis récursivement comme suit :

- $\text{Bridge}(K) = \emptyset$ si K est une ontologie ;
- $\text{Bridge}(K) = \text{Bridge}(K_1) \cup \dots \cup \text{Bridge}(K_n) \cup B(K)$.

Exemple 11 Nous reprenons les ontologies, alignements de l'exemple 1 et nous détaillons la représentation du réseau de connaissances multi-niveaux en DDL. La figure 4.3 illustre cette représentation quand au détail des systèmes distribués en DDL conçus pour former le réseau est donné par la table 4.2.

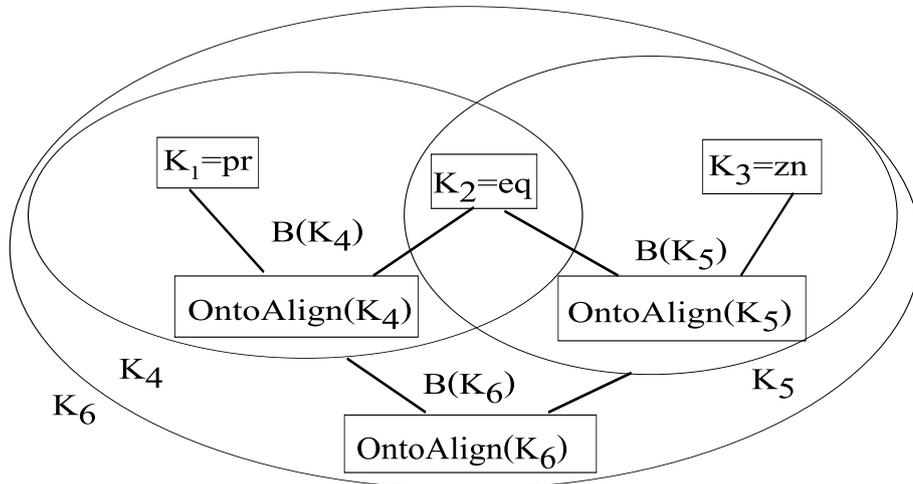


FIGURE 4.3 – Exemple d'un réseau de connaissances multi-niveaux en Logique de Description Distribuée

Soit un réseau de connaissances multi-niveaux $K = \langle \{O_1, O_2\}, \{A_{12}\} \rangle$, avec $A_{12} = \{1:A \xleftrightarrow{\sqsubseteq} 2:B, 1:a \xleftrightarrow{L} 2:b\}$ où A, B sont des concepts ou des rôles, a, b sont des individus et L est un *link*. On peut dire qu'une interprétation \mathcal{I} satisfait K si \mathcal{I} satisfait O_1 et O_2 et elle satisfait aussi les correspondances dans A_{12} . Pour interpréter K selon l'approche DDL, nous le transformons en un système distribué $\text{SystDis}(K)$ (voir la figure 4.4). avec :

- $b_1 = 1:A \xrightarrow{\equiv} 3:1:A$;
- $b_2 = 1:a \xrightarrow{\equiv} 3:1:a$;
- $b_3 = 2:B \xrightarrow{\equiv} 3:2:B$;
- $b_4 = 2:b \xrightarrow{\equiv} 3:2:b$;

b_1, b_2, b_3, b_4 sont des *bridges rules* et sont interprétées par les relations de domaines qui lient les interprétations locales correspondantes selon la sémantique de base DDL :

- $\mathcal{I}_1, \mathcal{I}_3 \models b_1$ si $r_{13}(A^{\mathcal{I}_1}) = 1 : A^{\mathcal{I}_3}$;
- $\mathcal{I}_1, \mathcal{I}_3 \models b_2$ si $r_{13}(a^{\mathcal{I}_1}) = 1 : a^{\mathcal{I}_3}$;

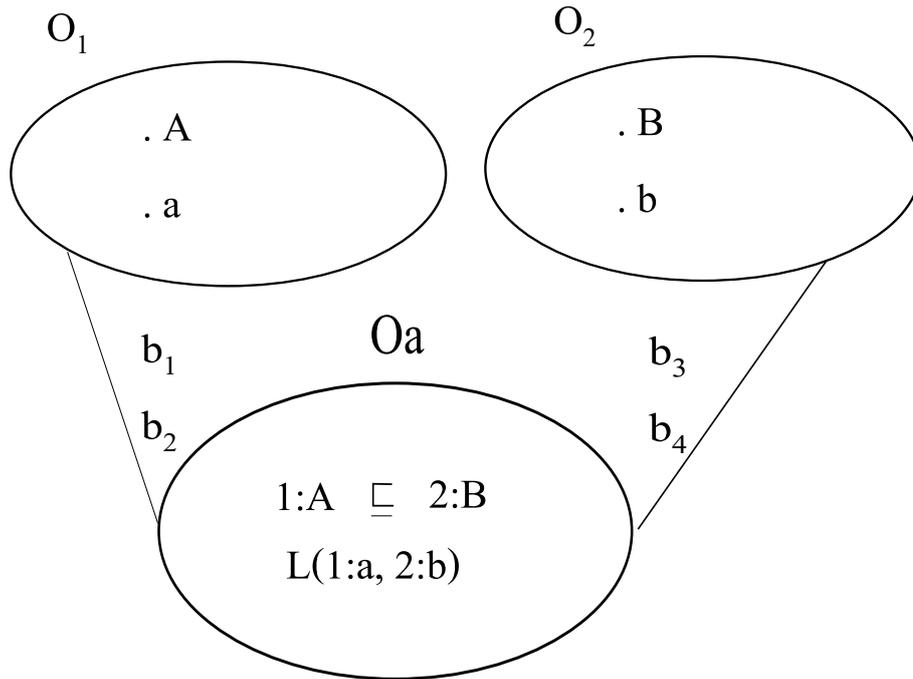


FIGURE 4.4 – $\text{SysDis}(K) = \langle \{O_1, O_2, O_a\}, \{b_1, b_2, b_3, b_4\} \rangle$

- $\mathcal{I}_2, \mathcal{I}_3 \models b_3$ si $r_{23}(B^{\mathcal{I}_2}) = 1 : B^{\mathcal{I}_3}$;
- $\mathcal{I}_2, \mathcal{I}_3 \models b_4$ si $r_{23}(b^{\mathcal{I}_2}) = 1 : b^{\mathcal{I}_3}$;

L'interprétation de K satisfait les correspondances de K (voir la figure 4.5) si :

- $\mathcal{I} \models 1:A \xleftrightarrow{\sqsubseteq} 2:B$ si $\mathcal{I}_1, \mathcal{I}_3 \models b_1$ et $(1 : A)^{\mathcal{I}_3} \subseteq (2 : B)^{\mathcal{I}_3}$ et $\mathcal{I}_2, \mathcal{I}_3 \models b_3$;
- $\mathcal{I} \models 1:a \xleftrightarrow{L} 2:b$ si $\mathcal{I}_1, \mathcal{I}_3 \models b_2$ et $L(1 : a^{\mathcal{I}_3}, 2 : b^{\mathcal{I}_3}) \in L^{\mathcal{I}_3}$ et $\mathcal{I}_2, \mathcal{I}_3 \models b_4$.

L'interprétation distribuée de K , $\mathcal{I} = \{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, r_{13}, r_{23}\}$ où $\mathcal{I}_1, \mathcal{I}_2$ sont les interprétations locales associées à O_1 et O_2 , \mathcal{I}_3 est l'interprétation de l'ontologie alignement générée et r_{13}, r_{23} sont des relations de domaine pour interpréter les bridges rules générées. On peut affirmer que \mathcal{I} satisfait le réseau de connaissances K dans l'approche DDL si \mathcal{I} satisfait $\text{SysDis}(K) = \langle \{O_1, O_2, O_a\}, \{b_1, b_2, b_3, b_4\} \rangle$ dans la sémantique de base DDL. En généralisant, nous pouvons dire qu'un réseau de connaissances multi-niveaux dans l'approche DDL est interprété comme un système distribué dans la sémantique de base DDL et l'interprétation distribuée du réseau de connaissances est une interprétation DDL du système distribué associé via la transformation SysDis.

Définition 4.19 (DDL-interprétation d'un nœud de connaissances) *Une DDL-interprétation*

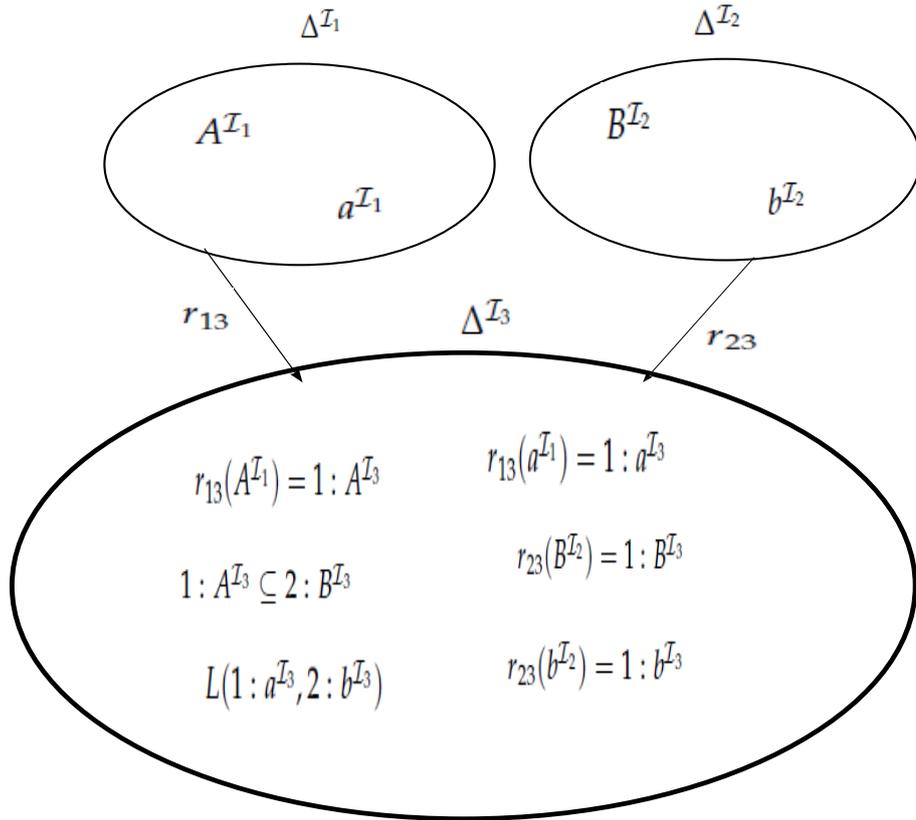


FIGURE 4.5 – L'interprétation du réseau K dans l'approche DDL

d'un nœud de connaissances X est une interprétation en DDL du système distribué de X ($\text{SystDis}(X)$).

L'interprétation distribuée est un modèle du réseau d'ontologies en DDL, si chaque interprétation locale est un modèle de son ontologie associée, et si la relation de domaine satisfait les passerelles.

Définition 4.20 (Relation de satisfaction dans un réseau de connaissances multi-niveaux en DDL)

Soit X un nœud de connaissances multi-niveaux et I une DDL-interprétation de X , alors I DDL-satisfait X (noté $I \models_{N-DDL} X$) si et seulement si I satisfait $\text{SystDis}(X)$ au sens de DDL (noté $I \models_{DDL} \text{SystDis}(X)$). Dans ce cas, on dit que I est un DDL-modèle de X .

Avec cette définition, on peut définir la notion d'implication de façon usuelle, à savoir que X DDL-implique Y si et seulement si tous les DDL-modèles de X sont des DDL-modèles de Y .

