International Journal of Artificial Intelligence,
ISSN 0974-0635; Int. J. Artif. Intell.
Spring (March) 2012, Volume 8, Number S12

Exploiting Majority Acceptable Arguments for Ontology Matching

Cassia Trojahn[1], Paulo Quaresma[2], Renata Vieira[3]

[1]INRIA & LIG
Grenoble, France
cassia.trojahn@inrialpes.fr

[2]Departamento de Informática, Universidade de Évora
Évora, Portugal
pq@di.uevora.pt

[3]Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul
Porto Alegre, Brazil
renata.vieira@pucrs.br

ABSTRACT

Ontology matching consists of generating an alignment (set of correspondences) from a pair of ontologies. This process has been seen as a mainstream solution to the semantic heterogeneity problem in ontology-based systems. A wide diversity of matching solutions has been proposed, which exploit different features within an ontology. Matching systems usually differ in their results and an important issue is to combine different matching results and deal with potential conflicts that arise from the different views. Our approach exploits argumentation theory as a way for dealing with that issue. Here, arguments are as positions that support or reject correspondences and argumentation frameworks support the creation and exchange of arguments, followed by the reasoning on their acceptability. First, matchers generate their correspondences and represent them as arguments. Next, they share their arguments and interpret them on the basis of argumentation frameworks and individual preferences. As a result, each matcher has a subset of acceptable arguments, from the set of arguments initially shared. The subset of globally acceptable arguments (consensus) is computed from the individual. In this paper, we exploit the notion of majority, where arguments being acceptable by the majority of matchers are considered as a consensus on the initial alignments. We evaluate our proposal on a standard set of alignments. Considering the correspondences represented as arguments acceptable for the majority of individual subsets, both precision and recall are improved, specially when compared with the subsets acceptable for every matcher or for some matchers.

Keywords: ontology matching, argumentation frameworks, evaluation.

Mathematics Subject Classification: 68T01.

1 Introduction

An ontology defines the set of concepts (by specifying the vocabulary) that represent the knowledge in a domain. In this way, ontologies aim at specifying a shared understanding of a domain that can be communicated between applications (Fensel, 2003). Although ontologies have been seen as a way for sharing knowledge in the Web, the existing ontologies might be themselves sources of heterogeneity. They may differ in granularity, representation and the way in which they model the concepts, properties and axioms. Ontology matching has been seen as a mainstream solution to the semantic heterogeneity problem in ontology-based systems.

Ontology matching is the process of finding correspondences between two ontologies. Many different techniques to the matching problem in found in the literature (Euzenat and Shvaiko, 2007). The distinction between them is accentuated by the manner in which they exploit the features within an ontology. Whereas lexical techniques consider measures of string similarity; semantic ones consider semantic relations usually on the basis of lexical oriented linguistic resources; while structural techniques consider entity locations in the ontology hierarchy. Each category offers a wide diversity of options.

Due to that diverse way of exploring the problem, matching systems generally differ in the alignments proposed for two ontologies. Some approaches will perform better than others for specific ontologies, depending on how well the technique fits the material available. Furthermore, approaches that perform well for some ontology or task may not be successful in others. A way to combine the different matching approaches and solve the potential conflicts in the alignments provided by them is via an agreement process. This process can be carried out by using argumentation theory.

Our approach exploits argumentation as a way for supporting the creation and exchange of arguments that represent positions in favour or against correspondences between entities of ontologies, as well as for supporting the reasoning on the acceptability of these correspondences. Different matchers, working on different approaches, generate their set of correspondences. Next, they share their arguments and interpret them on basis of argumentation and individual preferences. The subset of globally acceptable arguments (consensus) is computed from the individual subsets of acceptable arguments. The underlying argumentation theory we apply here is the strength-based argumentation framework (Trojahn, Quaresma, Vieira and Moraes, 2008), (Trojahn, Quaresma and Vieira, 2008a), (Trojahn, Quaresma and Vieira, 2008b) (SVAF), which has been specified for associating arguments to the confidence that matchers have in the similarity between the entities of the ontologies being matched. This confidence is usually derived from similarity assessments made during the matching process.

