

Travaux d'études et de recherches

Argumentation, contexte et alignement d'ontologies

Robin Lamarche-Perrin
(Robin.Lamarche-Perrin@ensimag.imag.fr)

Encadrant
Jérôme Euzenat
(Jerome.Euzenat@imag.fr)

INRIA Grenoble Rhône-Alpes
Laboratoire d'Informatique de Grenoble
Équipe EXMO

Table des matières

Table des matières.....	2
1. Introduction.....	3
1.1. Cadre de recherche.....	3
1.2. Sujet de recherche.....	3
2. Théories de l'argumentation.....	5
2.1. Exemple préliminaire.....	5
2.2. Théorie classique.....	6
2.3. Argumentation basée sur des valeurs.....	7
3. Pratiques de l'argumentation.....	10
3.1. Argumentation au service de l'alignement d'ontologies.....	10
3.2. Limites du modèle.....	12
3.3. Argumentation selon des degrés de confiance.....	13
4. Argumentation et contexte.....	15
4.1. Problématique.....	15
4.2. Arguments contextuels.....	16
4.3. Argumentation cohérente.....	20
4.4. Historique d'argumentation.....	21
5. Conclusion.....	23
Bibliographie.....	24
Remerciements.....	24

1. Introduction

1.1. Cadre de recherche

Les travaux qui suivent ont été menés au sein de l'équipe de recherche EXMO¹ de l'INRIA Grenoble Rhône-Alpes. Ils ont été encadrés par Jérôme Euzenat, directeur de recherche à l'INRIA, qui mène l'équipe en collaboration avec le LIG (Laboratoire d'Informatique de Grenoble).

L'équipe EXMO étudie les « échanges de connaissance structurée médiatisés par ordinateur ». Elle propose des outils théoriques et logiciels pour aider à la formalisation de la connaissance et à son échange via les systèmes d'informations. Elle participe entre autre au développement du Web Sémantique et de la communication de ses ressources entre les hommes. Dans cet objectif, l'équipe EXMO s'intéresse au traitement et à la mise en correspondance des différents modèles de représentations présents sur le Web, afin de les rendre interopérables.

La première phase du travail a consisté principalement à lire les documents publiés à propos du sujet de recherche et à se familiariser avec les notions utilisées dans le domaine. La rédaction de ce présent dossier a ensuite été réalisée en s'appuyant sur ces lectures et sur les entretiens bimensuels avec Jérôme Euzenat.

1.2. Sujet de recherche

Nous souhaitons favoriser l'échange massif des ressources du Web entre les hommes. Certaines technologies proposées par le Web Sémantique ont pour vocation de rendre ces ressources accessibles aux agents logiciels, afin de développer des algorithmes de classification, de formatage et de traitement des données. Dans ce but, le Web Sémantique offre différentes méthodes de représentation formelle de l'information. Certains langages de formalisation sont assez répandus, comme XML (*Extensible Markup Language*) qui permet, via des balises de marquage, la structuration des données en vue de leur stockage ou de leur échange. Le W3C (*World Wide Web Consortium*), fondé en 1994, est chargé de favoriser la standardisation de ces technologies, afin de les rendre interopérables.

Cependant, les modèles de représentation sont souvent hétérogènes, soit parce qu'ils utilisent des langages de formalisation différents, soit parce qu'ils sont référencés par des vocabulaires variés au sein d'un même langage. Nous cherchons donc des méthodes pour les mettre en correspondance, nous cherchons des moyens de traduction.

Les ontologies sont des ensembles structurés de termes et de concepts décrivant un champ de la connaissance. Elles sont ainsi des modèles de représentation qui permettent le raisonnement sur les objets du domaine ou l'échange d'informations. Le W3C propose un langage de description pour construire de telles ontologies, OWL (*Web Ontology Language*) qui est un dialecte du langage XML. Il existe sur le Web de nombreux cas d'utilisation de cette technologie, mais les représentations développées par ces ontologies ne sont pas toujours homogènes. Nous donnons un exemple d'ontologies hétérogènes qui représentent chacune un même objet de connaissance : un film (cf. **figure 1**).

Nous cherchons à mettre en correspondance les termes des différentes ontologies et à exprimer les liens qu'elles entretiennent entre elles. Par exemple, nous aimerions lier le terme *titre_du_film* de l'ontologie 1 et le terme *nom_film* de l'ontologie 2 puisqu'ils semblent désigner la

¹ Pour en savoir plus :

<http://www.inria.fr/recherche/equipes/exmo.fr.html>

<http://www.inrialpes.fr/exmo/>

même chose. Cette activité s'appelle l'alignement d'ontologies. Elle constitue une part importante du travail de l'équipe EXMO.

Plusieurs algorithmes, que nous nommons *matchers*, permettent de produire de tels alignements entre ontologies. Ils utilisent différents critères pour proposer des correspondances entre les termes de deux ontologies hétérogènes (comparaison lexicale du nom des termes, comparaison de la structure des deux ontologies, comparaison sémantique, etc.). Et parmi les critères utilisés, il faut choisir lesquels sont les plus pertinents.

Figure 1
Deux ontologies hétérogènes
représentant un même champ de la connaissance

Ontologie 1	Ontologie 2
<ul style="list-style-type: none"> ▶ <i>titre_du_film</i> ▶ <i>genre_du_film</i> ▶ <i>annee_de_sortie</i> ▶ <i>equipe</i> <ul style="list-style-type: none"> ▶ <i>nom_du_realisateur</i> ▶ <i>nom_du_producteur</i> ▶ <i>equipe_realisation</i> <ul style="list-style-type: none"> ▶ <i>scenaristes</i> ▶ <i>monteurs</i> ▶ <i>equipe_tournage</i> <ul style="list-style-type: none"> ▶ <i>acteurs</i> ▶ <i>cadreurs</i> 	<ul style="list-style-type: none"> ▶ <i>nom_film</i> ▶ <i>genre</i> ▶ <i>annee</i> ▶ <i>budget</i> ▶ <i>equipe</i> <ul style="list-style-type: none"> ▶ <i>realisateurs</i> ▶ <i>producteurs</i> ▶ <i>scenaristes</i> ▶ <i>acteurs_principaux</i> ▶ <i>acteurs_secondaires</i>

Nous nous plaçons dans le cadre d'un système multi-agents. Prenons l'exemple d'un agent 1 qui veut acheter une place de cinéma à un agent 2 qui propose ce service. Supposons que l'agent 1 utilise l'ontologie 1 pour organiser ses données concernant le film à voir alors que l'agent 2 utilise l'ontologie 2. Afin d'échanger des informations, les deux agents doivent convenir d'un alignement entre les ontologies qu'ils utilisent. Ils veulent en somme obtenir un moyen de traduction entre leurs deux vocabulaires.

Les choix des critères de mise en correspondance peuvent différer entre les deux agents. Par exemple, le premier préfère une simple corrélation lexicographique des termes, tandis que le second veut utiliser une banque de correspondances sémantiques du Web (de telles banques donnent les liens sémantiques qui peuvent exister entre différents termes du langage naturel). Comment résoudre un tel conflit ?

La théorie de l'argumentation de Phan Minh Dung [1] donne une solution à ce problème. Elle formalise ce que nous appelons un *système d'argumentation*. Il s'agit d'une structure logique au sein de laquelle des assertions (les arguments) sont exprimées et peuvent s'attaquer ou se défendre mutuellement.

Nous rappelons dans un premier temps les différentes théories qui ont été développées à propos de ces *systèmes d'argumentation* (cf. **partie 2**). Puis nous montrons comment elles sont utilisées pour servir l'alignement d'ontologies dans le cadre des systèmes multi-agents (cf. **partie 3**).

Cependant, nous pensons que ces travaux offrent une vue trop figée de l'argumentation. La négociation n'est alors pas sensible au contexte même de la communication. Nous expliquons cette limite et proposons des méthodes de résolution (cf. **partie 4**).

2. Théories de l'argumentation

2.1 Exemple préliminaire

Nous illustrons le principe de la théorie de l'argumentation par l'exemple suivant. Antoine et Benoît veulent aller au cinéma. Ils discutent à propos du lieu de la séance Il est 20h et ils sont sur le campus.