Propriété 4.2 Soit X un nœud de connaissances global comprenant des nœuds locaux

N_1, \dots, N_k , alors X DDL-implique tous ses nœuds locaux ($X \models_{N-DDL} N_i$) pour tout $i \in [1, k]$) mais l'inverse n'est pas vrai.

Preuve: Soit un nœud $X = \langle \{N_i\}, \{A_{ij}\} \rangle$. Démontrer que $X \models_{N-DDL} N_i$ revient à montrer que si une interprétation I est un DDL-modèle de X alors I est aussi un DDL-modèle de N_i . On suppose que $I \models_{N-DDL} X$. Ceci implique que $I \models_{DDL} \text{SystDis}(X)$ d'après la définition 4.20 d'où $I \models_{DDL} \langle \text{Onto}(X), \text{Bridge}(X) \rangle$ d'après la définition 4.18. Il s'en suit que $I \models_{DDL} \text{Onto}(N_1) \cup \text{Onto}(N_2) \cup \dots \cup \text{Onto}(N_n) \cup \text{OntoAlign}(X), \text{Bridge}(N_1) \cup \text{Bridge}(N_2) \cup \dots \cup \text{Bridge}(N_n)$ alors $I \models_{DDL} \text{SystDis}(N_i)$ pour $i = [1..n]$ par conséquent I est un DDL modèle de N_i CQFD

Remarque :

En se référant à l'exemple 11, $\text{SystDis}(K_6) = \langle \{K_1, K_2, K_3, oa_4, oa_5, oa_6\}, \text{Bridge}(K_6) \rangle$
 On remarque que $I \models_{DDL} I_1, I_2, I_4$, $I \models_{DDL} k_1:G_1 \xrightarrow{=} oa_4:k_1:G_1$ et $I \models_{DDL} k_2:DF_1 \xrightarrow{=} oa_4:k_2:DF_1$ alors I satisfait $\text{SystDis}(K_4)$ et cela implique que I est un modèle de K_4 . I satisfait aussi $\text{SystDis}(K_5)$ alors I est un modèle de K_5 . Donc on peut conclure que K_6 implique ses nœuds locaux K_4 et K_5 .

La transformation d'un nœud de connaissances multi-niveaux nous permet de rattacher ce dernier aux logiques de description distribuée qui dispose de l'outil de raisonnement opérationnel DRAGO (Serafini et Tamilin 2005) et qui peut être exploité pour le raisonnement sur un nœud en DDL.

4.3 COMPARAISON DES DEUX APPROCHES SÉMANTIQUES DU RÉSEAU DE CONNAISSANCES MULTI-NIVEAUX

Dans les paragraphes précédents, nous avons présenté et détaillé deux manières d'interpréter la sémantique du réseau de connaissances multi-niveaux à savoir l'approche DL et l'approche DDL. Dans cette section, nous allons essayer de comparer ces deux approches selon certains critères

afin de déterminer pour quels cas l'une est plus pertinente que l'autre (voir la table 4.3).

Nous commençons par étudier les deux approches pour chacun des critères puis un tableau récapitulatif de comparaison sera présenté par la suite.

Comparaison de consistance

Pour la consistance, on cherche à prouver que si un réseau de connaissances multi-niveaux est inconsistant dans l'approche DL, il ne l'est pas forcément dans l'approche DDL.

Théorème 4.1 *Il existe des réseaux de connaissances multi-niveaux qui sont DDL-consistants alors qu'ils sont DL-inconsistants.*

On peut prouver ce théorème en montrant que le réseau de connaissances multi-niveaux de l'exemple suivant est inconsistant selon la sémantique DL et il est consistant selon la sémantique DDL.

Exemple 12 Soit $O_1 = \{A_1 \sqsubseteq \neg B_1, A_1(a)\}$,
 $O_2 = \{A_2 \sqsubseteq B_3\}$,
 $A_{12} = \{1:A_1 \overset{\equiv}{\leftarrow} 2:A_2, 1:B_1 \overset{\equiv}{\leftarrow} 2:B_2\}$.

Lemme 4.1 *DL-consistance : Constituer une ontologie globale dont les éléments sont préfixés des ontologies sources et les correspondances sont transformés en axiomes $O_G = \{1:A_1 \sqsubseteq \neg 1:B_1, 1:A_1(1:a), 2:A_2 \sqsubseteq 2:B_3, 1:A_1 \equiv 2:A_2, 1:B_1 \equiv 2:B_2\}$ et qui*

représente un réseau de connaissances en DL.

$$1:A_1 \sqsubseteq \neg 1:B_1 \quad (4.1)$$

$$1:A_1(1:a) \quad (4.2)$$

$$2:A_2 \sqsubseteq 2:B_3 \quad (4.3)$$

$$1:A_1 \equiv 2:A_2 \quad (4.4)$$

$$1:B_1 \equiv 2:B_2 \quad (4.5)$$

$$1,4,5 \Rightarrow 1:A_2 \sqsubseteq \neg 1:B_2 \quad (4.6)$$

$$2,4 \Rightarrow 1:A_2(1:a) \quad (4.7)$$

$$6,7 \Rightarrow \neg 2:B_2(1:a) \quad (4.8)$$

$$7,3 \Rightarrow 2:B_2(1:a) \quad (4.9)$$

contradiction d'après (4.8) et (4.9) et ceci implique que O_G est DL-inconsistante.

Lemme 4.2 *DDL-consistance* : Reprenons le même exemple 12, construisons un système distribué S selon l'approche DDL, pour cela, il faut construire une ontologie alignement à partir des correspondances qu'on note O_{12} et il faut générer les bridges rules correspondantes \mathcal{B} . Nous obtenons une ontologie $O_{12} = \{1:A_1 \equiv 2:A_2, 1:B_1 \equiv 2:B_2\}$ et les bridges rules $\mathcal{B} = \{b_1, b_2, b_3, b_4\}$ où :

$$- b_1 = \{1:A_1 \xrightarrow{\equiv} 12:(1:A_1)\};$$

$$- b_2 = \{1:B_1 \xrightarrow{\equiv} 12:(1:B_1)\};$$

$$- b_3 = \{2:A_2 \xrightarrow{\equiv} 12:(2:A_2)\};$$

$$- b_4 = \{2:B_2 \xrightarrow{\equiv} 12:(2:B_2)\}.$$

Pour montrer que $S = \{O_1, O_2, O_{12}, \mathcal{B}\}$ est consistant, alors il faut trouver un modèle qui satisfait tous les axiomes et les bridges rules de S .

On suppose qu'il existe un modèle de S alors, il existe une interprétation distribuée $I = \{I_1, I_2, I_3, r_{12}, r_{13}, r_{23}, r_{21}, r_{31}, r_{32}\}$ tel que $I \models S$.

Ceci implique que :

$$- I_1 \models O_1;$$

$$- I_2 \models O_2;$$

$$- I_3 \models O_{12};$$

- $I_1, I_3, r_{13} \models b_1$;
- $I_1, I_3, r_{13} \models b_1$;
- $I_1, I_3, r_{13} \models b_2$;
- $I_2, I_3, r_{23} \models b_3$;
- $I_2, I_3, r_{23} \models b_4$;
- $I_1, I_3, r_{13} \models b_5$.

Ceci revient à montrer qu'il existe une interprétation $I = \{I_1, I_2, I_3, r_{13}, r_{23}\}$

tel que :

- $A_1^{I_1} \subseteq \neg B_1^{I_1} \subseteq \Delta_1$;
- $a^{I_1} \in A_1^{I_1}$;
- $A_2^{I_2} \subseteq B_2^{I_2} \subseteq \Delta_2$;
- $r_{13}(A_1^{I_1}) = (1:A_1)^{I_3}$;
- $r_{13}(B_1^{I_1}) = (1:B_1)^{I_3}$;
- $r_{23}(A_2^{I_2}) = (2:A_2)^{I_3}$;
- $r_{23}(B_2^{I_2}) = (2:B_2)^{I_3}$.

Considérons les domaines d'interprétation suivants : $\Delta_1 = \{1\}$, $\Delta_2 = \{2\}$,

$\Delta_3 = \{3\}$, et les fonctions d'interprétation définies comme suit :

- $A_1^{I_1} = \{1\}$;
- $a^{I_1} = 1$;
- $B_1^{I_1} = \emptyset$;
- $A_2^{I_2} = \emptyset$;
- $B_2^{I_2} = \emptyset$;
- $(1:A_1)^{I_3} = \emptyset$;
- $(1:B_1)^{I_3} = \emptyset$;
- $(2:A_2)^{I_3} = \emptyset$;
- $(2:B_2)^{I_3} = \emptyset$;
- $r_{13}(A_1^{I_1}) = \emptyset$;
- $r_{13}(B_1^{I_1}) = \emptyset$;
- $r_{23}(A_2^{I_2}) = \emptyset$;
- $r_{23}(B_2^{I_2}) = \emptyset$.

Donc pour $I = \langle (\{1\}, I_1), (\{2\}, I_2), (\{3\}, I_3), \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset \rangle$, on a $I \models S$, alors S est consistant.

Nous pouvons conclure que la manière dont on traite les alignements dans l'approche DL est fixée et permet ainsi une réconciliation des ontologies locales et des alignements. Cette approche peut être utilisée dans le cas des ontologies modulaires où chaque module fait partie d'un point de vue global dans un domaine plus large. Elle présente des limites s'il s'agit du Web, où les ontologies peuvent avoir des points de vues contradictoires. Par contre l'approche DDL permet de faire cohabiter des points de vue différents même s'ils sont incompatibles.

Complexité de transformation

- La complexité de transformation du réseau de connaissances multi-niveaux en une ontologie DL est linéaire au niveau des termes des ontologies, et peut être calculée en fonction du nombre d'opérations de prefix des termes des ontologies, des links et du nombre de création des axiomes à partir des correspondances.
- La complexité de transformation en un système distribué DDL est calculée en fonction du nombre d'opérations de création des axiomes dans l'ontologie alignement à partir des correspondances et du nombre d'opérations de création des bridge rules. Rappelons que pour une correspondance, on trouve deux termes (le terme à droite et le terme à gauche de la correspondance) et pour chaque terme, on crée une bridge rule.

La complexité de transformation dans le cas de la mise à jour des ontologies locales pour l'approche DL est proportionnelle à l'importance de la mise à jour. Ceci est dû au fait que la mise à jour des ontologies locales conduit à la reconstitution des termes de l'ontologie globale. Pour l'approche DDL, la mise à jour des ontologies locales n'affecte pas la transformation, on peut dire qu'elle est transparente.

Nous pouvons conclure que dans le cas où les ontologies locales sont

évolutives et dans le cas où elles sont plus importantes que les correspondances, l'approche de DDL est plus intéressante que l'approche DL.

Complexité de raisonnement

la complexité de raisonnement des deux approches se basent sur les complexités de raisonnement des sémantiques DL et DDL. Le raisonnement sur un réseau de connaissances doit être décidable tout en réduisant la complexité des inférences. C'est pour cette raison qu'il faut trouver un compromis entre l'expressivité de la logique et la complexité de celle-ci.

Le réseau de connaissances multi-niveaux tel qu'il a été présenté dans la section 4.2.1 est une ontologie en logique de description construite à partir d'une fusion des ontologies locales dont les termes ont été préfixés et les alignements transformés en axiomes.