In this paper, we exploit the notion of majority in combining the individual subsets of acceptable arguments. Arguments being acceptable by the majority of matchers are considered as the consensus on the alignments generated by the different matchers. According to (Coste-Marquis, Devred, Konieczny, Lagasquie-Schiex and Marquis, 2007), this kind of voting mechanism is inadequate to characterise the notion of acceptable arguments at a group level. However, this is only true when the different agents do not share the same initial set of arguments. In our case, voting makes sense because all matchers share their arguments with each others and have the same set of arguments at the start.

We focus on the practical evaluation of our proposal on a standard set of alignments. Considering the subset of arguments being acceptable for the majority of matchers we improve both precision and recall, when compared with the subsets being acceptable for every matcher (called objectively acceptable) and for some matcher (called subjectively acceptable). For this problem we give special attention to precision because usually task operations, processes or problem-solving issues are based on ontologies, thus higher precisions allow systems operate more on the safe side when interoperability is treated through ontology matching.

The rest of the paper is organised as follows. First, we introduce some definitions on ontology matching and present argumentation frameworks upon which our model is defined (Section 2). We then present the argumentation approach for combining different matching approaches (Section 3), detailing the notions of argument for the context of ontology matching and our argumentation model based on confidence of correspondences. Next, we present the evaluation

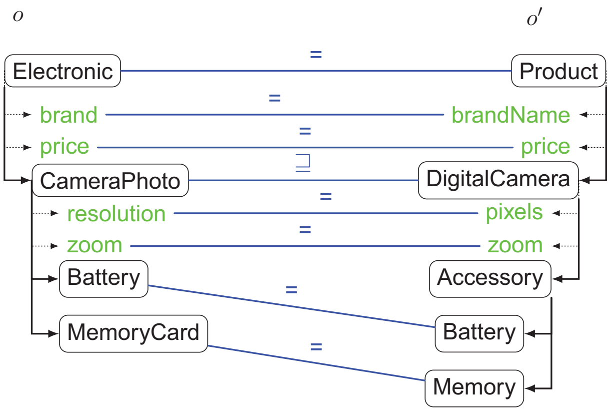

Figure 1: Fragments of ontologies o and o' with alignment A.

of our model (Section 4), which is carried out on standard alignments and a set of different matchers. Finally, we present the main related work (Section 5) and discuss the strengths and weaknesses of our approach (Section 6).

2 Foundations

2.1 Ontology matching

An ontology typically provides a vocabulary describing a domain of interest and a specification of the meaning of terms in that vocabulary (Euzenat and Shvaiko, 2007) usually identifying elements such as classes, individuals, relations, attributes and axioms. Ontology matching consists of generating an alignment (A) from a pair of ontologies (o and o'). An alignment A is a set of correspondences between vocabulary items representing entities of the same type:

Definition 2.1 (Correspondence)**.** Given two ontologies, o and o', a correspondence is a quadruple:

$$\langle e, e', r, n\rangle$$

where $e \in o$, $e' \in o'$ are the entities (e.g., formulas, terms, classes, individuals) of the ontologies; r is the relation between e and e', taken from the set of alignment relations (e.g., $\equiv$, $\sqsubseteq$, or $\sqsupseteq$); and $n \in [0\ 1]$ is a confidence level (e.g., measure of confidence in the fact that the correspondence holds).

Definition 2.2 (Alignment)**.** Given two ontologies, o and o', an alignment A is a set of correspondences.

Figure 1 illustrates an example of alignment A between two ontologies, o and o'. In A we have, for instance, the correspondences:

$$\langle Electronic_o, Product_{o'}, \equiv, 1.0\rangle \quad (2.1)$$
$$\langle CameraPhoto_o, DigitalCamera_{o'}, \sqsubseteq, 0.9\rangle \quad (2.2)$$

Many different approaches to the problem of ontology matching have been proposed in the literature (Euzenat and Shvaiko, 2007). The distinction between them is reinforced by the manner in which they exploit the features within an ontology. The matching techniques can be grouped into broad categories: lexical (detecting string similarities usually between labels of concepts and properties), semantic (the terms can be evaluated semantically, usually on the basis of semantic-oriented linguistic resources[1], structural (using the structure of the ontologies), and instance-based matching (using instance data to detect the similarity between concepts). However, each category offers a wide diversity of options as well as different classifications for the approaches have been proposed in the literature. Furthermore, many ontology matching systems rely not on a single approach, but combine several approaches.