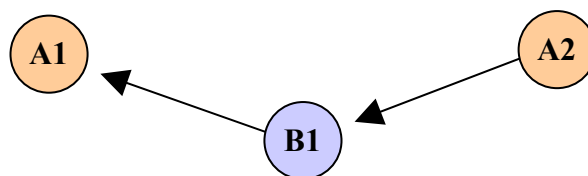
A1 - Antoine propose d'aller au Club en centre-ville où les places sont moins chères.

B1 - Benoît rétorque qu'ils n'auront pas le temps d'y arriver en tramway pour la séance de 20h15.

A2 - Antoine répond qu'il a deux vélos dans son garage.

L'argument **B1** de Benoît attaque la proposition initiale **A1** d'Antoine. Mais en proposant d'aller au Club en vélo, le contre-argument **A2** d'Antoine attaque **B1**. Nous pouvons schématiser le jeu d'attaque de ces trois arguments par un graphe (cf. **figure 2**).

Figure 2
Graphe d'attaque entre arguments



Il est évident que si nous nous arrêtons ici, la position d'Antoine est acceptable. En effet, puisqu'il a le dernier mot, les arguments qu'il a formulés forment un ensemble cohérent que Benoît ne peut pas remettre en question avec son argument **B1**. L'ensemble $\{A1, A2\}$ est donc accepté (tandis que l'argument **B1** est refusé) et ce pour deux raisons :

i) **Aucun argument de l'ensemble A n'attaque un autre argument de cet ensemble.** Il y a une cohérence interne.

ii) **Tous les arguments qui attaquent un argument de l'ensemble A sont à leur tour attaqués par un argument de cet ensemble.** Il y a une stabilité vis-à-vis du système pris dans sa totalité.

Bien entendu, si Benoît formule un argument qui contredit **A2** (**B2** - Benoît rappelle qu'il ne sait pas faire de vélo.), alors l'ensemble accepté est $\{B1, B2\}$. Et les relations d'attaque impliquent le rejet des arguments **A1** et **A2**.

Le calcul, sur des graphes plus complexes, de ces ensembles stables et cohérents est le principe de la théorie de l'argumentation. Ce sont les *extensions préférées* du système, elles définissent des positions stables vis-à-vis de l'argumentation. Nous allons voir maintenant comment formaliser le graphe et en définir ces *extensions préférées*.

2.2. Théorie classique

Nous rappelons ici les bases posées par la théorie de l'argumentation de Phan Minh Dung [1].

Les arguments ne sont pas ici définis. En effet, la théorie formalise le cadre général de l'argumentation. Elle ne précise pas ce sur quoi porte l'argument, mais tire son rôle de la relation qu'il entretient avec les autres dans le graphe constitué par le système.

Définition 2.2.1

Un *système d'argumentation* est un doublet $SA = \langle EA, att \rangle$ où :

- EA est un ensemble d'arguments
- att est une relation binaire sur EA que nous nommons relation d'attaque

Nous avons donc $att \subseteq (EA \times EA)$. Cette relation définit les attaques entre arguments et formalise ainsi la notion de *contre-argument*.

Définition 2.2.2

Un argument A *attaque* un autre argument B si et seulement si on a la relation $att(A, B)$. On dit alors que A est un *contre-argument* de B .

Définition 2.2.3

Un ensemble d'arguments S *attaque* un argument B si et seulement si $\exists X \in S, att(X, B)$.

On dit qu'un ensemble d'arguments soutient un argument A (il est alors *acceptable* selon cet ensemble) si il est capable de proposer un contre-argument à n'importe quel argument qui attaque A .

Définition 2.2.4

L'argument A est *acceptable* selon un ensemble d'arguments S si et seulement si $\forall X \in EA, att(X, A) \Rightarrow (\exists Y \in S, att(Y, X))$.

Il s'agit maintenant de définir des ensembles d'arguments cohérents, i.e. qui ne contiennent aucun argument attaquant un autre argument de cet ensemble. Ces ensembles représentent ainsi des positions sans conflit logique interne.

Définition 2.2.5

Un ensemble d'arguments S est *sans-conflit* si et seulement si $\neg (\exists (X, Y) \in S^2, att(X, Y))$.

De plus, de tels ensembles peuvent être stable vis-à-vis de la totalité des arguments du système. Il s'agit, en plus d'une position *sans-conflit*, d'une position défendable vis-à-vis de l'ensemble des arguments.

Définition 2.2.6

Un ensemble d'arguments *sans-conflit* S est *admissible* si et seulement si $\forall X \in S, X$ est *acceptable* selon S .

Un ensemble d'arguments est donc *admissible* si chacun de ses arguments est soutenu par le reste de l'ensemble. Nous avons ici une stabilité vis-à-vis du *système d'argumentation*. Enfin, nous voulons inclure le maximum d'arguments dans ce type d'ensemble pour représenter la position cohérente et défendable qui prend en compte le plus d'arguments possibles.

Définition 2.2.7

Un ensemble d'arguments S est une *extension préférée* si et seulement si il est un ensemble *admissible* de taille maximale (au sens de l'inclusion).

Cette notion d'*extension préférée* est la notion clé de la théorie de Phan Minh Dung. Elle permet de construire les positions d'argumentation *admissibles* par le système qui regroupe le plus grand nombre d'arguments. Il faut noter que pour certains systèmes plusieurs *extensions préférées* sont possibles, i.e. plusieurs positions différentes sont *admissibles*. Dans le cas d'un système multi-agents, nous pouvons nous retrouver avec des conclusions différentes après la négociation. C'est qu'alors le débat n'est pas allé assez loin, il faut formuler de nouveaux arguments pour unifier les extensions. Sinon, le débat se solde par un échec de la négociation. Chaque agent peut de bon droit rester campé sur sa position et aucune entente n'est alors possible.

Nous ne développerons pas ici comment les *extensions préférées* sont ensuite prises en compte dans la suite de la théorie classique. Nous préférons étendre ce système à l'aide de préférences sur les arguments par ce qui suit.

2.3. Argumentation basée sur des valeurs

Supposons que dans l'exemple énoncé plus haut, Benoît préfère les arguments de confort (ne pas trop se déplacer, prendre le tramway, etc.) aux arguments liés à l'économie d'argent (aller là où c'est moins cher, prendre les vélos, etc.). Dans cette optique, l'argument **A2** ne suffit pas à attaquer **B1**, il est négligé par Benoît, et donc il ne défend pas avec succès la proposition **A1**.

Cette notion de préférence sur les arguments est introduite par Leila Amgoud [2] à l'intérieur de la théorie classique de l'argumentation. Il s'agit d'ordonner les arguments et de leur accorder plus ou moins de crédits, modifiant ainsi la géométrie du graphe en fonction de ces préférences.

Trevor Bench-Capon [3] reprend ce travail et donne aux arguments des valeurs. Ces valeurs permettent de classer les arguments dans des catégories. Ainsi, nous exprimons des préférences sur ces valeurs et non plus sur les arguments eux-mêmes. Les différents classements possibles sont appelés *audiences*. Elles définissent des préférences d'argumentation qui permettent de juger de la validité des arguments les uns par rapport aux autres.

Nous allons voir comment sont redéfinies les *extensions préférées* dans cette théorie. Elles sont maintenant fonctions des audiences. Dans notre exemple, l'audience préférant les arguments de confort aux arguments économiques aura pour extension l'ensemble $\{\mathbf{B1}\}$. Mais l'audience inversant cette préférence aura toujours $\{\mathbf{A1}, \mathbf{A2}\}$ pour extension.

Les valeurs attribuées aux arguments (leurs catégories) ne sont pas précisées ici. Nous sommes toujours dans une description abstraite du *système d'argumentation*. Il faudra ensuite préciser ces valeurs en fonction de l'objet de l'argumentation.

Définition 2.3.1

Un *système d'argumentation basé sur des valeurs* est un 5-uplet $SAV = \langle EA, att, V, val, P \rangle$ où :

- $\langle EA, att \rangle$ est un *système d'argumentation*
- V est un ensemble non-vide de valeurs
- val est une fonction des éléments de EA vers les éléments de V
- P est un ensemble d'audiences possibles

Pour ne pas contraindre la définition des arguments, nous utilisons une fonction pour leur associer une valeur. C'est le rôle de la fonction *val*.