Sachant que les ontologies locales peuvent être formalisées en logiques différentes, que l'expressivité des axiomes d'origine alignement est très simple et que ces axiomes peuvent être formalisés dans le langage \mathcal{EL} décidable et dont la complexité est **NPcomplete**, alors, la décidabilité et la complexité du réseau de connaissances multi-niveaux interprété en DL peuvent être données par l'étude de la décidabilité et la complexité de la fusion des logiques de description des ontologies locales et de la logique des axiomes intégrés.

Cette étude n'a pas été traitée dans cette thèse, mais nous pouvons utiliser les résultats fournis dans l'article de Baader et al. (2002). Tout d'abord, ce travail montre que la fusion de deux logiques de description est un fragment de l'union de ces dernières car le raisonnement sur l'union des deux logiques nécessite l'implémentation d'une nouvelle méthode de raisonnement, mais le raisonnement sur la fusion peut être réduit au raisonnement sur les logiques composantes. En plus le raisonnement sur l'union de deux logiques décidables peut être indécidable alors que le raisonnement de la fusion des mêmes logiques reste décidable.

À titre d'exemple, l'union de la logique \mathcal{ALCF} (qui est une extension

Approche	point de vue	Impact de la mise à jour des ontologies	Implémentation du raisonneur	Complexité de transformation	Complexité de raisonnement
DL	points de vue compatibles	conduit à la mise à jour du réseau	prend en charge plusieurs logiques	augmente avec l'importance des ontologies et des links	= complexité de la fusion des logiques composantes
DDL	points de vue incompatibles	transparente par rapport au réseau	un raisonneur par logique	augmente avec l'importance des alignements	= complexité la plus élevée entre complexité des logiques locales et complexité des inférences des mappings

TABLE 4.3 – Tableau comparatif

de \mathcal{ALC} par l'ajout de rôles fonctionnels) et de la logique $\mathcal{ALC}^{+,o,\cup}$ (qui est une extension de \mathcal{ALC} par l'ajout de la transitivité, la composition et l'union de rôles), est indécidable, alors que leur fusion est décidable.

Concernant la complexité, d'après le même article, la complexité de la fusion des logiques de description dont la complexité **Pspace** est également **Pspace**. Ceci n'est pas valable pour l'union de ces logiques. À titre d'exemple, la complexité de l'union de la logique \mathcal{ALCFOQ} (qui est une extension de \mathcal{ALC} par l'ajout de rôle fonctionnel, des nominaux et de la restriction de nombre qualifié) et de la logique \mathcal{ALCI} (qui est une extension de \mathcal{ALC} par l'ajout de rôle inverse) est **NExpTime** alors que la complexité des logiques composantes est **Pspace**.

C'est différent pour l'approche DDL, où les logiques ne sont pas fusionnées mais connectées par des relations. Ghidini et al. (2008) présentent une étude montrant que l'inférence sur les *mappings* est décidable et la complexité est comprise entre **ExpTime** et **2ExpTime**.

Nous pouvons conclure que la complexité du réseau de connaissances multi-niveaux interprété en DDL peut être égale à la complexité la plus élevée parmi celles des ontologies locales et des inférences sur les mappings.

4.4 DISCUSSION

La gestion indépendante des alignements nous donne la possibilité d'augmenter l'expressivité du langage d'alignement et de réaliser des opérations intéressantes sur les alignements telle que la composition des alignements. Cette partie n'a pas été traitée dans cette thèse, mais on peut s'inspirer de la proposition de Zimmermann (2008) afin de faire évoluer l'expressivité du langage d'alignement associé au réseau de connaissances multi-niveaux et définir la composition des *mappings* et des *links* à travers les alignements existants. Parmi les sémantiques proposées par les travaux précédents, nous avons remarqué qu'aucune sémantique ne peut être adoptée directement pour interpréter notre formalisme. Le formalisme \mathcal{E} -connection supporte les *links*, il propose une interprétation des *links* mais selon le point de vue de l'ontologie cible de la correspondance, en plus, il ne tolère pas l'hétérogénéité entre les *links* car il ne supporte pas les *mappings*, donc il y a absence d'interprétation pour ce type de correspondance. Le formalisme P-DL interprète les entités importées en tant que termes étrangers, donc pas d'interprétation pour les alignements. Le point commun de notre formalisme avec le formalisme IDDL est que les correspondances sont exprimées selon un point de vue global, il propose une sémantique adéquate à cette formalisation, car il propose une interprétation des alignements dans un domaine d'interprétation indépendant. Mais cette sémantique ne peut être exploitée directement car il n'y a pas d'interprétation proposée pour les *links*. Le formalisme $E - SHIQ$ supporte les *mappings* et les *links*, il propose une interprétation pour les deux types de correspondances mais ils sont interprétés selon le point de vue de l'ontologie cible des correspondances de la même façon que les *mappings* dans DDL et les *links* dans \mathcal{E} -connection. Néanmoins, nous avons proposé d'interpréter notre réseau de connaissances multi-niveaux selon deux sémantiques existantes possibles, mais ceci ne peut être réalisé qu'après une transformation adéquate du réseau. La première approche concerne la sémantique classique des logiques de description et on a vu

que cette approche est utilisée dans le cas où les ontologies et les alignements sont développés dans des contextes différents à condition que leurs points de vue soient compatibles. La deuxième approche proposée adopte une sémantique distribuée qui tolère l'incompatibilité des points de vues des ontologies et des alignements à intégrer. Cette approche utilise la sémantique DDL pour interpréter le réseau de connaissances multi-niveaux, mais après sa transformation en un système distribué. On a vu que DDL ne supporte pas les *links*, donc il ya absence d'interprétation de ces derniers et les *mappings* sont interprétés du point de vue de l'ontologie cible de la correspondance. L'idée donc est de générer une ontologies alignements qui permettra d'interpréter les alignements. Des règles passerelles dans le cadre DDL entre les entités les ontologies locales et les entités correspondantes dans l'ontologie alignement sont générées également. Ces dernières sont interprétées par des relations de domaine comme le définit la sémantique DDL.

CONCLUSION DU CHAPITRE

Dans ce chapitre, nous avons détaillé le formalisme qu'on a proposé pour représenter un réseau d'ontologies alignées à plusieurs niveaux auquel nous avons associé deux sémantiques possibles. Dans le prochain chapitre, nous présentons, deux prototypes pour son implémentation selon ses sémantiques.

IMPLÉMENTATION ET EXPÉRIMENTATION

SOMMAIRE

5.1	IMPLÉMENTATION ET EXPÉRIMENTATION DE L'APPROCHE DL	99
5.1.1	Architecture proposée	99
5.1.2	Format d'alignement adopté	100
5.1.3	Description du format proposé	100
5.1.4	Expérimentation et évaluation	105
5.1.5	Discussion	107
5.2	IMPLÉMENTATION DE L'APPROCHE DDL	112
5.2.1	Drago	112
5.2.2	Implémentation de l'approche DDL	114

DANS ce chapitre, nous décrivons dans un premier lieu l'implémentation du premier prototype pour le raisonnement sur le réseau de connaissances multi-niveaux pour l'approche DL présentée dans la section 4.2.1 (Klai et al. 2016b). Dans un second lieu, nous décrivons l'implémentation du deuxième prototype pour l'approche DDL présentée dans la section 4.2.2.

L'API qu'on a nommé "Description Logic for Multi-Level Networked Reasoner" (DLMLNKR) permet le raisonnement sur les connaissances de

ce type de réseau, soit pour vérifier la consistance, tester la satisfaction des axiomes ou pour l'inférence de nouvelles connaissances en utilisant les raisonneurs Pellet ou Hermit selon le choix de l'utilisateur.

5.1 IMPLÉMENTATION ET EXPÉRIMENTATION DE L'APPROCHE DL

5.1.1 Architecture proposée

DLMLNKR accepte en entrée des ontologies en OWL/RDF, des alignements dans plusieurs niveaux stockés dans des fichiers RDF sous le format API Alignment que nous avons étendu pour stocker les *links*. Nous commençons par définir les composants de DLMLNKR (voir la figure 5.1) :

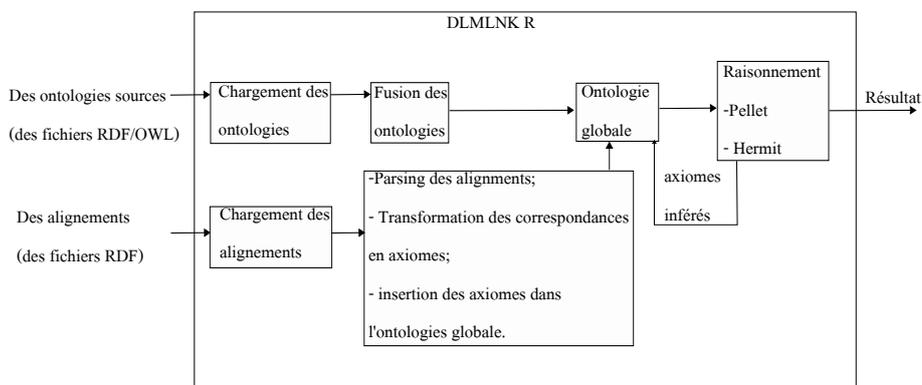


FIGURE 5.1 – Architecture du raisonneur DLMLNKR

- Chargeur des ontologies sources en spécifiant leurs localisations ou leurs IRI ;
- Chargeur des alignements qui peuvent être effectués entre une paire d'ontologies (alignement inter-ontologies) ou entre une paire d'alignements (il s'agit donc d'alignement de niveau supérieur) : cette opération consiste à charger des alignements stockés dans des fichiers RDF résultats des outils de **matching** disponibles sur le Web. Les alignements peuvent être enrichis d'une manière semi automatique par des **links**. Les outils d'alignement peuvent être étendus pour l'insertion des **links** ou faire un appariement des fichiers alignement ;
- Module de fusion des ontologies sources et produire une ontologie globale ;
- Parseur des alignements, ceci consiste à extraire les **links** et les **map-**

pings. Les **links** sont transformés en rôles spécifiques et les **map-pings** en axiomes puis on procède à l'insertion de ces axiomes dans l'ontologie globale ;

- Raisonnement pour le test de la consistance de l'ontologie globale ainsi que tester la satisfaction des axiomes ;
- Raisonnement pour l'inférence de nouvelles connaissances permettant ainsi la composition des alignements transformés en axiomes ;
- Générateur des nœuds dans le but de les sauvegarder pour d'éventuels alignements avec d'autres ontologies ou d'autres réseaux.

5.1.2 Format d'alignement adopté

Nous utilisons le format d'alignement proposé dans l'API Alignment dont les auteurs se sont inspirés du format général de l'alignement pour des raisons de consensualité. Initialement, ce format est décrit en XML, puis il a été transformé en RDF pour supporter les correspondances entre les ontologies en OWL. L'espace nom utilisé dans ce format est "<http://knowledgeweb.semanticweb.org/heterogeneity/alignment#>".

5.1.3 Description du format proposé

Les éléments de l'alignement sont les suivants :

– **xml**

Cet attribut prend pour valeur ("yes"/"no"), ce qui indique si l'alignement peut être lu en tant que fichier XML et compilé avec une DTD ;

– **onto1**

indique l'ontologie à gauche du *mapping* ;

– **onto2**

indique l'ontologie à droite du *mapping* ;

– **map**

Cet attribut prend pour valeur (Cell) et qui signifie une correspondance de type *mapping* ;

– **link**

Cet attribut prend pour valeur (Celll) et qui signifie une correspondance de type *link*.