2.2 Argumentation frameworks

Argumentation can be seen as a process based on the construction and comparison of arguments, followed by the reasoning on their acceptability. The central notion in argumentation systems is the notion of *acceptability*. Two kinds of *acceptability* have been considered (Amgoud and Cayrol, 2002):

- *Individual acceptability*: where a level is assigned to each argument depending on the existence of direct defeaters (or counter-arguments). In this case, the acceptability of an argument depends only on its properties.
- *Joint acceptability (or collective)*: which relies on a notion of defense. The set of all arguments that a rational agent may accept must defend itself against any defeater.

Following the *joint acceptability*, an argument can be defended by other arguments. Such notion was introduced by (Dung, 1995). In Dung's model, the acceptability of an argument is based on a reasonable view: an argument should be accepted only if every attack on it is attacked by an accepted argument. An argument is an abstract entity whose role is determined by its relation to other arguments.

Dung defines an argumentation framework as follows.

Definition 2.3 (Argumentation Framework (AF) (Dung, 1995))**.** An AF is a pair $\langle \mathcal{A}, \ltimes \rangle$, where $\mathcal{A}$ is a set of arguments and $\ltimes$ is a binary relation on $\mathcal{A}$, i.e., $\ltimes \subseteq \mathcal{A} \times \mathcal{A}$. $a \ \ltimes \ b$ means that the argument a attacks the argument b. A set of arguments S attacks an argument b iff b is attacked by an argument in S.

One argument is acceptable if every attack on it is attacked by an accepted argument. This notion produces the following definitions:

Definition 2.4 (Acceptable argument (Dung, 1995))**.** An argument $a \in \mathcal{A}$ is acceptable with respect to a set of arguments S (noted $acceptable(a, S)$), iff ($\forall x \in \mathcal{A} \wedge x \ltimes a) \longrightarrow (\exists y \in S \wedge y \ltimes x$).

An argument is acceptable with respect to a set S of arguments if it is defended by that S against all its defeaters.

Definition 2.5 (Preferred extension (Dung, 1995))**.** A set of arguments S is conflict-free iff ($\forall a \in S \wedge \forall b \in S)(\neg a \ltimes b$). A conflict-free set of arguments S is admissible iff ($\forall a \in S) \longrightarrow acceptable(a, S)$ A set of arguments S is a preferred extension if it is a maximal (with respect to inclusion set) admissible set of $\mathcal{A}$.

A preferred extension represents a consistent position within AF, which can defend itself against all attacks and which cannot be further extended without introducing a conflict.

[1] In some classifications, semantic approaches consider model-theoretic semantics to determine whether or not a correspondence exists between two entities

In Dung's model, all arguments have equal strength and an attack always succeeds. (Amgoud and Cayrol, 1998) has introduced the notion of preference between arguments, where an argument can defend itself against weaker arguments. This model defines a global preference between arguments. In order to relate preferences to different audiences, (Bench-Capon, 2003) has proposed to associate arguments to the values which support them. Different audiences can have different preferences over these values. This leads to the notion of *successful attacks*, i.e., those which defeat the attacked argument, with respect to an ordering on the preferences that are associated with the arguments. It allows for accommodating different audiences with different interests and preferences:

Definition 2.6 (Value-based Argumentation Framework (VAF) (Bench-Capon, 2003))**.** A VAF is a 5-tuple $\langle \mathcal{A}, \ltimes, \mathcal{V}, v, \succeq \rangle$, where $\langle \mathcal{A}, \ltimes \rangle$ is a AF; $\mathcal{V}$ is a nonempty set of values; $v : \mathcal{A} \rightarrow \mathcal{V}$, $\succeq$ is the preference relation over $\mathcal{V}$ ($v_1 \succeq v_2$ means that v_1 is preferred over v_2).