Une audience $a \in P$ représente un classement des valeurs de l'ensemble V . Elle définit donc une relation binaire de préférence (transitive, irréflexive et asymétrique) sur cet ensemble que nous notons $valpref_a$.

Définition 2.3.2

Une valeur v_1 est *préférée* à une valeur v_2 selon une audience a si et seulement si on a la relation $valpref_a(v_1, v_2)$.

Nous redéfinissons les différentes notions de la théorie classique pour prendre en compte les préférences formalisées par les audiences.

Définition 2.3.3

Un argument A *attaque avec succès* selon l'audience a un autre argument B si et seulement si $att(A, B) \wedge \neg valpref_a(val(B), val(A))$. On note $att_a(A, B)$.

Définition 2.3.4

Un argument A est *acceptable* selon une audience a pour un ensemble d'arguments S si et seulement si $\forall X \in EA, att_a(X, A) \Rightarrow (\exists Y \in S, att_a(Y, X))$.

Un ensemble est maintenant *sans-conflit* si aucun argument n'attaque un autre argument de cet ensemble, ou alors il faut que l'argument attaqué ait une valeur *préférée* à celle de l'argument qui l'attaque.

Définition 2.3.5

Un ensemble d'arguments S est *sans-conflit* selon une audience a si et seulement si $\neg (\exists (X, Y) \in S^2, att(X, Y) \wedge \neg valpref_a(val(Y), val(X)))$, i.e. si $\neg (\exists (X, Y) \in S^2, att_a(Y, X))$.

Définition 2.3.6

Un ensemble d'arguments S *sans-conflit* selon une audience a est *admissible* selon cette même audience a si et seulement si $\forall X \in S, X$ est *acceptable* selon l'audience a pour l'ensemble S .

Définition 2.3.7

Un ensemble d'argument S est une *extension préférée* selon l'audience a si et seulement si il est un ensemble *admissible* selon l'audience a de taille maximale (au sens de l'inclusion).

Après avoir redéfini la notion d'*extension préférée* (qui est maintenant fonction d'une audience donnée), nous précisons plusieurs catégories d'arguments. Ils n'ont pas la même place au sein de l'argumentation suivant qu'ils apparaissent dans certaines *extensions préférées*, dans toute, ou encore dans aucune d'entre elles.

Définition 2.3.8

Un argument X est *objectivement acceptable* si et seulement si il appartient aux *extensions préférées* selon toutes les audiences de P .

Définition 2.3.9

Un argument X est *subjectivement acceptable* si et seulement si il appartient aux *extensions préférées* selon certaines audiences de P , mais pas toutes.

Définition 2.3.10

Un argument X est *non-défendable* si et seulement si il n'appartient à aucune *extension préférée*.

Le graphe des arguments n'est plus figé. En fonction de l'audience avec laquelle nous l'observons, le résultat de l'argumentation peut différer. Les arguments ne sont plus simplement acceptés ou refusés, ils acquièrent du crédit par rapport au nombre des *extensions préférées* auxquelles ils appartiennent. C'est de la confrontation de ces diverses audiences que naît le résultat de l'argumentation.

Nous nous intéressons maintenant à l'utilisation de cette théorie au service de l'alignement d'ontologies dans les systèmes multi-agents.

3. Pratiques de l'argumentation

Problématique : Deux agents dotés d'ontologies hétérogènes veulent communiquer. Ils disposent de plusieurs correspondances entre les termes de ces ontologies et doivent choisir lesquelles seront utilisées lors de l'échange d'informations. Leurs choix peuvent différer, ils ont alors besoin de négocier à propos de ces correspondances.

Comment donner un cadre formel à cette négociation ?

3.1. Argumentation au service de l'alignement d'ontologies

Loredana Laera [4] propose une utilisation de la théorie de l'argumentation pour résoudre le problème de l'alignement d'ontologies dans un système multi-agents.

Il s'agit dans un premier temps de donner une définition des arguments, de leurs valeurs et de formaliser un processus de négociation. Nous nous intéressons tout d'abord aux objets sur lesquels portera l'argumentation : les correspondances entre ontologies.

Définition 3.1.1

Une *ontologie* est un 5-uplet $\langle C, \text{ordre}C, R, \text{ordre}R, I \rangle$ où :

- C est un ensemble de concepts
- R est un ensemble de relations
- I est un ensemble d'instances des concepts de C qui peuvent être mis en relation avec R
- $\text{ordre}C$ est une relation d'ordre partiel sur C
- $\text{ordre}R$ est une relation d'ordre partiel sur R

Il faut noter que ce n'est qu'une définition possible parmi d'autres. Elle sert de base pour le processus décrit par Loredana Laera. Les relations $\text{ordre}C$ et $\text{ordre}R$ permettent d'exprimer la structure de l'ontologie (quels concepts ou relations sont plus généraux que les autres).

Définition 3.1.2

Une *correspondance* est un triplet $\langle e, e', r \rangle$ où :

- e et e' sont des entités (concept, relation ou instance) de deux ontologies
- r est la relation liant e à e' (entités équivalentes, e plus générale que e' , etc.)

Définition 3.1.3

Un *argument* est un triplet $\langle v, c, \sigma \rangle$ où :

- $v \in V$ est la valeur de l'argument
- c est la correspondance sur laquelle on argumente
- $\sigma \in \{-, +\}$ suivant si l'argument soutient ou attaque la correspondance

Définition 3.1.4

L'ensemble V des valeurs que peuvent prendre les arguments est $\{\text{semantic}, \text{internal structural}, \text{external structural}, \text{terminological}, \text{extensional}\}$.

Ces différentes valeurs font référence aux différents types de *matchers* qui peuvent exister.

- *semantic matching* – les deux entités partagent plus ou moins de sens (utilise une banque de correspondances sémantiques librement disponible sur le Web)
- *internal structural matching* – la structure interne des deux entités est plus ou moins similaire

(cardinalité de leurs attributs, rang dans l'ontologie, etc.)

- *external structural matching* – les ensembles de relations des deux entités sont comparables
- *terminological matching* – les noms des entités sont plus ou moins semblables sur le plan lexical
- *extensional matching* – les extensions connues des deux entités sont plus ou moins comparables

Définition 3.1.5

Un *agent* est un quadruplet $\langle o, SAV, a, \epsilon \rangle$ où :

- O est une ontologie utilisée par l'agent
- $SAV = \langle EA, att, V, val, P \rangle$ est le *système d'argumentation basé sur des valeurs* de l'agent
- a est l'audience de P propre à l'agent
- ϵ est le seuil de confiance de l'agent

Soit un système multi-agents de taille n , $Ag = \{Ag_1, Ag_2, \dots, Ag_n\}$. Chacun des agents utilise une ontologie qui lui est spécifique, $O = \{o_1, o_2, \dots, o_n\}$. Nous disposons d'un ensemble de correspondances entre les entités de ces ontologies $C = \{c_1, c_2, \dots, c_m\}$ (cet ensemble est généralement obtenu via une banque de correspondances du Web).

Il nous faut définir un processus de négociation afin de convenir de l'ensemble des correspondances qui formeront les alignements entre ontologies. Le processus décrit par Loredana Laera a lieu en trois étapes. Nous les rappelons ici :

- 1) Pour chaque correspondance de C , chaque agent de Ag produit un *système d'argumentation* à partir des arguments qu'il possède dans son propre système. Ces arguments attaquent ou défendent la correspondance en question. Puis, il calcule les relations d'attaque entre ces arguments.
- 2) Les différents *systèmes d'argumentation* sont ensuite échangés entre les agents et chacun étend les relations d'attaque entre les arguments de ces systèmes.
- 3) Enfin, chaque agent calcule l'ensemble des arguments objectivement *acceptables*, en tenant compte des audiences spécifiques à chaque agent. Les correspondances qui n'ont alors que des arguments favorables sont retenues pour la communication.

Ce modèle permet effectivement de former des alignements en fonction des préférences de chaque agent. Ces préférences sont exprimées d'une part par les arguments dont chaque agent dispose dans son propre *système d'argumentation* (et qui forme le graphe général de l'argumentation). Mais également par la notion d'audience qui permet à chaque agent d'exprimer une position d'argumentation bien plus générale. En effet, classer les arguments en fonction de leur valeur permet de définir des directives de négociation indépendantes des correspondances discutées et des arguments formulés. Le déterministe logique de la théorie de l'argumentation est ici effacé par le choix de ces audiences. Elles donnent aux agents de véritables moyens d'expression quant aux termes de la négociation.