Nous tenons à préciser que l'élément (Celll) n'est pas présent dans le format d'alignement de l'API d'origine et que nous avons rajouté pour stocker les links.

```
<?xml version='1.0' encoding='utf-8' standalone='no'?>
<rdf:RDF xmlns="http://alignment#"
  xmlns:align="http://alignment#"
  xmlns:rdf="http://www.w3.org/-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/XMLSchema#">
<Alignment>
<alignment IRI = "http://alignment-12.rdf#" />
  <xml>yes</xml>
  <onto1>....</onto1>
  <onto2>....</onto2>
    <map>.....<\map>
    <link>.....<\link>
</Alignment>
</rdf:RDF>
```

L'élément *ontology* contient trois attributs :

- **rdf:about** contient l'URI qui identifie l'ontologie ;
- **location** contient l'URL correspondant à la localisation de l'ontologie ;
- **formalism** décrit le nom et l'URI du langage dont l'ontologie est exprimée.

```

<onto1>
<Ontology
  rdf:about="http://www.univ-annaba.com/pr.owl\#">
<location>D:/pr.owl</location>
<formalism>
<Formalism
  align:name="OWL2.o" align:uri="http://.../owl\#" />
</formalism>
</Ontology>
</onto1>
<onto2>
<Ontology
  rdf:about="http://www.univ-annaba.com/eq.owl\#">
<location>D:/eq.owl</location>
<formalism>
<Formalism
  align:name="OWL2.o" align:uri="http://.../owl\#" />
</formalism>
</Ontology>
</onto2>

```

Dans le cas d'un alignement inter-alignements, l'élément *ontology* contient dans l'attribut **rdf:about** l'URI qui identifie l'alignement situé à gauche ou à droite de la correspondance, l'attribut **location** contient l'URL correspondant à la localisation de l'alignement et l'attribut **formalism** décrit le nom et l'URI du format dont l'alignement est exprimé.

```

<Alignment>
<alignment
  IRI = "http://heterogeneity/alignment-12-23.rdf" />
<xml>yes</xml>

```

```

<onto1>
<Ontology
rdf:about="http://heterogeneity/alignment-12.rdf\#">
<location>D:/alignment-12.rdf</location>
<formalism>
<Formalism align:name="API-Alignment"
align:uri="http://heterogeneity/alignment\#" />
</formalism>
</Ontology>
</onto1>
<onto2>
<Ontology
rdf:about="http://heterogeneity/alignment-23.rdf\#">
<location>D:/alignment-23.rdf</location>
<formalism>
<Formalism align:name="API-Alignment"
  align:uri="http://heterogeneity/alignment\#" />
</formalism>
</Ontology>
</onto2>

```

L'élément **Cell** identifie une correspondance de type *link* et il est défini par les attributs suivants :

- **entity1** contient l'URI de l'entité située à gauche de la correspondance;
- **entity2** contient l'URI de l'entité située à droite de la correspondance; les entités peuvent appartenir à une ontologie source ou à un alignement.
- **relation** contient le **link** de la correspondance.

```

<link>
<Cell>

```

```

<entity1
rdf:resource="http://www.univ-annaba.com/pr.owl\#G1"/>
<entity2
rdf:resource="http://www.univ-annaba.com/eq.owl\#DF1"/>
<relation
IRI="http://knowledgeweb../alignment-12.rdf\#compose"/>
</Cell>
</link>

```

Et enfin l'élément **Cell** identifie une correspondance de type *mapping* et il est défini par les attributs suivants :

- **entity1** contient l'URI de l'entité située à gauche de la correspondance ;
- **entity2** contient l'URI de l'entité située à droite de la correspondance ;
- **relation** contient la relation qui existe entre les deux entités. Cette relation peut être définie par les symboles suivants : > pour une relation (subsumes), < pour la relation (is subsumed), = pour (équivalent), % pour (incompatible), HasInstance et InstanceOf.

```

<map>
<Cell>
<entity1
rdf:resource="http://.../alignment-12.rdf\#compose"/>
<entity2
rdf:resource="http://alignment-23.rdf\#fait-partie-de"/>
<relation>=</relation>
<measure
rdf:datatype='http://XMLSchema\#float'>1.0</measure>
</Cell>
</map>

```

5.1.4 Expérimentation et évaluation

Nous avons effectué des tests sur des ontologies complexes et des alignements respectifs (benchmark) téléchargeables au niveau du site "http://oaei.ontologymatching.org/2014/". Nous avons enrichi les alignements d'une manière semi automatique par des links et les correspondances entre les alignements ont été insérées manuellement, mais un travail est en cours de réalisation pour étendre l'outil d'appariement **XMAP++** (Djeddi et Khadir 2013) afin de prendre en charge les links et aligner les fichiers RDF sous le format des alignements. Le tableau 5.1 décrit la taille des ontologies, les alignements inter-ontologies et les alignements inter-alignements constituant notre réseau de connaissances multi-niveaux.

Les tests ont été réalisés sur trois typologies possibles pour le raisonnement sur un réseau de connaissances multi-niveaux constitué de trois ontologies (O_1, O_2, O_3), des alignements inter-ontologie (A_{12}, A_{23}) et un alignement inter-alignement (A_{12-23}) à titre d'exemple. La première typologie (T_1) consiste à charger l'ensemble des ontologies et des alignements en une seule opérations puis raisonner sur l'ensemble. La seconde typologie (T_2) consiste à créer un nœud par paire d'ontologies avec l'alignement respectif, puis raisonner sur le réseau constitué de la paire des nœuds avec leur alignement. Et la troisième typologie (T_3) consiste à raisonner sur un nœud déjà construit et aligné avec la troisième ontologie. Le détail des trois typologies est illustré par la figure 5.2 .

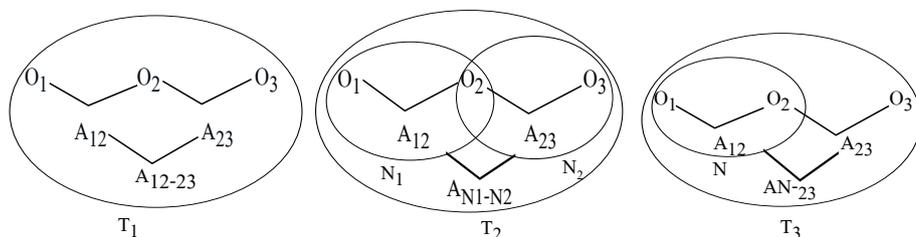


FIGURE 5.2 – Les typologies possibles pour un réseau multi-niveaux de trois ontologies et les alignements respectifs

L'évaluation concerne le temps écoulé pour l'accomplissement des deux opérations principales qui sont :

Exemple	Ontologies/alignements	Taille
Anatomy	Human	3.27 MB
	mouse	1.33 MB
	$A_{human-mouse}$	371 kB
Conference	cmt	27.8kB
	conference	38.4kB
	confOf	36.5kB
	edas	100kB
	ekaw	35.1kB
	iasted	70kB
	sigkdd	15.1 kB
	Alignment	65.76kB
Benchmarks	101	71.5kB
	103	80.4kB
	104	46.1kB
	$A_{101-103}$	44.4kB
	$A_{101-104}$	49.3kB
	$A_{101-103-101-104}$	4.45kB

TABLE 5.1 – Taille ontologies/alignements

- La transformation du réseau de connaissances multi-niveaux en une ontologie globale détaillée selon le temps de la fusion des ontologies sources **fus** et le temps de parsing et de transfert des alignements en axiomes ainsi que l'insertion de ces derniers dans l'ontologie globale. on note **pars** A_1 pour le parsing de l'alignement A_{12} , **pars** A_2 pour le parsing de l'alignement A_{23} et **pars** A_3 pour le parsing de l'alignement A_{12-23} ;
- Le test de consistance de l'ontologie résultante (**Consist**).

Impact de la taille des ontologies et des alignements sur la transformation du réseau multi-niveaux et le test de consistance Les tests ont été effectués sur les ontologies figurant dans l'exemple de "Anatomy" (voir Table 5.1). Afin de mesurer l'impact de la taille, nous avons conduit en premier lieu les tests sur approximativement les 25%, puis les 50% et enfin les 100% des ontologies "human", "mouse" et les alignements $A_{human-mouse}$. Les résultats des tests sont présentés dans la table 5.2.

L'impact du nombre d'ontologies et d'alignements dans le réseau sur le transfert du réseau et sur le test de consistance de l'ontologie DL résultat

Les tests ont été effectués sur l'exemple de "Conference" (voir table 5.1). Et pour mesurer l'impact du nombre des ontologies et des alignements, nous avons construit le réseau avec deux ontologies pour le premier cas, et pour chaque cas, nous introduisons dans le réseau une nouvelle ontologie ses alignements avec le reste des ontologies déjà intégrées dans le réseau. Nous pouvons consulter les résultats dans la table 5.3

L'impact de niveaux et par typologie Les tests ont été effectués sur les ontologies de "benchmark" avec les alignements correspondants en adoptant les typologies possibles (voir la figure 5.2). La première typologie (T_1) consiste à charger l'ensemble des ontologies ainsi que les alignements en une seule opération puis raisonner sur l'ensemble complet. La seconde typologie consiste à créer un nœud par paire d'ontologies et d'alignements puis raisonner sur le réseau construit. La troisième typologie consiste à raisonner sur le nœud de connaissance construit et aligné à la troisième ontologie et les résultats sont donnés dans la table 5.4.

5.1.5 Discussion

Les résultats fournis dans la table 5.2 montrent que la taille des ontologies et des alignements n'a pas d'impact ni sur le transfert du réseau pour une interprétation en logique de description ni sur le test de consistance de l'ontologie en DL. Il est clair que l'importance de la taille fait augmenter le temps de transfert et le temps de test de consistance. Notons que c'est le même cas pour l'impact du nombre d'ontologies et d'alignements (voir les résultats dans la table 5.3). Par contre, la complexité des ontologies influence largement le test de consistance, c'est le cas 5 et le cas 6 où l'intégration de l'ontologie 'iasted' malgré que sa taille est plus petite que la taille de l'ontologie 'edas', mais le temps de test de consistance est plus élevé. Ceci se confirme dans le cas 7 avec les ontologies cmt, iasted

Ontologies /alignements	fus (ms)	pars (ms)	Total (ms)	Consist (ms)
$\simeq 25\%$ Human $\simeq 25\%$ mouse $\simeq 25\% A_{human-mouse}$	152	248	400	332
$\simeq 50\%$ Human $\simeq 50\%$ mouse $\simeq 50\% A_{human-mouse}$	293	490	783	540
100% Human 100% mouse 100% $A_{human-mouse}$	570	821	1391	1290

TABLE 5.2 – Exemple de l'impact de taille pour un réseau de connaissances à un niveau

case	Ontologies /alignements	fus (ms)	pars (ms)	Total	Consist (ms)
case1	cmt, conference, $A_{cmt-conference}$	30	50	80	320
case2	cmt, conference, confOf	40	130	170	330
case3	cmt, conference, confOf, edas,	60	260	320	440
case4	cmt, conference, confOf, edas, ekaw	70	882	892	450
case5	cmt, conference, confOf, edas, ekaw, iasted	80	917	997	3090
case6	cmt, conference, confOf, edas, edas, ekaw, iasted, sigkdd	130	1101	1231	3110
case7	cmt, iasted, $A_{cmt-istaded}$	60	80	140	2280
case8	cmt, edas, $A_{cmt-ekaw}$	40	60	100	360