Definition 2.7 (Successful attack)**.** An argument $a \in \mathcal{A}$ successfully attacks (or defeats) an argument $b \in \mathcal{A}$ for the audience α (noted $a \dagger_\alpha b$) iff $a \ltimes b \wedge \neg v(b) \succeq_\alpha v(a)$.

Definition 2.8 (Acceptable argument)**.** An argument $a \in \mathcal{A}$ is acceptable to the audience α with respect to set of arguments S (noted $acceptable_\alpha(a, S)$) iff ($\forall x \in A x \dagger_\alpha a$) $\longrightarrow$ ($\exists y \in S$) $\wedge y \dagger_\alpha x$.

The notion of audience introduces the following definition of preferred extension in the VAF:

Definition 2.9 (Preferred extension)**.** A set S of arguments is *conflict-free* for an audience α iff $\forall a, b \in S, \neg(a \ltimes b) \vee a \dagger_\alpha b$. A conflict-free set of arguments S is *admissible* for an audience α iff $\forall a \in S, acceptable_\alpha(a, S)$. A set of arguments S in the VAF is a *preferred extension* for an audience α iff it is a maximal admissible set (with respect to set inclusion) for α.

For determining the status of the VAF with respect to a value ordering promoted by distinct audiences, the arguments will have one of the three statuses (Bench-Capon, 2003):

- Some arguments will be in the preferred extension, irrespective of value ordering. Such arguments will either have no attacker or have their inclusion forced by the argumentation process. These will be *objectively acceptable*.
- Some arguments will be in the preferred extension for some ordering of values. These will be *subjectively acceptable*.
- Some arguments will not be included in the preferred extension whatever the ordering on values. Such arguments are *indefensible*.

3 Argumentation for Ontology Matching

Our approach exploits argumentation theory as a way for supporting the creation and exchange of arguments, followed by the reasoning on their acceptability, where arguments can be seen as positions that support or reject correspondences. The arguments interact following the notion of attack and are selected according to the notion of acceptability. We have proposed an argumentation framework (Trojahn, Quaresma, Vieira and Moraes, 2008) that redefines the notion of acceptability from VAF, taking into account the confidence of the correspondences. In this section, firstly, we introduce the definition of argument in the context of ontology matching and next we detail our argumentation framework.

3.1 Arguments on correspondences

An argument represents a position in favour or against a correspondence. According to the degree of confidence associated to a correspondence (Section 2.1), we can provide means to indicate that an argument is stronger or weaker than other arguments. The confidence is a measure of the trust in the fact that the correspondence is appropriate and is usually derived from the similarity assessment made during the ontology matching process, *e.g.* from an edit distance measure between labels or an overlap measure between instance sets.

We define an argument as follows:

Definition 3.1 (Argument)**.** An argument $a \in AF$ is a tuple $a = \langle c, v, s, h \rangle$, such that c is a correspondence $\langle e, e', r, n \rangle$; $v \in \mathcal{V}$ is the value of the argument (associated with the matching approach, as we will detail in the next section); s is the strength of the arguments (from n); and h is one of $\{+, -\}$ depending on whether the argument is that c does or does not hold.

Arguments interact based on the notion of attack, which has been initially defined by (Laera, Tamma, Euzenat, Bench-Capon and Payne, 2006a):

Definition 3.2 (Attack)**.** An argument $\langle c, v, s, h \rangle \in \mathcal{A}$ attacks an argument $\langle c', v', s', h' \rangle \in \mathcal{A}$ iff $c = c'$ and $h \neq h'$.

For instance, if $a = \langle c, v_1, 1.0, + \rangle$ and $b = \langle c, v_2, 0.8, - \rangle$, $a \ltimes b$ and vice-versa (b is the counter-argument of a, and a is the counter-argument of b).

However, (Laera et al., 2006a) do not consider the strength of the arguments as an element in the arguments, as we present in the next section.

3.2 Argumentation model

We have extended the VAF model in order to consider the confidence of the correspondences in the notion of *successful attack*. The values in $\mathcal{V}$ correspond to the different matching approaches and each matcher has a preference ordering $\succeq$ over $\mathcal{V}$ such that its preferred values are those it associates to its arguments. We then aggregate both confidence of arguments and matcher preferences in our model.