Nous donnons cependant quelques limites à ce modèle.

3.2. Limites du modèle

Limite 3.2.1

Nous pouvons faire plusieurs reproches au modèle de l'argumentation formulé par Loredana Laera [4]. Le premier concerne l'évaluation qu'elle en propose.

Cette évaluation est faite uniquement de manière empirique. Mais puisque la base du modèle est issue d'une théorie logique (les *systèmes d'argumentation* de Phan Minh Dung [1]), nous pouvons *a priori* donner une preuve analytique de la correspondance entre le résultat de la négociation inter-agent d'une part et la confrontation des *extensions préférées* calculées par la théorie d'autre part. De cette façon, le modèle gagnerait en légitimité.

Mais, l'évaluation proposée s'appuie essentiellement sur la comparaison du processus avec d'autres algorithmes d'alignement connus. Il faudrait formaliser par l'analyse cette évaluation.

Limite 3.2.2

Un second problème concerne les échanges faits entre les agents lors du processus. En effet, lors de la seconde étape, des *systèmes d'argumentations* entiers sont échangés pour que chaque agent puisse ensuite y répondre. Or par souci de confidentialité, nous pouvons exiger que le moins d'informations possible soit transmis aux agents avec lesquels on communique. Dans l'idéal, ne doit être transmis que l'information strictement nécessaire.

Puisque la négociation entre agents peut être assimilée à une application de la théorie des jeux (Phan Minh Dung le souligne d'ailleurs dans son article [1]), il est raisonnable de ne pas vouloir dévoiler notre entière stratégie lors de la négociation.

Il faut au contraire concevoir l'argumentation comme un véritable dialogue. Un peu à la manière de la négociation entre *matchers* formalisée par Cássia Trojahn [5]. Dans ce processus de négociation, un agent médiateur récupère tour-à-tour les arguments proposés par chaque agent. Il en transmet ensuite un à l'ensemble des agents pour récupérer les contre-arguments éventuellement formulés. Il procède ainsi par étapes, ne présentant qu'un argument à la fois. Ainsi, les stratégies d'argumentation se forment pas-à-pas, en fonction des réponses des autres agents. Il n'est pas nécessaire de donner d'un seul trait l'ensemble des arguments privés.

Pourtant, le processus défini à proprement parler par Loredana Laera (nous ne l'avons pas reformulé dans ce présent document) est composé d'actes, de mouvements, de réponses, etc. Il offre ainsi la possibilité d'un dialogue de type pas à pas. Nous nous interrogeons sur la non-utilisation de cette méthode par Loredana Laera.

Limite 3.2.3

Les arguments du modèle de Loredana Laera n'argumentent qu'en faveur ou en défaveur de correspondances. Or, nous pouvons envisager que des arguments plus généraux (qui n'appartiennent pas au domaine de l'alignement d'ontologies) viennent s'insérer dans l'argumentation et défendre ou attaquer d'autres arguments. Ce cas n'est pas envisagé ici.

De plus, les correspondances elles-mêmes ne participent pas à l'argumentation. Pourtant, lorsqu'un agent défend une correspondance, ce choix implique le respect d'une certaine cohérence interne à l'alignement ainsi partiellement formé. Nous pouvons imaginer, par souci de maintien de cette cohérence, qu'une correspondance puisse remettre en cause une autre correspondance. Au contraire, une correspondance qui s'intègre particulièrement bien à un alignement formé par d'autres correspondances pourrait être défendue par ces dernières.

Dans ce modèle, une correspondance peut éventuellement tenir un rôle dans le graphe des arguments. Un travail sur les inconsistances créées par le choix simultané de deux correspondances pourrait formaliser ce phénomène. Ceci permettrait d'améliorer la négociation pas à pas. Les choix d'un agent pouvant alors influencer ou contraindre les choix futurs.

3.3. Argumentation selon des degrés de confiance

De nouvelles approches sont introduites par Cássia Trojahn pour modifier encore le jeu des relations d'attaque.

Il s'agit dans un premier temps de donner un cadre à la coopération entre les différents *matchers* [5] afin de multiplier les sources d'arguments en faveur ou en défaveur des correspondances. Lors de cette négociation, ce sont les *matchers* à proprement parler qui argumentent et non des agents aux buts privés. L'objectif est alors de construire le meilleur alignement possible en prenant en compte les différentes expertises de ces *matchers*. Il n'y a pas ici de position à défendre par des agents autonomes (via des choix d'audiences), mais plutôt une volonté de faire coopérer les différents points de vue.

Cássia Trojahn étend ensuite le modèle de Trevor Bench-Capon en octroyant des degrés de confiance aux arguments [5]. Ces degrés représentent la crédibilité des arguments basée sur la confiance que nous leur accordons.

La démarche est la même que précédemment. Nous introduisons un second axe de classification pour juger des relations d'attaque entre les arguments. Il faut noter que ce nouveau classement n'est fonction d'aucune audience. En effet, les degrés de confiance sont des valeurs objectives (obtenues « de manière scientifique ») dont l'évaluation n'est pas fonction d'une quelconque position d'argumentation.

Un troisième axe est ensuite donné à l'aide de la fréquence d'apparition d'une correspondance au cours de la négociation entre différents *matchers* [6]. Il s'agit d'introduire la notion de *force* d'un argument en fonction des votes de ces *matchers*. Une correspondance *majoritaire* est alors plus forte qu'une correspondance qui a recueilli peu de votes. Cette théorie n'étant pas nécessaire pour les présents travaux, nous ne la développons pas.

Enfin, il faut remarquer que toutes ces précisions ont pour but de multiplier les moyens d'évaluation des arguments (et donc de sélection des correspondances). Nous voulons disposer de moyens d'action sur le sens de l'argumentation en fonction de différentes optiques (préférences privées, puissance objective des arguments, etc.). Il faut en effet rendre dynamique le jeu des influences et lui donner un sens.

Revenons à l'utilisation des degrés de confiance par Cássia Trojahn [5]. Nous définissons des *systèmes d'argumentations étendus*.

Définition 3.3.1

Un *système d'argumentation entendu basé sur des valeurs* est un 7-uplet $SAEV = \langle EA, att, V, val, P, C, valC \rangle$ où :

- $\langle EA, att, V, val, P \rangle$ est un *système d'argumentation basé sur des valeurs*
- C est un ensemble non-vide de degrés de confiance
- $valC$ est une fonction des éléments de EA vers les éléments de C

Encore une fois, nous redéfinissons les notions vues précédemment. La relation binaire de classement objectif des degrés de confiance est notée $valprefC$.

Définition 3.3.2

Un degré de confiance c_1 est *préféré* à un degré de confiance c_2 si et seulement si on a la relation $valprefC(c_1, c_2)$.

Un argument *attaque avec succès* un autre argument si il l'attaque dans le système d'argumentation classique et que nous avons plus confiance en lui, ou alors que nous ne préférons pas l'autre et que nous n'avons pas plus confiance en l'autre. Ainsi, le second axe (lié au degré de confiance) est prioritaire sur le premier axe (celui des préférences sur les arguments). En effet, il suffit qu'on ait confiance en un argument donné pour qu'il *attaque avec succès* un autre argument (peut importe alors l'audience choisie).

Définition 3.3.3

Un argument A *attaque avec succès* un autre argument B selon l'audience a si et seulement si $att(A,B) \wedge (valprefC(valC(A),valC(B)) \vee (\neg valpref(val(B),val(A)) \wedge \neg valprefC(valC(B),valC(A))))$. On note $att_a(A,B)$.

Définition 3.3.4

Un ensemble d'arguments S est *sans-conflit* selon une audience a si et seulement si $\neg (\exists (X,Y) \in S^2, att(X,Y) \wedge (valprefC(valC(X),valC(Y)) \vee (\neg valpref(val(Y),val(X)) \wedge \neg valprefC(valC(Y),valC(X))))$, i.e. si $\neg (\exists (X,Y) \in S^2, att_a(Y,X))$.

Nous définissons ensuite la notion d'*extension préférée* de la même manière que précédemment...