TABLE 5.3 – Exemple d'impact de nombre d'ontologies et d'alignements dans le réseau

et l'alignement $A_{cmt-istaded}$ puis, le cas 8 avec les ontologies cmt, edas et l'alignement $A_{cmt-edas}$. La table 5.4 montre que le temps de consistance est approximativement égal pour les différents niveaux du réseau de connaissances (niveau 0, niveau 1 et niveau 2), par contre, le temps de transfert en DL-ontologie progresse avec l'évolution du niveau, donc il n'y a pas d'impact sur le niveau. Le changement de typologies a un impact sur le transfert du réseau. Les résultats fournis dans la table 5.4 montre que la typologie 2 (T2) donne de meilleurs résultats point de vue temps de transfert d'un réseau à un niveau et à deux niveaux, mais ce n'est pas le cas pour le test de consistance, notons que la typologie 3 est la meilleure point de vue test de consistance.

	fus(ms)	pars(ms) (101-103)	pars(ms) (101-104)	pars(ms) (101-103 -101-104)	Total(ms)	Consist(ms)
niveau 0						
T_1	(101, 103, 104) = 70				70	460
T_2	(101, 103) = 50 (101, 104) = 40 (N_1, N_2) = 80				170	460
T_3	(101, 103) = 50 ($N_1, 104$) = 70				120	440
niveau 1						
T_1	(101, 103, 104) = 70	950	1030		2050	470
T_2	(101, 103) = 50 (101, 104) = 40 (N_1, N_2) = 80	430	470		1070	450
T_3	(101, 103) = 50 ($N_1, 104$) = 60	430	640		1180	450
niveau 2						
T_1	(101, 103, 104) = 100	600	1050	1000	2650	450
T_2	(101, 103) = 50 (101, 104) = 40 (N_1, N_2) = 80	430	470	70	1140	460
T_3	(101, 103) = 50 ($N_1, 104$) = 60	430	660	590	1790	400