We define SVAF as follows:

Definition 3.3 (Strength-based argumentation framework (SVAF) (Trojahn, Quaresma, Vieira and Moraes, 2008))**.** A SVAF is a sextuple $\langle \mathcal{A}, \ltimes, \mathcal{V}, v, \succeq, s \rangle$ such that $\langle \mathcal{A}, \ltimes \rangle$ is an AF, $\mathcal{V}$ is a nonempty set of values, $v : \mathcal{A} \rightarrow \mathcal{V}$, $\succeq$ is the preference relation over $\mathcal{V}$ ($v_1 \succeq v_2$ means that, in this framework, v_1 is preferred over v_2), and $s : \mathcal{A} \rightarrow [0, 1]$ represents a function that maps strengths to arguments.

Each audience α is associated with its own argumentation framework in which only the preference relation $\succeq_\alpha$ differs. In order to accommodate the notion of *strength*, the notion of *successful attack* has been extended:

Definition 3.4 (Successful attack (Trojahn, Quaresma, Vieira and Moraes, 2008))**.** An argument $a \in \mathcal{A}$ *successfully attacks* (or *defeats*, noted $a \dagger_\alpha b$) an argument $b \in \mathcal{A}$ for an audience α iff

$$a \ltimes b \wedge (s(a) > s(b) \vee (s(a) = s(b) \wedge \neg v(b) \succeq_\alpha v(a)))$$

An AF can be represented as a directed graph where the arguments are vertices and edges represent the attacks between arguments. The plurality of preferred extensions is associated with cycles in the graph. An AF (here VAF or SVAF), where cycles contain arguments with different associated values (polychromatic cycle), has a unique, non-empty preferred extension, given an ordering on values (Bench-Capon, 2003). Each value ordering represents an audience. In our model, each argument in a cycle has a different value and the algorithm from (Bench-Capon and Dunne, 2002) is used to compute the unique preferred extension for each audience.

From a preferred extension (as defined above), we define the alignment associated with it:

Definition 3.5 (Alignment associated with an extension)**.** Given an extension S in a SVAF, the alignment associated with this extensions is:

$$A(S) = \{c; \exists\langle c, v, s, +\rangle \in S\}$$

In this paper, we assume that arguments being acceptable by the majority of matchers can be considered as the consensus on the alignments generated by them:

Definition 3.6. Considering n audiences, $\alpha_1, ..., \alpha_n$, an argument $a \in A(S)$ is in the *set of majority acceptable arguments* iff a appears in n' subsets $A(S)$, $A(S)_1, ..., A(S)_{n'}$, where $n' > n/2$.

Our approach is based on the idea of voting. According to (Coste-Marquis et al., 2007), this mechanism is inadequate to characterise the notion of acceptable arguments at a group level and can potentially result in a conflicting set. However, this inadequateness is only true when the different agents do not share the same initial set of arguments. In our setting, all matchers share their arguments with each others and have the same set of arguments at the start. Thus, voting can be applied for generating the final set.

Although, in the general case, the set of majority acceptable arguments may contain conflicts, this does not happen in the context of this work because there are only positive arguments in the individual sets, i.e., from each preferred extension, only arguments with $h = +$ are taken into account. Furthermore, ontologies do not have definitions of negative information and do not allow the inference of contradictions.

In addition, in the general situation this potential problem could be solved through the calculus of the preferred extension of the majority acceptable arguments.

Finally, the arguments selected by voting are ”theoretical” defensible because they appear in at least the majority of preferred extensions.

3.3 Arguing on correspondences

The way arguments are generated differs from different applications and scenarios. In other terms, arguments are generated following different strategies. For instance, in ontology alignment agreement for agent communication (Laera, Blacoe, Tamma, Payne, Euzenat and Bench-Capon, 2007), arguments for and against correspondences between two ontologies are generated. A correspondence from an alignment service is accepted by an agent if the justification for this correspondence (why the correspondence was found) corresponds to the highest agent's preference (i.e, an argument in favour to the correspondence is generated, with $h = +$, otherwise an argument against the correspondence is generated, $h = -$).