Les différentes valeurs données aux degrés sont fonctions de la méthode de mise en correspondance. Par exemple, pour un argument de type terminologique, la métrique de Levenshtein est utilisée (la distance entre deux chaînes de caractères est égale au nombre minimal de caractères qu'il faut supprimer, insérer ou remplacer pour passer de l'une à l'autre). Pour les arguments sémantiques, lorsque le *matcher* cherche dans une banque de relations sémantiques, si les deux termes comparés sont synonymes le degré de confiance est alors maximal. Les autres types de relations (holonymies, meronymies, hyponymies, hypernymies) sont liés à des degrés de confiance inférieurs.

Nous synthétisons l'état de l'art présenté dans les deux parties précédentes de cette manière :

[1] **Théorie de l'argumentation classique** (Phan Minh Dung)

↳ [2] **Introduction de préférences sur les arguments** (Leila Amgoud)

↳ [3] **Introduction de préférences sur la valeur des arguments** (Trevor Bench-Capon)

↳ [4] **Utilisation de l'argumentation pour l'alignement d'ontologies** (Loredana Laera)

↳ [5] **Introduction des degrés de confiance** (Cássia Trojahn)

↳ [6] **Introduction de la notion de fréquence d'apparition** (Cássia Trojahn)

4. Argumentation et contexte

4.1. Problématique

Nous disposons de plusieurs outils pour donner aux agents des directives d'argumentation (préférences sur les techniques de mise en correspondance, différents degrés de confiance accordés aux résultats de ces techniques, etc.). Cependant, ces directives sont très fortement liées aux techniques d'alignement elles-mêmes. Les travaux entrepris précédemment offrent une vue trop figée des arguments impliqués.

Nous aimerions donc inscrire le processus de négociation au sein d'un cadre plus général, lié directement aux nécessités de la communication des agents en présence. Un agent doit pouvoir sélectionner des correspondances en fonction du contexte dans lequel il se trouve et non seulement en fonction des techniques de mise en correspondance.

Nous introduisons la notion de contexte de l'argumentation pour désigner le cadre général dans lequel prend place la communication. Ce contexte représente en fait la totalité des informations auxquelles les agents ont accès au moment du dialogue et l'ensemble des objectifs qu'ils doivent atteindre. Et chacun de ces éléments peut *a priori* servir au choix des alignements.

L'objectif de nos recherches est donc de formaliser cette notion de contexte et de permettre aux agents d'orienter la négociation en fonction de celle-ci. Nous envisageons pour cela quatre solutions compatibles les unes avec les autres.

Solution 4.1.1 – Bien choisir les audiences

La première remarque que nous faisons concerne l'implémentation des agents eux-mêmes. Lorsque nous programmons un agent autonome destiné à accomplir une tâche précise, il faut se poser la question de ses positions d'argumentation. Quelles audiences choisira-t-il si il vient à devoir négocier avec d'autres agents ? Le thème de la conversation, directement lié aux objectifs qu'il cherche à atteindre, doit orienter le choix de ces audiences.

Mais il n'est pas toujours évident de faire un lien logique entre des préférences sur les *matchers* et une tâche à accomplir.

Solution 4.1.2 – Définir des arguments contextuels

Cette solution consiste à produire des arguments qui dépendent du cadre de la négociation. La plupart des *matchers* que nous avons évoqués font un travail préliminaire à la négociation. Ils proposent des arguments statiques indépendant de systèmes multi-agents donnés. En dotant les agents de leurs propres *matchers*, nous pouvons leur permettre de produire dynamiquement des arguments en fonction de la conversation spécifique qu'ils se préparent à avoir.

Cette approche est développée dans la **partie 4.2** à l'aide d'arguments de type *extensional*.

Solution 4.1.3 – Utiliser des arguments de cohérence

Nous avons constaté que les correspondances pouvaient avoir un rôle au sein de l'argumentation (cf. **limite 3.2.3**). En développant cette idée, sur des bases de cohérence interne aux alignements, nous exprimons des arguments de type *contextual*. Ces arguments prennent en compte les choix faits par les différents agents au sein d'un processus de négociation donné pour juger des autres choix. Ils sont donc créés de manière dynamique (cf. **partie 4.3**).

Solution 4.1.4 – Avoir un historique de l'argumentation

Enfin, nous proposons une formalisation du contexte de l'argumentation et du processus de négociation qui y prend place. Nous introduisons à partir de ces bases une notion d'historique de l'argumentation. Cet historique est composé des correspondances précédemment acceptées ou refusées par le système, posant ainsi une base préliminaire à la négociation.

Cette notion, formalisée dans la **partie 4.4**, permet d'avoir un regard critique sur les argumentations présentes à l'aide des choix effectués dans le passé.

L'intérêt de contextualiser l'argumentation est d'orienter l'alignement d'ontologies en fonction du rôle qu'auront ces ontologies dans la conversation elle-même. Les correspondances entre les termes de ces ontologies doivent dépendre de l'information qu'ils sont sensés véhiculer ou de la connaissance qu'ils sont sensés représenter. Ainsi, les alignements formés seront sensibles aux problèmes spécifiques que les agents essaient de résoudre.

Il ne s'agit plus de faire une traduction « mot-à-mot » en fonction d'arguments d'ordre linguistique uniquement. Au contraire, la traduction doit être liée aux nécessités du dialogue, à son contexte, et prendre en compte sa raison d'être.

4.2. Arguments contextuels

Dans cette section, nous proposons une solution pour le cas particulier des ontologies destinées à recueillir des données. Elles sont ici chargées de regrouper et de structurer des informations sous différentes étiquettes (les termes de l'ontologie). Nous distinguons alors l'ontologie sous sa forme abstraite, c'est-à-dire telle qu'elle a été créée pour représenter un champ de la connaissance (cf. **figure 1**), et l'extension donnée à une telle ontologie, qui représente l'ensemble de ses cas d'utilisation (cf. **figure 3**).

Nous appelons (de manière abusive) *instance de l'ontologie* un cas concret d'utilisation de cette ontologie. Et nous appelons *extension d'une ontologie* l'ensemble de ces instances.

Pour mettre en place une correspondance entre deux termes, nous pouvons discuter à propos du nom donné à ce terme (argument terminologique), à propos du ou des sens que nous donnons à ce terme (argument sémantique) ou à propos de l'agencement des différents termes les uns vis-à-vis des autres (argument structurel). Toutes ses approches sont liées à la forme abstraite de l'ontologie et non pas aux données qu'elle ordonne effectivement. Le but étant de formuler des règles générales sur les termes qui impliquent des règles particulières sur les données.

Mais lorsqu'un agent cherche à échanger des informations avec un autre agent, il dispose des données elles-mêmes. Et elles sont exploitables pour la mise en correspondance. Loredana Laera regroupe ce genre d'argument sous la valeur *extensional*.

Pour revenir à notre exemple, supposons que l'agent 1 veuille acheter à l'agent 2 une place de cinéma pour aller voir un film de science-fiction. Nous cherchons à aligner les deux ontologies utilisées.

Nous pouvons supposer que l'agent 1 dispose d'un certain nombre d'instances de l'ontologie 1 (il dispose d'une extension cette ontologie) qui représentent des films du genre. Ainsi, le terme *titre du film* peut être étendu (c'est-à-dire instancié) par plusieurs chaînes de caractères comme "2001 : l'Odyssée de l'Espace", "Star Wars" et "Sunshine". Une recherche des chaînes de caractères qui étendent le terme *nom film* de l'ontologie 2 (dans l'extension possédée par l'agent 2) pourrait mettre en corrélation les deux termes. En effet, si suffisamment de noms sont similaires au sein des deux extensions, et si l'on trouve ces noms uniquement pour des deux termes, il y a de fortes chances pour qu'ils puissent être mis en correspondance.