TABLE 5.4 – Exemple d'impact des niveaux et des typologies

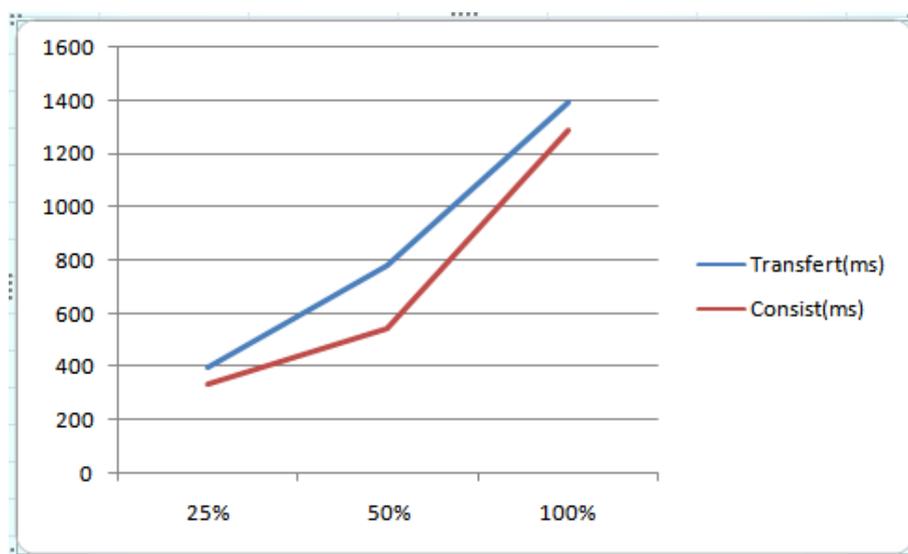


FIGURE 5.3 – Exemple d'impact de taille du réseau d'ontologies

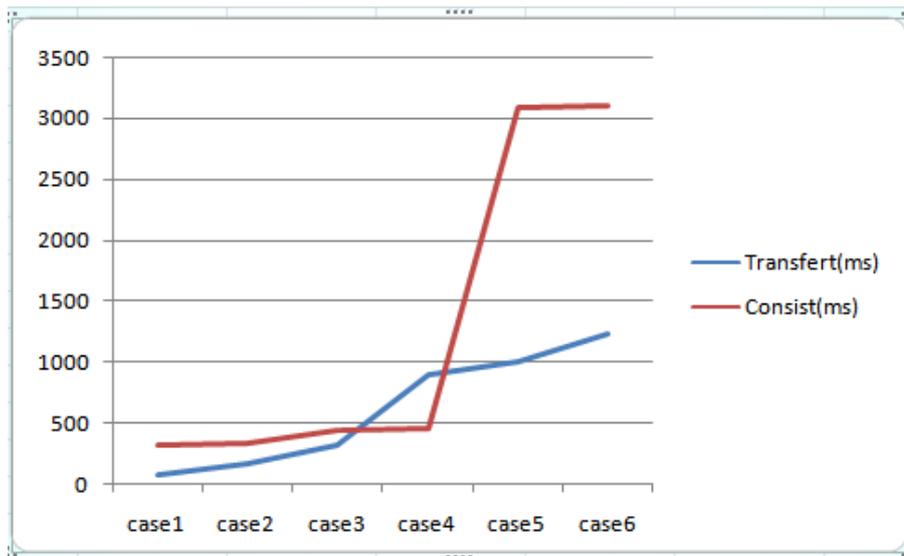


FIGURE 5.4 – Exemple d'impact de nombre d'ontologies

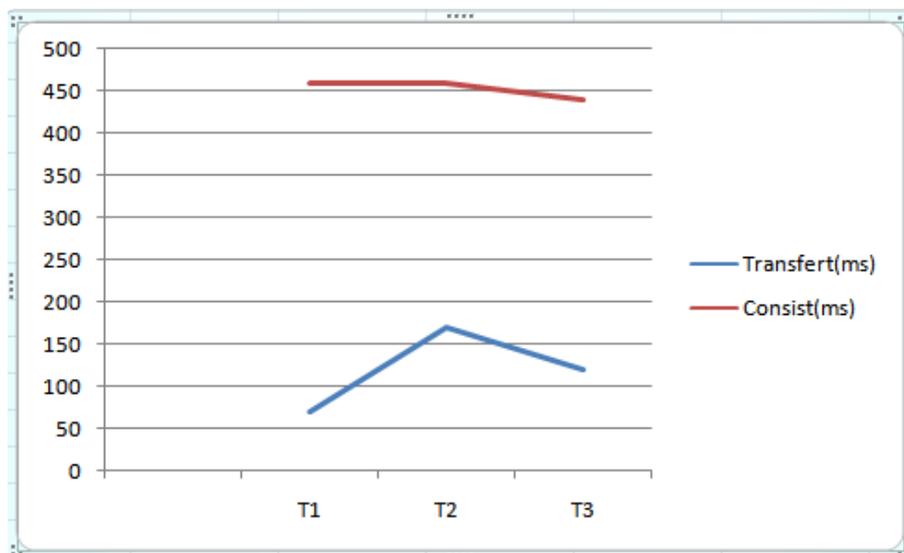


FIGURE 5.5 – Exemple d'impact de typologies pour le niveau0

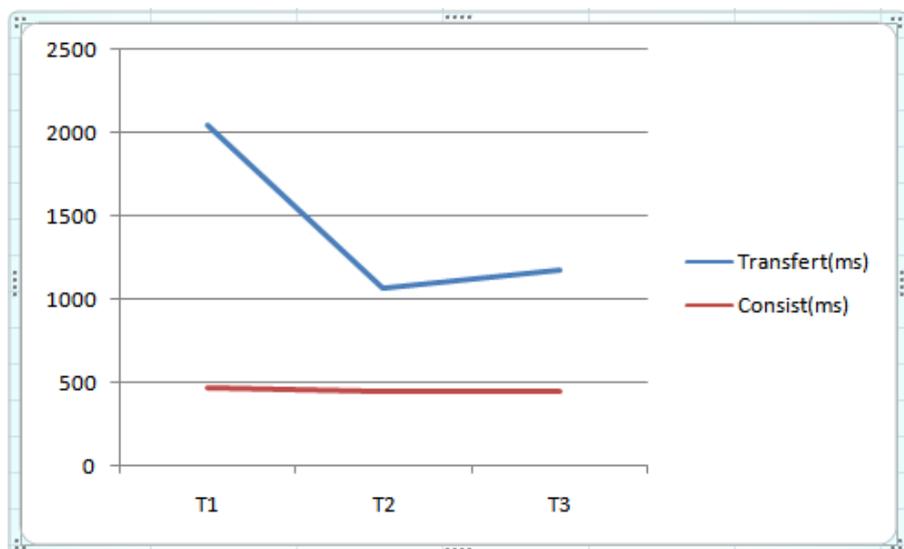


FIGURE 5.6 – Exemple d'impact de typologies pour le niveau1

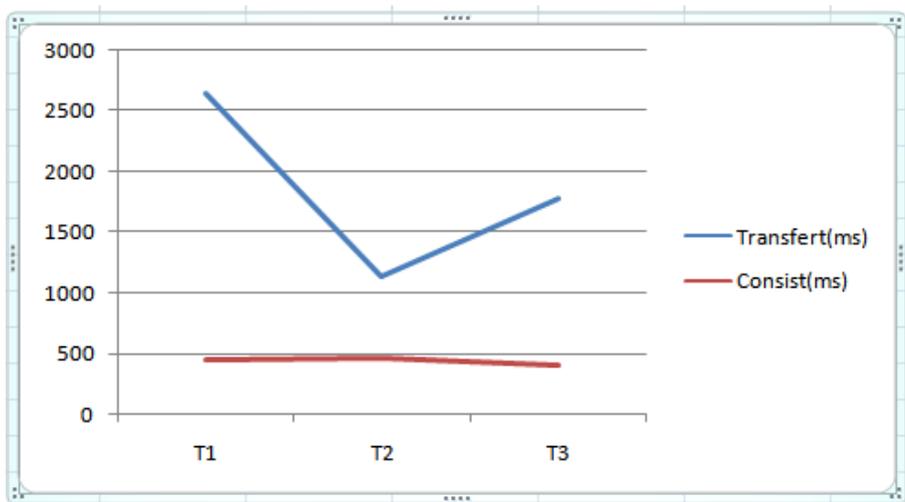


FIGURE 5.7 – Exemple d'impact de typologie pour le niveau 2

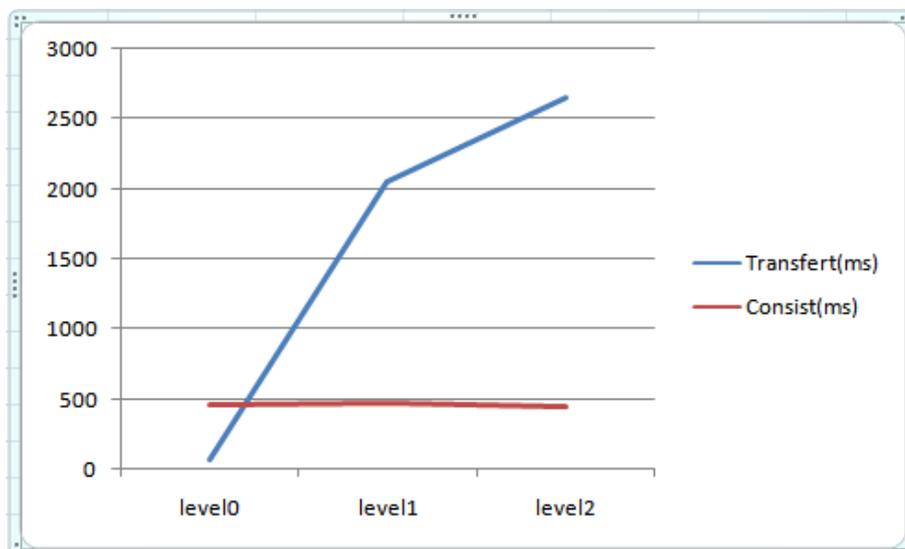


FIGURE 5.8 – Exemple d'impact de niveau pour la typologie 1

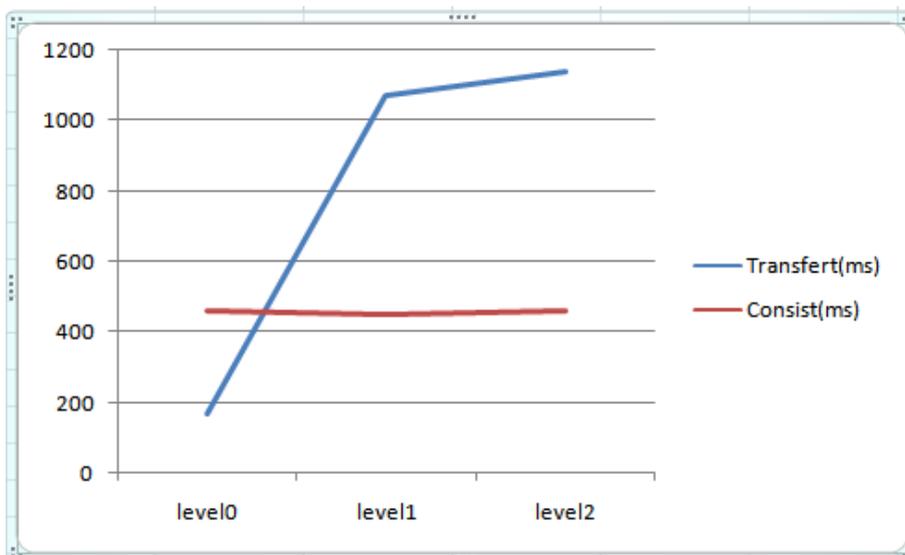


FIGURE 5.9 – Exemple d'impact de niveau pour la typologie 2

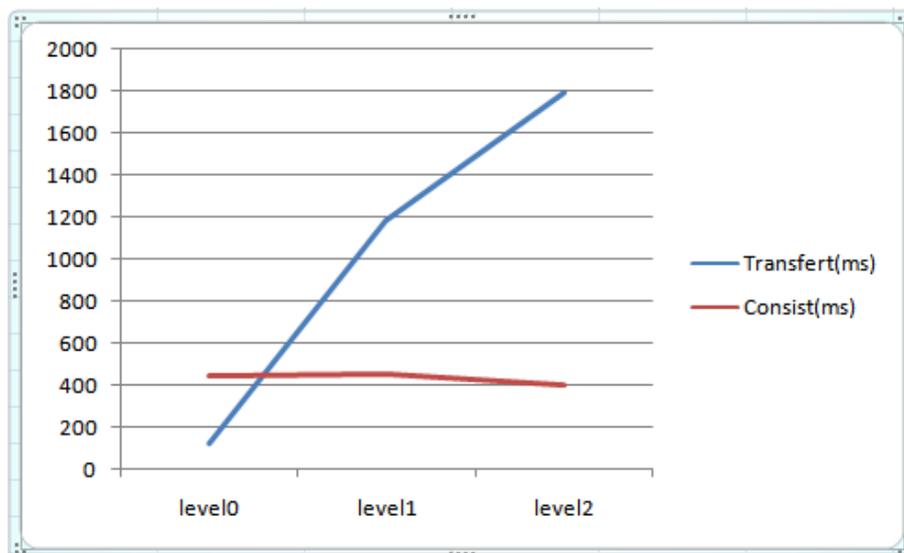


FIGURE 5.10 – Exemple d'impact de niveau pour la typologie 3

5.2 IMPLÉMENTATION DE L'APPROCHE DDL

5.2.1 Le raisonneur distribué Drago

(Distributed Reasoning Architecture for a Galaxy of Ontologies) c'est un outil qui raisonne sur un ensemble d'ontologies où chaque paire d'ontologies est interconnectée par des mappings. Chaque DRP (Drago Reasoning Peers) concerne une ontologies cible, les mappings et les ontologies sources correspondants. Il utilise des raisonneurs locaux tels que (Fact, Racer, Pellet,..) pour un raisonnement local. Il accepte en entrée des ontologies en OWL et les mappings en C-OWL.

Architecture de Drago

La figure 5.11 montre les composants principaux de Drago :

- **Registration service** : permet d'enregistrer les détails de chaque DRP :
 - URL de l'ontologie cibles ;
 - URLs des mappings attachés à l'ontologie cible ;
 - pour chaque ontologie étrangère qui figure dans les mappings, les adresses des DRPs les accueillant sont spécifiées afin de lier le

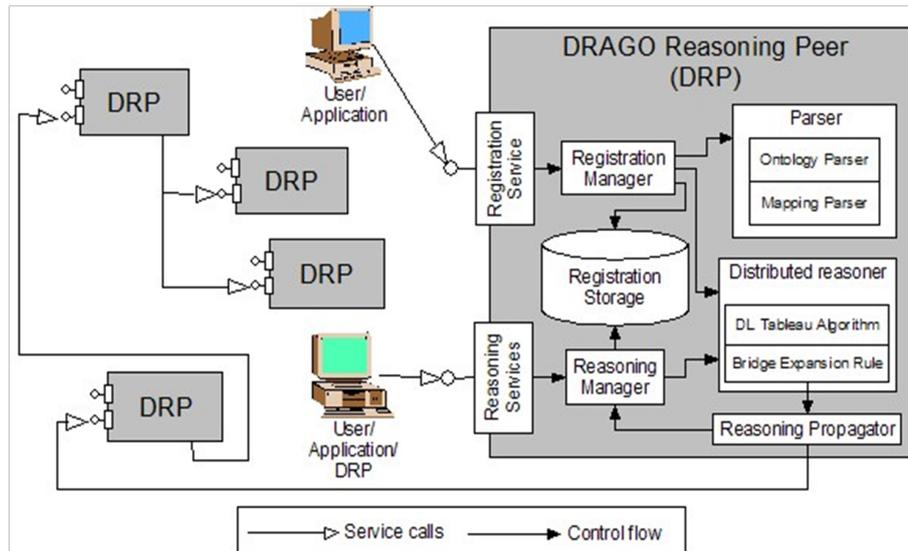


FIGURE 5.11 – Architecture de Drago raisonnement par paire

DRP courant avec les autres paires pour établir le raisonnement distribué ;

- le parseur translate les ontologies locales et les mappings sous format de T-box et M-box pour constituer la base de connaissances distribuée.
- **Reasoning service** : Le raisonnement sur le DRP se base sur un raisonnement standard (algorithme des tableaux) modifié avec l'ajout des règles de complétion pour permettre le raisonnement à travers les mappings. La modification permet au raisonneur de fonctionner selon deux modes : local et distribué :
 - le mode local : ignore les mappings utilisant l'algorithme des tableaux ;
 - le mode distribué : utilise les mappings contenus dans M-box et l'algorithme des tableaux distribués.

Reasoner Request Propagator est utilisé pour chercher l'adresse des DRP à exploiter pour la propagation des requêtes.

- **Communication service** : est impliqué dans le cas de la communication entre les paires et a pour tâche principale les dispatching des requêtes aux DRP étrangers et récupère les réponses.

5.2.2 Implémentation de l'approche DDL

Notre outil exploite le raisonneur distribué Drago, il accepte en entrée des ontologies OWL, et représente les ontologies locales du réseau de connaissances multi-niveaux et des fichiers XML contenant les correspondances de type *mappings* et *links*. Les composants principaux de cet outil sont illustrés par la figure 5.12 qui montre l'architecture générale de l'implémentation de l'approche DDL.

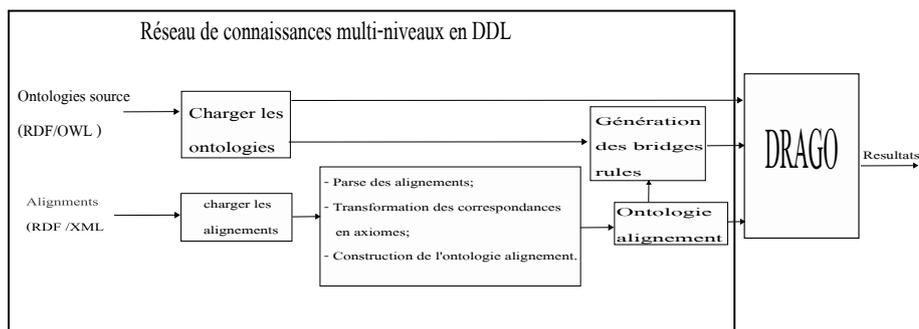


FIGURE 5.12 – Architecture du raisonneur DDLMLNKR

- **Chargeur des ontologies et les alignements**
- **Parseur** : il permet de parser les fichiers XML contenant les alignements, permet de reconnaître les *mappings* qui sont convertis en axiomes et les *links* qui sont convertis en rôles spécifiques et qui constituent l'ontologie alignement générée ;
- **Générateur de l'ontologie alignement** : il s'agit de la construction d'une ontologie en logique de description dont les entités figurent à gauche ou à droite des correspondances, on y intègre aussi les axiomes produits à partir de la transformation des *mappings*, et les rôles produits à partir des *links*.
- **Générateur des Bridges rules** : ce composant permet de générer les bridges rules (des passerelles d'équivalence) entre les entités appartenant aux ontologies locales et les entités correspondantes appartenant aux ontologies alignement et de les stocker dans des fichiers C-OWL ;
- **Module d'exécution du raisonneur distribué DRAGO** : on introduit des URLs de l'ontologies cible (l'ontologie alignement) et des

bridges rules, les ontologies sources sont déterminées par DRAGO, on pourra par la suite déterminer si le réseau d'ontologies est consistant.

CONCLUSION GÉNÉRALE

Résumé de la contribution

Dans cette thèse, nous avons montré la nécessité de mettre en place un formalisme qui permet de combiner des ontologies sémantiquement hétérogènes qui nécessitent des correspondances de type *mappings* pour réduire cette hétérogénéité, et des correspondances de type *links* pour les lier. Une définition contextuelle des *links* par paire d'ontologies génère de nouveau une hétérogénéité sémantique entre ces derniers. Le formalisme proposé permet à la fois de combiner des *mappings* et des *links* et de représenter des relations entre *links* issus de contextes différents. Nous avons formalisé syntaxiquement ce formalisme de manière très générale en introduisant ce que nous appelons les *réseaux de connaissances multi-niveaux*. Ceux-ci dépendent d'un langage d'alignement qui permet de décrire les *mappings* et les *links* entre deux vocabulaires quelconques. Dans la définition récursive des nœuds de connaissances dans un réseau de connaissances multi-niveaux, chaque alignement entre une paire de nœuds dans un certain niveau, génère un autre nœud dans un niveau supérieur. La deuxième partie de ce travail s'est focalisée sur la description de la sémantique de ces réseaux. Nous avons choisi d'exploiter des paradigmes existants disposant des procédures de raisonnement complètes ainsi que des outils opérationnels tels que les logiques de description et les logiques de description distribuées. Les deux manières d'interpréter ont été étudiées pour pouvoir raisonner sur le réseau de connaissances multi-niveaux. La section 2.3.1 montre l'existence de cas réels nécessitant la gestion de l'hétérogénéité sémantique entre les alignements inter-ontologies et justifie ainsi qu'une organisation de la connaissance

en plusieurs niveaux peut être nécessaire. Une implémentation a été réalisée avec les prototypes DLMLNKR (Description Logic Multi-Level Networked Knowledge Reasoner) et DDDMLNKR (DistributeDescription Logic Multi-Level Networked Knowledge Reasoner) afin de valider la faisabilité des approches, permettre de tester la cohérence du réseau de connaissances multi-niveaux et inférer de nouveaux axiomes. Pour les deux prototypes, on exploite des raisonneurs existants tels que Pellet, Hermit et Drago.