Contrary to the Laeras's strategy, the strategy we adopt in this paper, *negative arguments as failure*, relies on the assumption that matchers return complete results. Each possible pair of ontology entities which is not returned by the matcher is considered to be a risk, and a negative argument is generated ($h = -$).

In our specific scenario, different matchers argue with each other in order to obtain an agreement on their alignments. To do this, each matcher represents a different audience. The values

Table 1: Correspondences and arguments generated by m_l and m_s.

id	correspondence	argument	matcher
A	$c_{l,1} = \langle zoom_o, zoom_{o'}, \equiv, 1.0\rangle$	$\langle c_{l,1}, l, 1.0, +\rangle$	m_l
B	$c_{l,2} = \langle Battery_o, Battery_{o'}, \equiv, 1.0\rangle$	$\langle c_{l,2}, l, 1.0+\rangle$	m_l
C	$c_{l,3} = \langle MemoryCard_o Memory_{o'}, \equiv, 0.33\rangle$	$\langle c_{l,3}, l, 0.33, +\rangle$	m_l
D	$c_{l,4} = \langle brand_o, brandName_{o'}, \equiv, 0.22\rangle$	$\langle c_{l,4}, l, 0.22, +\rangle$	m_l
E	$c_{l,5} = \langle price_o, price_{o'}, \equiv, 1.0\rangle$	$\langle c_{l,5}, l, 1.0, +\rangle$	m_l
F	$c_{s,1} = \langle CameraPhoto_o, DigitalCamera_{o'}, \equiv, 1.0\rangle$	$\langle c_{s,1}, s, 1.0, +\rangle$	m_s
G	$c_{s,2} = \langle zoom_o, zoom_{o'}, \equiv, 1.0\rangle$	$\langle c_{s,2}, s, 1.0, +\rangle$	m_s
H	$c_{s,3} = \langle brand_o, brandName_{o'}, \equiv, 1.0\rangle$	$\langle c_{s,3}, s, 1.0, +\rangle$	m_s
I	$c_{s,4} = \langle resolution_o, pixels_{o'}, \equiv, 1.0\rangle$	$\langle c_{s,4}, s, 1.0, +\rangle$	m_s
J	$c_{s,5} = \langle price_o, price_{o'}, \equiv, 1.0\rangle$	$\langle c_{s,5}, s, 1.0, +\rangle$	m_s

Table 2: Counter-arguments (attacks) for the arguments in Table 1.

id	correspondence	counter-argument	matcher
L	$c_{l,6} = \langle CameraPhoto_o, DigitalCamera_{o'}, \equiv, 0.5\rangle$	$\langle c_{l,6}, l, 0.5, -\rangle$	m_l
M	$c_{l,7} = \langle resolution_o, pixels_{o'}, \equiv, 0.5\rangle$	$\langle c_{l,7}, l, 0.5, -\rangle$	m_l
N	$c_{s,6} = \langle Battery_o, Battery_{o'}, \equiv, 0.5\rangle$	$\langle c_{s,6}, s, 0.5, -\rangle$	m_s
O	$c_{s,7} = \langle MemoryCard_o, Memory_{o'}, \equiv, 0.5\rangle$	$\langle c_{s,7}, s, 0.5, -\rangle$	m_s

in $\mathcal{V}$ correspond to the different matching approaches and each matcher m has a preference ordering $\succeq_m$ over $\mathcal{V}$ such that its preferred values are those it associates to its arguments. For instance, consider $\mathcal{V}$= $\{l, s, w\}$, i.e., *lexical*, *structural* and *wordnet-based* approaches, respectively, and three matchers m_l, m_s and m_w, using such approaches. The matcher m_l could have as preference order $l \succeq_{m_l} s \succeq_{m_l} w$.