Figure 3
Exemple d'instanciation des ontologies 1 et 2

Instance 1 (de l'ontologie 1)
<ul style="list-style-type: none"> ▶ <i>titre_du_film</i> : "2001 : l'Odyssée de l'Espace" ▶ <i>genre_du_film</i> : "Science-Fiction" ▶ <i>annee_de_sortie</i> : "1968" ▶ <i>equipe</i> <ul style="list-style-type: none"> ▶ <i>nom_du_realisateur</i> : "Stanley Kubrick" ▶ <i>nom_du_producteur</i> : "Stanley Kubrick" ▶ <i>equipe_realisation</i> <ul style="list-style-type: none"> ▶ <i>scenaristes</i> : "Arthur C. Clarke, Stanley Kubrick" ▶ <i>monteurs</i> : "Ray Lovejoy" ▶ <i>equipe_tournage</i> <ul style="list-style-type: none"> ▶ <i>acteurs</i> : "Keir Dullea, Gary Lockwood, William Sylvester" ▶ <i>cadreurs</i> : "Kelvin Pike"
Instance 2 (de l'ontologie 2)
<ul style="list-style-type: none"> ▶ <i>nom_film</i> : "2001 : a Space Odyssey" ▶ <i>genre</i> : "SF" ▶ <i>date</i> : "1968" ▶ <i>budget</i> : "NC" ▶ <i>equipe</i> <ul style="list-style-type: none"> ▶ <i>realisateurs</i> : "Stanley Kubrick" ▶ <i>producteurs</i> : "Stanley Kubrick, Victor Lyndon" ▶ <i>scenaristes</i> : "Stanley Kubrick, Arthur C. Clarke" ▶ <i>acteurs_principaux</i> : "Keir Dullea, Gary Lockwood, William Sylvester" ▶ <i>acteurs_secondaires</i> : "Daniel Richter, Douglas Rain"

Nous voulons formaliser cette méthode. Pour cela, nous définissons la notion d'*ontologie simple*, qui est une réduction de la définition d'ontologie donnée par Loredana Laera [4], afin de simplifier notre raisonnement.

Définition 4.2.1

Une *ontologie simple* est un couple $\langle T, \text{ordre}T \rangle$ où :

- T est l'ensemble des termes de l'ontologie
- $\text{ordre}T$ est une relation d'ordre partiel sur T (définissant la structure de l'ontologie)

Pour l'ontologie 1, nous avons par exemple :

- $T = \{ \text{titre_du_film}, \text{genre_du_film}, \text{annee_de_sortie}, \text{equipe}, \text{nom_du_realisateur}, \text{nom_du_producteur}, \text{equipe_realisation}, \text{scenaristes}, \text{monteurs}, \text{equipe_tournage}, \text{acteurs}, \text{cadreurs} \}$
- $\text{ordre}T = \{ (\text{equipe}, \text{nom_du_realisateur}), (\text{equipe}, \text{nom_du_producteur}), (\text{equipe}, \text{equipe_realisation}), (\text{equipe}, \text{equipe_tournage}), (\text{equipe_realisation}, \text{scenaristes}), (\text{equipe_realisation}, \text{monteurs}), (\text{equipe_tournage}, \text{acteurs}), (\text{equipe_tournage}, \text{cadreurs}) \}$

Définition 4.2.2

Une *instance* est un triplet $\langle T, \text{ordre}T, D, \text{etend} \rangle$ où :

- $\langle T, \text{ordre}T \rangle$ est une *ontologie simple*
- D est un ensemble de données (ici des chaînes de caractères) étendant les termes de T
- *etend* est une fonction des termes de T vers les données de D

Pour l'instance 1 nous avons par exemple :

- T et *ordre* T sont ceux de l'ontologie 1
- $D = \{ \text{"2001 : l'Odyssée de l'Espace"}, \text{"Science-Fiction"}, \text{"1968"}, \text{"Stanley Kubrick"}, \text{"Arthur C. Clarke, Stanley Kubrick"}, \text{"Ray Lovejoy"}, \text{"Keir Dullea, Gary Lockwood, William Sylvester"}, \text{"Kelvin Pike"} \}$
- et la fonction *etend* associe correctement les termes de T aux chaînes de caractères de D

Puisque nous allons comparer des données, il faut définir les règles de similarité qui peuvent exister entre elles. Nous donnons une définition informelle et nous notons que le choix de ces similarités sont fonctions de différentes approches (comme pour la mise en correspondance).

Définition 4.2.3

Deux données $d1$ et $d2$ sont *similaires* si et seulement si elles peuvent être mises en correspondance de manière terminologique ou sémantique.

Dans l'exemple, certaines données sont *similaires* : "2001 : l'Odyssée de l'Espace" et "2001 : a Space Odyssey" sont sémantiquement *similaires*, "Science-Fiction" et "SF" également, "1968" et "1968" sont lexicalement *similaires* (en fait elle sont lexicalement *identiques*), etc.

Définition 4.2.4

Soient deux instances $i1 = \langle T1, \text{ordre}T1, D1, \text{etend}1 \rangle$ et $i2 = \langle T2, \text{ordre}T2, D2, \text{etend}2 \rangle$.

Soient deux termes $t1$ et $t2$ appartenant respectivement aux ensembles $T1$ et $T2$.

Les termes $t1$ et $t2$ sont *similaires en extension* selon $i1$ et $i2$ si et seulement si $\text{etend}1(t1)$ et $\text{etend}2(t2)$ sont *similaires*.

Dans l'exemple, les termes *similaires en extensions* selon l'instance 1 et l'instance 2 sont les couples suivants : (*titre_du_film*, *nom_film*), (*genre_du_film*, *genre*), (*annee_de_sortie*, *date*), (*nom_du_realisateur*, *realisateurs*), (*acteurs*, *acteurs_principaux*), etc. Cette similarité est issue de la similarité des données que nous avons notée précédemment.

Il faut maintenant généraliser cette recherche de similarités à des extensions.

Définition 4.2.5

Une *extension* E est un ensemble d'instances qui étendent une seule et même *ontologie simple*. Nous notons $\langle TE, \text{ordre}TE \rangle$ cette ontologie.

Définition 4.2.6

Soient deux *extensions* $E1$ et $E2$.

Soient deux termes $t1$ et $t2$ appartenant respectivement aux ensembles $TE1$ et $TE2$.

Le *nombre de similarités en extension* des termes $t1$ et $t2$ vis-à-vis des extensions $E1$ et $E2$ est égal au cardinal de l'ensemble $\{(e1, e2) \in E1 \times E2 \mid t1 \text{ et } t2 \text{ sont similaires en extension selon } e1 \text{ et } e2\}$.

Le *nombre de similarités en extension* est donc le nombre de couples d'instances pour lesquels les données des termes $t1$ et $t2$ sont *similaires*. Nous pouvons exploiter ce nombre pour mettre en correspondance les deux termes.

Cette méthode produit en fait un argument de valeur *extensional*. Nous définissons ici un nouveau type de *matcher* qui s'appuie sur les données en présence pour mettre en correspondance des termes.

Il faut impérativement utiliser les degrés de confiance introduits par le travail de Cássia Trojahn [5] pour décrire ce genre d'argument. La confiance portée à un argument *extensional* est en effet directement liée au *nombre de similarités en extension*. Nous pourrions définir celui-ci par le rapport entre le *nombre de similarités* et le minimum des nombres d'instances présentes dans les extensions E1 et E2.

Définition 4.2.7

Soient la correspondance $c = \langle t1, t2, \equiv \rangle$ et l'argument $a = \langle \text{extensional}, c, + \rangle$ formulé à l'aide des *extensions* précédentes.

Le degré de confiance de l'argument a est égal au *nombre de similarités en extension* des termes $t1$ et $t2$ vis-à-vis des extensions E1 et E2 divisé par $\min(\#E1, \#E2)$.

Lorsque toutes les instances de E1 ont été liées de cette manière à au moins une instance de E2, le degré de confiance est supérieur ou égal à 1 (on le réduit à alors 1). Si aucun couple n'a pu être formé, la confiance est nulle.

Il faut noter que les extensions doivent être de taille conséquente pour éviter d'obtenir des résultats dégénérés (par exemple, une extension contenant une unique instance qui a par hasard une similarité avec une instance contenue dans une autre extension très fournie).

Les arguments produits par ces *matchers* ne sont plus liés uniquement à la forme abstraite de l'ontologie. Ils dépendent des extensions utilisées pour la mise en correspondance. Ces extensions fournissent donc une base pour la négociation. Nous avons défini des arguments qui sont fonctions du contexte d'utilisation des ontologies que l'on cherche à aligner.