Perspective

Grâce à ce travail, plusieurs questions fondamentales peuvent être évoquées et ouvrent de nouvelles perspectives de recherche.

Ainsi que l'introduction de telles structures pose de nouveaux problèmes, pratiques ou théoriques, que nous souhaiterions explorer par la suite et qui sont :

1. On peut s'interroger sur le problème de découverte automatique de correspondances entre alignements : les outils et les techniques utilisés pour la construction d'alignement d'ontologies sont-ils adaptés à tous les niveaux d'un réseau de connaissances ? Peut-on exploiter des alignements à un certain niveau pour la découverte des alignements de niveau supérieur ?
2. nécessité d'une représentation concrète de tels réseaux dans un format et si possible standard ;
3. des outils de gestion ou de visualisation de connaissances doivent être construits pour organiser et observer simplement des réseaux à plusieurs niveaux afin de les maintenir le long de leur cycle de vie. En outre, la construction hiérarchique de réseaux de connaissance exige de réévaluer les méthodologies de modélisation de connaissances en détaillant les étapes à suivre pour leur développement ;
4. Concernant la partie sémantique, nous avons choisi dans ce tra-

vail d'utiliser des paradigmes existants. Il serait intéressant de réfléchir sur une autre façon d'interpréter la sémantique du réseau de connaissances multi-niveaux en définissant une sémantique formelle construite directement sur la structure des réseaux de connaissances multi-niveaux et proposer alors un algorithme de raisonnement correct et complet mieux adapté à cette structure ;

5. Nous nous poserons certainement la question sur la possibilité de transformation du réseau de connaissances multi-niveaux afin d'adopter les autres sémantiques telles que \mathcal{E} -connection, $E - SHIQ$ et IDDL, ceci n'est pas traité dans le cadre de cette thèse mais peut être proposé comme travail à faire en perspective.
6. Enfin, il nous paraît important de réaliser des systèmes capables d'interroger les connaissances de ce type de réseaux et répondre à des requêtes dans un formalisme d'interrogation qui reste à définir.

BIBLIOGRAPHIE

Philippe Adjiman, Philippe Chatalic, François Goasdoué, Marie-Christine Rousset, et Laurent Simon. SomeWhere in the Semantic Web. Dans François Fages et Sylvain Soliman, éditeurs, *Principles and Practice of Semantic Web Reasoning, Third International Workshop, PPSWR 2005, Dagstuhl Castle, Germany, September 11-16, 2005, Proceedings*, volume 3703 de *Lecture Notes in Computer Science*, pages 1–16. Springer-Verlag, 2005. ISBN 3-540-28793-0. (Cité pages 3 et 41.)

Philippe Adjiman, Philippe Chatalic, François Goasdoué, Marie-Christine Rousset, et Laurent Simon. Distributed Reasoning in a Peer-to-Peer Setting : Application to the Semantic Web. *Journal of Artificial Intelligence Research*, 25 :269–314, 2006. (Cité pages 3 et 41.)

Philippe Adjiman, François Goasdoué, et Marie-Christine Rousset. SomeRDFS in the Semantic Web. *Journal on Data Semantics*, 8 :158–181, 2007. (Cité pages 3 et 41.)

Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, et Peter F. Patel-Schneider, éditeurs. *The Description Logic Handbook : Theory, Implementation, and Applications*. Cambridge University Press, 2003. ISBN 0-521-78176-0. (Cité page 17.)

Franz Baader, Carsten Lutz, Holger Sturm, et Frank Wolter. Fusions of Description Logics and Abstract Description Systems. *Journal of Artificial Intelligence Research*, 16 :1–58, 2002. (Cité page 93.)

Franz Baader et Ulrike Sattler. Tableau Algorithms for Description Logics. Dans Roy Dyckhoff, éditeur, *Automated Reasoning with Analytic Tableaux and Related Methods, International Conference, TABLEAUX 2000*,

- St Andrews, Scotland, UK, July 3-7, 2000, Proceedings*, volume 1847 de *Lecture Notes in Computer Science*, pages 1–18. Springer-Verlag, 2000. ISBN 3-540-67697-X. (Cité page 21.)
- Thanh-Le Bach. *Construction d'un web sémantique multi-points de vue*. PhD thesis, Ecole de Mines de Paris à Sophia Antipolis, 2006. (Cité page 25.)
- Jie Bao, jiao Tao, et Deborah L. McGuinness. Context Representation for the Semantic Web . Dans *Proceedings of the web science conference 2010*, 2010. URL <http://www.websci10.org/2010>. (Cité page 24.)
- Jie Bao, George Voutsadakis, Giora Slutzki, et Vasant G. Honavar. Package-Based Description Logics. Dans Heiner Stuckenschmidt, Christine Parent, et Stefano Spaccapietra, éditeurs, *Modular Ontologies : Concepts, Theories and Techniques for Knowledge Modularization*, volume 5445 de *Lecture Notes in Computer Science*, pages 349–371. Springer-Verlag, 2009. ISBN 978-3-642-01906-7. (Cité pages 5, 24, 25, 41 et 48.)
- M. Benerecetti, Paolo Bouquet, et Chiara Ghidini. On the Dimensions of context Dependence : Partiality, Approximation, and Perspective. Dans *proc. of context*, pages 59–72, 2001. (Cité page 23.)
- Camila Bezerra, Fred Freitas, Jérôme Euzenat, et Antoine Zimmermann. ModOnto : A tool for modularizing ontologies. Dans Fred Freitas, Heiner Stuckenschmidt, Helena Sofia Pinto, Andreia Malucelli, et Óscar Corcho, éditeurs, *WONTO 2008, 3rd Workshop on Ontologies and their Applications, Proceedings of the 3rd Workshop on Ontologies and their Applications Salvador, Bahia, Brazil, October 26, 2008*, volume 427 de *CEUR Workshop Proceedings*. Sun SITE Central Europe (CEUR), Octobre 2008. URL http://www.cin.ufpe.br/~wonto2008/wonto2008_arquivos/wonto2008-final.pdf. (Cité page 57.)
- Harold Boley, Said Tabet, et Gerd Wagner. Design Rationale for RuleML : A Markup Language for Semantic Web Rules. Dans *SWWS*, 2001. (Cité page 22.)

- Alex Borgida et Luciano Serafini. Distributed Description Logics : Directed domain correspondences in federated information sources. Dans Robert Meersman et Zahir Tari, éditeurs, *On the Move to Meaningful Internet Systems, 2002 - DOA/CoopIS/ODBASE 2002 Confederated International Conferences DOA, CoopIS and ODBASE 2002 Irvine, California, USA, October 30 - November 1, 2002, Proceedings*, volume 2519 de *Lecture Notes in Computer Science*. Springer-Verlag, 2002. ISBN 3-540-00106-9. URL citeseer.ist.psu.edu/borgida02distributed.html. (Cité page 42.)
- Alex Borgida et Luciano Serafini. Distributed Description Logics : Assimilating Information from Peer Sources. *Journal on Data Semantics*, 1 : 153–184, 2003. (Cité pages 4, 24, 25 et 41.)
- Paolo Bouquet, Fausto Giunchiglia, Frank van Harmelen, Luciano Serafini, et Heiner Stuckenschmidt. C-OWL : Contextualizing Ontologies. Dans Dieter Fensel, Katia P. Sycara, et John Mylopoulos, éditeurs, *The Semantic Web - ISWC 2003, Second International Semantic Web Conference, Sanibel Island, FL, USA, October 20-23, 2003, Proceedings*, volume 2870 de *Lecture Notes in Computer Science*, pages 164–179. Springer-Verlag, Octobre 20 2003. ISBN 3-540-20362-1. URL citeseer.ist.psu.edu/bouquet03cowl.html. (Cité page 29.)
- Bernardo Cuenca-Grau et Oliver Kutz. Modular Ontology Languages Revisited. Dans Vasant G. Honavar, Timothy W. Finin, Doina Caragea, Dunja Mladenic, et York Sure, éditeurs, *SWeCKa 2007 : Proceedings of the IJCAI-2007 Workshop on Semantic Web for Collaborative Knowledge Acquisition, Hyderabad, India, January 7, 2007*, 2007. (Cité page 56.)
- Richard Cyganiak, David Wood, et Markus Lanthaler. RDF 1.1 Concepts and Abstract Syntax, W3C Recommendation 25 February 2014. W3C Recommendation, World Wide Web Consortium (W3C), Février 25 2014. URL <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>. (Cité page 55.)

- Mathieu d'Aquin. Modularizing Ontologies. Dans María del Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta, et Aldo Gangemi, éditeurs, *Ontology Engineering in a Networked World*, pages 213–233. Springer-Verlag, 2012. ISBN 978-3-642-24793-4. (Cité page 56.)
- Mike Dean et Guus Schreiber. OWL Web Ontology Language Reference, W3C Recommendation 10 February 2004. W3C Recommendation, World Wide Web Consortium (W3C), Février 10 2004. URL <http://www.w3.org/TR/2004/REC-owl-ref-20040210/>. (Cité page 21.)
- Lynda Djakhdjakha, Mounir Hemam, et Zizette Boufaïda. Semantics of multi-view points ontology alignments. Dans *Computer System and Applications, 11th International Conference on IEEE*, 2014. (Cité page 25.)
- Warith Djeddi et Mohamed Tarek Khadir. Ontology alignment using artificial neural network for large-scale ontologies. *International Journal of Metadata, Semantics and ontologies*, 2013. (Cité page 105.)
- Chan Le Duc, Myriam Lamolle, Antoine Zimmermann, et Olivier Curé. DRAOn : A Distributed Reasoner for Aligned Ontologies. Dans Samantha Bail, Birte Glimm, Rafael Gonçalves, Ernesto Jiménez-Ruiz, Yevgeny Kazakov, Nicolas Matentzoglou, et Bijan Parsia, éditeurs, *Informal Proceedings of the 2nd International Workshop on OWL Reasoner Evaluation (ORE-2013)*, volume 1015 de *CEUR Workshop Proceedings*, pages 81–86. Sun SITE Central Europe (CEUR), Juillet 2013. (Cité page 55.)
- Faezeh Ensan et Weichang Du. A knowledge encapsulation approach to ontology modularization. *Knowledge and Information Systems*, 26(2) :249–283, 2011. (Cité page 57.)
- Jérôme Euzenat. An API for Ontology Alignment. Dans Frank van Harmelen, Sheila McIlraith, et Dimitri Plexousakis, éditeurs, *The Semantic Web - ISWC 2004 : Third International Semantic Web Conference, Hiroshima, Japan, November 7-11, 2004. Proceedings*, volume 3298 de

- Lecture Notes in Computer Science*, pages 698–712. Springer-Verlag, 2004. ISBN 3-540-23798-4. URL <http://www.springerlink.com/index/DY8Y9F31A9GT9762>. (Cité page 31.)
- Jérôme Euzenat et Pavel Shvaiko. *Ontology Matching*. Springer-Verlag, Heidelberg (DE), 2007. ISBN 3-540-49611-4. (Cité pages 25 et 26.)
- Jérôme Euzenat, Antoine Zimmermann, et Fred Freitas. Alignment-based modules for encapsulating ontologies. Dans Bernardo Cuenca-Grau, Vasant G. Honavar, Anne Schlicht, et Frank Wolter, éditeurs, *Proceedings of the 2nd International Workshop on Modular Ontologies (WoMO-2007), Whistler, Canada, October 28, 2007*, volume 315 de *CEUR Workshop Proceedings*, pages 32–45. Sun SITE Central Europe (CEUR), Octobre 2007. (Cité page 57.)
- Chiara Ghidini et Fausto Giunchiglia. Local Models Semantics, or contextual reasoning=Locality+Compatibility. *Artificial Intelligence*, 127(2) : 221–259, 2001. (Cité page 24.)
- Chiara Ghidini et Luciano Serafini. Distributed First Order Logics. Dans Dov M. Gabbay et Maarten de Rijke, éditeurs, *Frontiers of Combining Systems 2*, volume 7 de *Studies in Logic and Computation*, pages 121–139. Research Studies Press, 2000. ISBN 0863802524. URL citeseer.ist.psu.edu/ghidini98distributed.html. (Cité pages 24 et 55.)
- Chiara Ghidini, Luciano Serafini, et Sergio Tessaris. On Relating Heterogeneous Elements from Different Ontologies. Dans *Modeling and Using Context, 6th International and Interdisciplinary Conference, CONTEXT 2007, Roskilde, Denmark, August 20-24, 2007, Proceedings*, volume 4635 de *Lecture Notes in Computer Science*, pages 234–247. Springer-Verlag, 2007. ISBN 978-3-540-74254-8. (Cité page 42.)
- Chiara Ghidini, Luciano Serafini, et Sergio Tessaris. Complexity of reasoning with expressive ontology mappings. Dans *FOIS*, pages 151–163, 2008. (Cité pages 42 et 94.)