In order to illustrate the argumentation process, consider two matchers, m_l (lexical) and m_s (structural), trying to reach a consensus on the alignment between the ontologies in Figure 1. m_l uses an edit distance measure to compute the similarity between labels of concepts and properties of the ontologies, while m_s is based on the comparison of the direct super-classes of the classes or classes of properties. Table 1 shows the correspondences and arguments generated by each matcher. The matchers generate complete alignments, i.e., if a correspondence is not found, an argument with value of $h = -$ is generated. It includes correspondences that are not relevant to the task at hand. For the sake of brevity, we show only the arguments with $h = +$ and the corresponding counter-arguments (Table 2). We consider 0.5 as the confidence level c for negative arguments ($h = -$). Considering $\mathcal{V} = \{l, v\}$, m_l associates to its arguments the value l, while m_s generates arguments with value s. m_l has as preference ordering: $l \succ_{m_l} s$, while m_s has the preference: $s \succ_{m_s} l$.

Having their arguments $\mathcal{A}$, the matchers exchange them. m_l sends to m_s its set of arguments $\mathcal{A}_l$ and vice-versa. Next, based on the attack notion, each matcher m_i generates its attack relation $\ltimes_i$ and then instantiates its SVAF$_i$. The arguments A, D, E, G, H and J (Table 1) are acceptable in both SVAFs (they are not attacked by counter-arguments with $h = -$). F, I, and B ($h = +$) successfully attack their counter-arguments ($h = -$) L, M and N, respectively, because they have highest confidence in their correspondences. C ($h = +$) is successfully attacked by its counter-argument O.

The arguments in the preferred extension of both matchers m_l and m_s are: A, D, E, F, G, H, J, F, I, B and O. While $\langle resolution_o, pixels_{o'}, \equiv, 1.0\rangle$, $\langle Battery_o, Battery_{o'}, \equiv, 1.0\rangle$ and $\langle CameraPhoto_o, DigitalCamera_{o'}, \equiv, 1.0\rangle$ have been accepted, $\langle MemoryCard_o, Memory_{o'}, \equiv, 0.33\rangle$ has been discarded.

4 Evaluation

We have experimented the proposed notion of set of majority acceptable arguments on basis of our argumentation model and a group of twelve matchers. With this methodology we are able to compare our approach with existent top-level state of the art ontology matchers. The individual results will be used as baseline to evaluate if the proposed system is able to perform as good as the best individual ontology matchers.

4.1 Dataset and matchers

We evaluate our argumentation model on basis of the benchmark (bibliography domain) provided by the OAEI community[2]. Each test is based on one particular (reference) ontology, which contains 33 named classes, 24 object properties, 40 data properties, 56 named individuals and 20 anonymous individuals. Basically, the reference ontology is matched with different alterations of itself. The test cases involving such alterations are grouped into three categories: (a) *concept test* (tests 101, 102, 103 and 104); (b) *systematic* (tests 201 – 266); and (c) *real ontologies* (tests 301–304).

We use a group of matchers, which has participated of the OAEI Benchmark Track 2007[3]: ASMOV (Jean-Mary and Kabuka, 2007a), (Jean-Mary and Kabuka, 2007b), DSSim (Nagy, Vargas-Vera and Motta, 2007), Falcon (Qu, Hu and Cheng, 2006), (Hu and Qu, 2008), Lily (Wang and Xu, 2007), Ola (Euzenat and Valtchev, 2004), OntoDNA (Kiu and Lee, 2006), (Kiu and Lee, 2007), PriorPlus (Mao and Peng, 2007), RiMON (Tang, Liang, Li and Wang, 2004), Sambo (Lambrix and Tan, 2006), (Tan and Lambrix, 2007), SEMA (Spiliopoulos, Valarakos, Vouros and Karkaletsis, 2007), TaxoMap (Zargayouna, Safar and Reynaud, 2007) and XSOM (Curino, Orsi and Tanca, 2007a), (Curino, Orsi and Tanca, 2007b).

DSSim, OntoDNA, PriorPlus, TaxoMap, and XSOM are based on the use of ontology-level information, such as labels of classes and properties, and ontology hierarchy, while ASMOV, Falcon, Lily, Ola, RiMON, Sambo, SEMA use both ontology-level and data-level (instances) information. When considering the techniques used in the matching process, DSSim, PriorPlus, and XSOM are based on edit-distance similarity, where DSSim, and X-SOM combine the string-based approaches with the synonymous relations provided by a thesaurus (WordNet). Regarding the structural approaches, several heuristics are used, such as number of common descendants, or number of similar nodes in the path between the root and the element (PriorPlus).