Pour une ontologie donnée, si nous prenons pour base la totalité de ses extensions présentes sur le Web, nous avons un contexte regroupant tous les cas d'utilisation jamais effectués. Les alignements créés à partir d'un tel contexte d'utilisation sont donc indépendants des agents en présence. Mais nous constatons que ce contexte est modulable. En effet, en prenant un ensemble d'instances restreint, nous orientons la recherche de similarités vers un domaine de données plus spécifique.

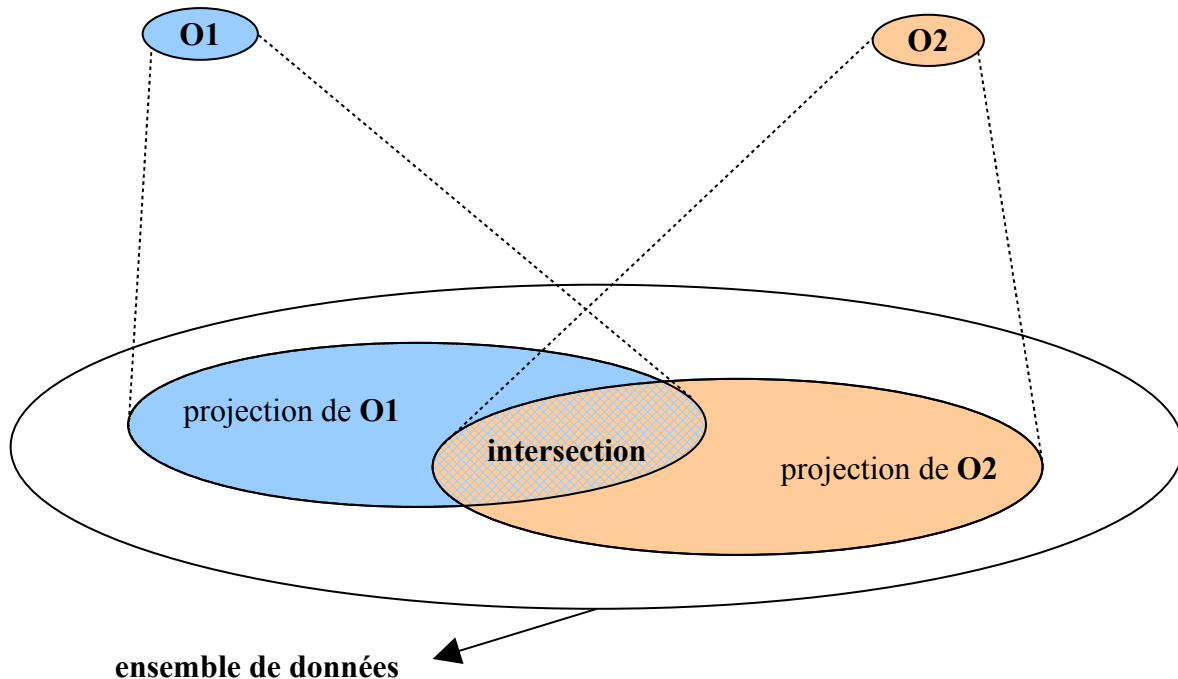
Les arguments *extensional* dépendent donc des données effectives que les agents vont ensuite échanger. Ils sont liés directement au contexte de l'argumentation parce qu'ils s'adaptent à la finalité de cette argumentation : l'échange des données en question. C'est en ce sens que nous avons ici produits des *arguments contextuels*.

Avant d'entamer la négociation, les deux agents peuvent décider sur quelles bases ils vont effectuer celle-ci. En définissant les extensions qui vont être utilisées, ils définissent le contexte d'utilisation qui servira à l'argumentation. Le choix de ces extensions est donc important, il doit correspondre à l'utilisation des données que feront effectivement les agents.

Pour résumer, cette méthode de création d'argument se base sur la projection des ontologies sur un ensemble de données (cf. **figure 4**). Une extension correspond à une projection selon plusieurs instances. A quelques similarités près, les projections peuvent être d'intersection non-vide. Cette intersection varie bien entendu en fonction de l'ensemble de données que nous avons choisi et des extensions utilisées.

Nous nous servons du choix de cet ensemble pour lier de manière contextuelle les ontologies projetées en fonction de l'intersection ainsi réalisée.

Figure 4
Principe de la projection d'ontologies



4.3. Argumentation cohérente

Dans un premier temps, nous souhaitons apporter une amélioration aux *systèmes d'argumentation*. Il s'agit de donner une place aux correspondances dans le graphe des arguments (cf. **limite 3.2.3**). Nous donnons ici des principes informels modifiant les relations d'attaque entre les arguments.

L'argumentation a pour but de produire des alignements cohérents. Or, deux correspondances peuvent sembler incompatibles ou, au contraire, liées l'une à l'autre. Nous évoquons diverses possibilités d'incohérence ou de liaison entre les correspondances.

Exemple 4.3.1

Soient les entités e_1 et e_1' d'une ontologie 1 telles que e_1 est plus générale que e_1' . Soient les entités e_2 et e_2' d'une ontologie 2 qui répondent aux mêmes critères. Si nous argumentons en faveur d'une correspondance c_1 entre e_1' et e_2 , il est illogique d'argumenter en faveur d'une correspondance c_2 entre e_1 et e_2' .

Un argument qui défend c_1 attaque également c_2 .

Exemple 4.3.2

Au contraire, si deux entités e_{11} et e_{21} (précisant l'entité e_1 dans une ontologie 1) ont été mises en correspondance avec deux autres entités e_{21} et e_{22} (précisant l'entité e_2 dans une ontologie 2), il est raisonnable de penser que les entités e_1 et e_2 peuvent être mises en correspondance, afin de garder une structure cohérente vis-à-vis des choix précédents.

Un argument qui défend les deux premières correspondances défend également la troisième.

4.4. Historique d'argumentation

Dans cette section, nous donnons une définition au contexte de l'argumentation. Puis nous formalisons un processus de négociation à partir de l'utilisation de ce contexte.

Définition 5.1

Le *contexte d'une argumentation* est un 9-uplet $\langle EA, att, V, val, P, Ag, O, okC, koC \rangle$ où :

- $\langle EA, att, V, val, P \rangle$ est un *système d'argumentation basé sur des valeurs*
- $Ag = \{Ag_1, \dots, Ag_n\}$ est l'ensemble des n agents en présence
- O est l'ensemble des ontologies utilisées par les agents
- okC est l'ensemble des correspondances qui ont été acceptées par les agents
- koC est l'ensemble des correspondances qui ont été refusées par les agents

Nous construisons le contexte dynamiquement lors du processus d'argumentation. Les seuls éléments statiques sont Ag et O , respectivement l'ensemble des agents en présence et l'ensemble des ontologies qu'ils utilisent.

Il faut également noter que les ensembles okC et koC sont disjoints.

Nous résumons le processus d'argumentation de la manière suivante.

Initialisation :

- les ensembles EA, att, V, val, P, okC et koC sont vides
- les ensembles Ag et O sont initialisés avec les données du système multi-agents

Processus :

- 1) Un agent Ag_i propose d'ouvrir la négociation à propos d'une correspondance c entre les entités de deux ontologies de O .
- 2) Chaque agent formule des arguments à partir de son *système d'argumentation* privé en faveur ou en défaveur de la correspondance c . Pour chaque argument formulé de cette manière, nous mettons à jour les ensembles EA, att, V, val et P .
- 3) Lorsqu'aucun agent n'a de nouvel argument à formuler, chaque agent choisit une audience parmi celles de P et transmet son choix aux autres. Chacun calcule ensuite les différentes *extensions préférées* ainsi créées.
- 4) Les arguments *objectivement acceptables* sont retenus. Si la correspondance n'a alors que des arguments favorables, elle est acceptée par le système et nous l'ajoutons à l'ensemble okC . Sinon, elle est refusée et nous l'ajoutons à l'ensemble koC .
- 5) Nous réitérons le processus à partir de l'étape 1 jusqu'à ce qu'aucune nouvelle correspondance ne soit proposée.

Remarques :

- Lors de l'étape 4, la sélection des arguments retenus doit être déterminée comme une règle de l'argumentation. Ainsi, nous pouvons envisager de retenir un argument si il est *objectivement acceptable*, ou alors si il est *subjectivement acceptable*, ou encore si il appartient à une majorité d'*extensions préférées* (nous disons alors qu'il est *majoritairement acceptable*).