- Fausto Giunchiglia. Contextual Reasoning. *Epistemologica*, 16 :345–364, 1993. (Cité pages 24 et 36.)
- Fausto Giunchiglia et Luciano Serafini. Multilanguage hierarchical logics, or : How we can do without modal logics. *Artificial Intelligence*, 65(1) : 29–70, 1994. (Cité page 24.)
- Ramanathan V. Guha. *a Formalization and some Applications*. PhD thesis, Stanford University, 1992. (Cité page 24.)
- Martin Homola. Semantic Investigations in Distributed Ontologies. Dans *PHD dissertation*, 2010. (Cité page 42.)
- Martin Homola et Luciano Serafini. Contextualized Knowledge Repositories for the Semantic Web. *Journal of Web Semantics*, 2012. (Cité pages 25 et 58.)
- Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosz, et Mike Dean. SWRL : A Semantic Web Rule Language - Combining OWL and RuleML. W3C member submission, World Wide Web Consortium (W3C), Mai 21 2004. URL <http://www.w3.org/TR/2004/REC-owl-ref-20040210/>. (Cité pages 22 et 29.)
- Mathew Joseph. *Query Answering over Contextualized RDF/OWL rules : Decidable classes*. PhD thesis, University of Trento, 2015. (Cité pages 25 et 58.)
- Mathew Joseph, Gabriel M. Kuper, et Luciano Serafini. Query Answering over Contextualized RDF Knowledge with Forall-Existential Bridge Rules : Attaining Decidability Using Acyclicity. Dans Laura Giordano, Valentina Gliozzi, et Gian Luca Pozzato, éditeurs, *Proceedings of the 29th Italian Conference on Computational Logic, Torino, Italy, June 16-18, 2014.*, volume 1195 de *CEUR Workshop Proceedings*, pages 210–224. Sun SITE Central Europe (CEUR), Juin 2014. URL <http://ceur-ws.org/Vol-1195/long14.pdf>. (Cité page 55.)

- Mathew Joseph et Luciano Serafini. Simple Reasoning for Contextualized RDF Knowledge. Dans Oliver Kutz et Thomas Schneider, éditeurs, *Modular Ontologies - Proceedings of the Fifth International Workshop, WoMO 2011, Ljubljana, Slovenia, August 2011*, volume 230 de *Frontiers in Artificial Intelligence and Applications*, pages 79–93. IOS Press, Août 2011. ISBN 978-1-60750-798-7. URL <http://dx.doi.org/10.3233/978-1-60750-799-4-79>. (Cité pages 25 et 58.)
- Sihem Klai et Mohamed Tarek Khadir. Approach for a rule based ontologies integration . Dans *5th International Conference Computer Science and Information Technology (CSIT) IEEE. Amman Jordan March 27-28 , 2013*. ISBN 978-9957-8592-1-3. (Cité page 29.)
- Sihem Klai, Antoine Zimmermann, et Mohamed Tarek Khadir. *Multi-level Networked Knowledge Base : DDL reasoning*, volume 9893 de *Lecture Notes in Computer Science*. Springer-Verlag, 2016a. (Cité pages 65 et 81.)
- Sihem Klai, Antoine Zimmermann, et Mohamed Tarek Khadir. Multi-level Networked Knowledge : Rerepresentation and DL reasoning. *International Journal of Metadata, Semantics and ontologies*, 11(01) :1–15, 2016b. (Cité pages 65, 72 et 97.)
- Szymon Klarman. *Reasoning with Contexts in Description Logics*. PhD thesis, Vrije Universiteit, Amsterdam (Netherland), Janvier 2013. (Cité page 58.)
- Boris Konev, Carsten Lutz, Dirk Walther, et Frank Wolter. Formal Properties of Modularisation. Dans Heiner Stuckenschmidt, Christine Parent, et Stefano Spaccapietra, éditeurs, *Modular Ontologies : Concepts, Theories and Techniques for Knowledge Modularization*, volume 5445 de *Lecture Notes in Computer Science*, pages 25–66. Springer-Verlag, 2009. ISBN 978-3-642-01906-7. (Cité page 56.)
- Oliver Kutz, Carsten Lutz, Frank Wolter, et Michael Zakharyashev. \mathcal{E} -connections of abstract description systems. *Artificial Intelligence*, 156(1) :1–73, 2004. (Cité pages 4, 24, 25, 41 et 45.)

Alexander Maedche, Boris Motik, Nuno Silva, et Raphael Voltz. MAFRA - A MAPPING FRAMework for Distributed Ontologies. Dans *Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web, 13th International Conference, EKAW 2002, Sigüenza, Spain, October 1-4, 2002, Proceedings*, pages 235–250. Springer-Verlag, 2002. ISBN 3-540-44268-5. (Cité page 31.)

John L. McCarthy. Generality in Artificial Intelligence. *Communications of the ACM*, 30(12) :1029–1035, 1987. (Cité pages 23 et 24.)

John L. McCarthy. Notes on formalizing context. Dans *Proceeding of the 13th international joint conference on Artificial Intelligence, IJAI'93, 1993*. (Cité pages 23 et 24.)

Till Mossakowski, Oliver Kutz, Mihai Codescu, et Christoph Lange. The Distributed Ontology, Modeling and Specification Language. Dans Chiara Del Vescovo, Torsten Hahmann, David Pearce, et Dirk Walther, éditeurs, *Proceedings of the 7th International Workshop on Modular Ontologies, co-located with the 12th International Conference on Logic Programming and Non-monotonic Reasoning (LPNMR 2013), Corunna, Spain, September 15, 2013*, volume 1081 de *CEUR Workshop Proceedings*. Sun SITE Central Europe (CEUR), Septembre 2013a. URL http://ceur-ws.org/Vol-1081/womo2013_invited_paper_1.pdf. (Cité page 56.)

Till Mossakowski, Christoph Lange, et Oliver Kutz. Three Semantics for the Core of the Distributed Ontology Language (Extended Abstract). Dans Francesca Rossi, éditeur, *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*. IJCAI/AAAI, Août 2013b. ISBN 978-1-57735-633-2. URL <http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6611>. (Cité page 56.)

Bijan Parsia et Bernardo Cuenca-Grau. Generalized link properties for expressive \mathcal{E} -connection of description logics, 2005. (Cité page 47.)

Luciano Serafini et Andrei Tamin. Local Tableaux for Reasoning in Distributed Description Logics. Dans Volker Haarslev et Ralf Möller, éditeurs, *Proceedings of the 2004 International Workshop on Description Logics (DL2004)*, Whistler, British Columbia, Canada, June 6-8, 2004, volume 104 de *CEUR Workshop Proceedings*. Sun SITE Central Europe (CEUR), 2004. (Cité page 42.)

Luciano Serafini et Andrei Tamin. DRAGO : Distributed Reasoning Architecture for the Semantic Web. Dans Asunción Gómez-Pérez et Jérôme Euzenat, éditeurs, *The Semantic Web : Research and Applications, Second European Semantic Web Conference, ESWC 2005, Heraklion, Crete, Greece, May 29 - June 1, 2005, Proceedings*, volume 3532 de *Lecture Notes in Computer Science*, pages 361–376. Springer-Verlag, Mai 2005. ISBN 3-540-26124-9. (Cité pages 45 et 88.)

Luciano Serafini et Andrei Tamin. Instance Migration in Heterogeneous Ontology Environments. Dans Karl Aberer, Key-Sun Choi, Natasha Fridman Noy, Dean Allemang, Kyung-Il Lee, Lyndon J. B. Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, Guus Schreiber, et Philippe Cudré-Mauroux, éditeurs, *The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 11-15, 2007*, volume 4825 de *Lecture Notes in Computer Science*, pages 452–465. Springer-Verlag, 2007. ISBN 978-3-540-76297-3. (Cité page 42.)

John F. Sowa. *Knowledge Representation : Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co., Pacific Grove, CA, 2000. ISBN 0-534-94965-7. (Cité page 15.)

Andrei Tamin. *Distributed Ontological Reasoning Theory, Algorithms And Applications*. PhD thesis, University of Trento, Février 2007. (Cité page 45.)

George Vouros et Georgios M. Santipantakis. Distributed Reasoning with

$E^{DDL}_{HQ^+}$ SHIQ. Dans Jodi Schneider et Dirk Walther, éditeurs, *Proceedings of the 6th International Workshop on Modular Ontologies, Graz, Austria, July 24, 2012*, volume 875 de *CEUR Workshop Proceedings*. Sun SITE Central Europe (CEUR), Juillet 2012. URL http://ceur-ws.org/Vol-875/short_paper_2.pdf. (Cité pages 24, 25, 30, 41 et 50.)

Antoine Zimmermann. *Integrated Distributed Description Logics*. Dans Diego Calvanese, Enrico Franconi, Volker Haarslev, Domenico Lembo, Boris Motik, Sergio Tessaris, et Anni-Yasmin Tuhran, éditeurs, *Proceedings of the 20th International Workshop on Description Logics DL'07, June 8 - 10, 2007, Brixen/Bressanone, Italy*, pages 507–514. Bolzano University Press, Juin 2007. ISBN 978-88-6046-008-5. URL http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-250/paper_37.pdf. (Cité pages 4, 5, 24, 25, 41 et 52.)

Antoine Zimmermann. *sémantique des réseaux des connaissances : gestion de l'hétérogénéité fondée sur le principe de médiation*. PhD thesis, Université Joseph-Fourier-Grenoble, Décembre 2008. (Cité page 95.)

Antoine Zimmermann et Jérôme Euzenat. *Three Semantics for Distributed Systems and their Relations with Alignment Composition*. Dans Isabel F. Cruz, Stephan Decker, Dean Allemang, Christ Preist, Daniel Schwabe, Peter Mika, Michael Uschold, et Lora Aroyo, éditeurs, *The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006, Proceedings*, volume 4273 de *Lecture Notes in Computer Science*, pages 16–29. Springer-Verlag, Novembre 2006. ISBN 3-540-49029-9. URL <http://iswc2006.semanticweb.org/items/Zimmermann2006jw.pdf>. (Cité page 56.)