Quoi qu'il en soit, il faut fixer ce mode de sélection lors de l'initialisation du processus, et ce en fonction de l'importance de la conversation. Pour une négociation critique par exemple (dont les conséquences sont à prendre « au sérieux ») nous préférons ne retenir que les arguments *objectivement acceptables*.

- Lors de l'étape 3, les agents choisissent une position d'argumentation en sélectionnant une audience parmi toutes celles qui sont possibles. Nous pouvons exiger que ce choix soit restreint par les choix effectués lors des tours de négociation précédents. Un agent peut alors choisir une

audience si et seulement si elle ne modifie pas le classement des valeurs qu'il avait auparavant. En effet, puisque l'argumentation se fait pas-à-pas, modifier les préférences d'un agent consisterait à revenir sur l'argumentation élaborée pour les correspondances précédentes. En outre, nous pouvons exiger qu'un agent garde une position cohérente tout au long du processus (i.e. un classement des valeurs qui reste stable).

Au contraire, nous pouvons décider de permettre aux agents de changer de position à n'importe quel moment du processus. Ceci n'est pas en contradiction avec la plupart des débats menés par l'homme où l'on adapte sa position au thème de la conversation à un temps donné. Mais il peut y avoir une rétroaction sur l'argumentation à propos des thèmes discutés auparavant. Cette rétroaction, si nous la prenons en compte et que nous recalculons l'issue des argumentations précédentes, peut ouvrir sur un débat sans fin où les changements de positions réaffectent constamment les bases établies.

C'est pourquoi nous préférons stocker les décisions prises à propos des correspondances dans les deux ensembles okC et koC pour poser à chaque étape les nouvelles bases du dialogue.

- Il faut enfin préciser que l'ordre d'évaluation des correspondances a ici une grande importance. En effet, le rejet (ou l'acceptation) des premières correspondances peut fortement influencer la sélection des suivantes (cf. **partie 4.3**). Il faut alors définir clairement cet ordre d'évaluation, ce qui impose une négociation préalable entre les agents.

Pour répondre à ce problème, nous pouvons envisager de soumettre à l'argumentation des ensembles de correspondances. Les arguments sont alors formulés à propos de ces ensembles.

- En outre, nous répondons à la **limite 3.2.2** concernant l'échange entre les agents. Ici, les arguments sont donnés un à un. Nous n'avons pas besoin de transmettre des *systèmes d'argumentation* entiers pour procéder à la négociation.

- Enfin, nous pouvons introduire de nouveaux arguments de type *contextual* en fonction des considérations faites dans la section précédente (cf. **partie 4.3**). Ces arguments sont formulés à partir des ensembles okC et koC pour rappeler les règles de cohérence interne. Ils imposent le respect des bases données au dialogue.

Enfin, l'intérêt de ce processus d'argumentation dynamique est de pouvoir poser petit à petit les bases de la négociation. Il faut de préférence commencer par discuter à propos des correspondances qui font l'unanimité (si nous disposons un moyen de les reconnaître) pour entamer les débats épineux avec des bases pré-établies.

Mais surtout, nous pouvons conserver un historique de la négociation, et utiliser cet historique pour les négociations futures. Ainsi, si deux agents utilisent des alignements pour résoudre avec succès un problème précis, ils peuvent commencer leur prochain processus de négociation avec des ensembles okC et koC non-vides. Nous pouvons également envisager de noter ces historiques en fonction de l'efficacité qu'ils ont eu sur le dialogue. Ainsi, les agents peuvent avoir une mémoire de leurs précédentes communications. Ils peuvent adapter le choix de leurs alignements en fonction de la pertinence de ceux qu'ils ont sélectionnés auparavant.

5. Conclusion

Nous résumons les différentes parties de ce dossier de la manière suivante :

- **Partie 1.2** – Introduction au problème de l'alignement d'ontologies pour le Web Sémantique dans le cas de communications inter-agents.
- **Parties 2 et 3** – État de l'art concernant les différentes théories de l'argumentation et étude de leur utilisation pour résoudre ce problème.
- **Partie 3.2** – Constat sur les différentes limites que présente cette utilisation.
- **Partie 4.1** – Problématique de notre sujet concernant la volonté de contextualiser l'argumentation afin de réduire certaines de ces limites.
- **Partie 4** – Tentative de résolution de notre problématique :
 - 1) A l'aide d'arguments liés au contexte du dialogue inter-agent.
 - 2) En proposant un nouveau processus de négociation basé sur un historique d'argumentation.

La **solution 1** utilise les données structurées par les ontologies pour définir des arguments contextuels. A partir de ces données, les agents produisent des alignements spécifiques aux finalités du dialogue. Ils utilisent l'information qu'ils vont échanger pour convenir des modalités mêmes de cet échange.

La **solution 2** propose de garder en mémoire les anciennes négociations afin d'entamer les nouvelles sur des bases solides et contrôlées. Cette méthode permet au système d'avoir un recul sur les alignements produits dans le passé. Cet historique de la négociation permet ainsi de perfectionner les alignements.

Bilan :

Nous pouvons être satisfaits des apports de la **solution 1** à notre modèle d'argumentation. En effet, elle prend en compte les différentes théories exprimées à ce sujet et s'y inscrit harmonieusement. La définition de *matchers* contextuels semble en effet répondre correctement à notre problématique et utilise les bases de notre recherche.

Par contre, la formalisation incomplète de la **solution 2** ne la rend pas exploitable immédiatement. Les règles de cohérence interne gagneraient à être explicitées. Mais nous exprimons cependant une notion attrayante qui peut se révéler féconde si nous la retravaillons de manière plus formelle.

Enfin, ce qu'il manque définitivement à ces travaux, c'est une évaluation au moins empirique des solutions qu'ils proposent. Des outils ont été développés sur le Web pour comparer les différents algorithmes d'alignement. OAEI (*Ontology Alignment Evaluation Initiative*) propose en effet des méthodes de comparaison de performances pour ces algorithmes. Certains travaux référencés dans notre bibliographie utilisent d'ailleurs ces outils pour évaluer leurs propres solutions.

Mais le temps dont nous disposons pour effectuer les présents travaux ne nous a pas permis de les utiliser.

Bibliographie

- [1] Phan Minh Dung, **On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n -Person Games**, *Artificial Intelligence* 77(2):321-358, 1995
- [2] Yannis Dimopoulos, Pavlos Moraitis, Leila Amgoud, **Theoretical and Computational Properties of Preference-based Argumentation**, *European Conference on Artificial Intelligence (ECAI 2008)*, Patras, Grèce, 21/07/08-25/07/08, IOS Press, p. 463-467, 2008
- [3] Trevor Bench-Capon, Sylvie Doutre, Paul Dunne, **Audiences in Argumentation Frameworks**, *Artificial Intelligence*, Elsevier, Vol. 171 N. 1, p. 42-71, janvier 2007
- [4] Loredana Laera, Ian Blacoe, Valentina Tamma, Terry Payne, Jérôme Euzenat, Trevor Bench-Capon, **Argumentation over Ontology Correspondences in MAS**, in: Proc. 6th International conference on Autonomous Agents and Multiagent Systems (AAMAS), Honolulu (HA US), pp1285-1292, 2007
- [5] Cássia Trojahn dos Santos, Márcia Cristina Moraes, Paulo Quaresma, Renata Vieira, **A Cooperative Approach for Composite Ontology Mapping**, *Journal on Data Semantics* 10:237-263, 2008
- [6] Antoine Isaac, Cassia Trojahn, Shenghui Wang, Paulo Quaresma, **Using quantitative aspects of alignment generation for argumentation on mappings**, Proc. ISWC-2008 Ontology latching workshop, pp49-60, Karlsruhe (DE), 2008

Remerciements

Je tiens à remercier Jérôme Euzenat pour m'avoir permis de travailler sur ce sujet, pour toutes les conversations fascinantes que nous avons eues, pour ses rectifications minutieuses et pour avoir corrigé un nombre incalculable de mes fautes d'orthographe.

Je tiens également à remercier Florence Maraninchi pour m'avoir mis sur une si bonne piste.