Most of the systems execute different matchers in parallel and combine their results: simple weighted formula (PriorPlus), Dempster's rule of combination (DSSim), feed-forward neural network (XSOM), systems of equations (OLA), linear interpolation (RiMON), weighted sum (SAMBO) and experimental weighted (Lily). Falcon executes sequentially a TFIDF linguistic matcher that combines concepts and instances, and a graph-based matcher. Moreover, ASMOV iteratively combines several matchers using a single weighted sum. Instance-based matchers are commonly based on Naive-Bayes classifiers (RiMON), statistics (Falcon and SEMA) or probabilistic methods (SAMBO).

4.2 Argumentation frameworks and arguments

Each matcher has a SVAF and a private preference order, which is based on the f–measure ordering for all matchers (as detailed in Section 4.4). For that, we consider a scenario where

[2] Ontology Alignment Evaluation Initiative (OAEI: http://oaei.ontologymatching.org/

[3] in http://oaei.ontologymatching.org/2007/results/

available matchers performance is known in advance for specific tasks, assuming that web matchers, available for contract, advertise themselves with performance results.

The highest preferred value of each matcher is the value that it associates to its arguments[4]. For instance, ASMOV and Lily have as preference ordering:

- $v(ASMOV) \succeq_{ASMOV} v(Lily) \succeq_{ASMOV} v(RiMON) \succeq_{ASMOV} v(Falcon) \succeq_{ASMOV} v(Ola) \succeq_{ASMOV} v(PriorPlus) \succeq_{ASMOV} v(Sema) \succeq_{ASMOV} v(DSSim) \succeq_{ASMOV} v(XSom) \succeq_{ASMOV} v(Sambo) \succeq_{ASMOV} v(OntoDNA)$;
- $v(Lily) \succeq_{Lily} v(ASMOV) \succeq_{Lily} v(RiMON) \succeq_{Lily} v(Falcon) \succeq_{Lily} v(Ola) \succeq_{Lily} v(PriorPlus) \succeq_{Lily} v(Sema) \succeq_{Lily} v(DSSim) \succeq_{Lily} v(XSom) \succeq_{Lily} v(Sambo) \succeq_{ASMOV} v(OntoDNA)$.

For negative arguments ($h = -$), we use str=1.0 and str=0.5, assuming that matchers strongly reject correspondences that they do not find (it could be the case when the information about the matcher quality is not available), or slightly reject them, respectively.

The number of arguments generated by each matcher depends on the size of the ontologies (here, the number of concepts and properties). As matchers are supposed to generate complete alignments (arguments for and against correspondences), the number of arguments, n_m, for the matcher m is:

$$n_m = |o| \times |o'|$$

4.3 Evaluation measures

As quality measures, the classical precision, recall and f–measure are used. Such measures are derived from a contingency table (Table 3).

Table 3: Contingency table for binary classification.

	manual h = +	manual h = -
output h = +	m_{++}	m_{+-}
output h = -	m_{-+}	m_{--}

Precision (P) is defined by the number of correct automated correspondences (m_{++}) divided by the number of correspondences returned by the system (m_{++} + m_{+-}). It measures the system's correctness (for instance, we measure the correctness of the set of correspondences extracted from the set of majority acceptable arguments, with respect to a reference alignment). *Recall* (R) indicates the number of correct correspondences returned by the system divided by the number of manual correspondences (m_{++} + m_{-+}). It measures how complete or comprehensive the system is in its extraction of relevant correspondences. *F–measure* (F) is a harmonic mean of precision and recall.

$$P = \frac{m_{++}}{(m_{++} + m_{+-})}, \quad R = \frac{m_{++}}{(m_{++} + m_{-+})}, \quad F = \frac{(2 * P * R)}{(P + R)}$$

To measure the global performance of the system, *macro-averaging* and *micro-averaging* (Joachims, 2002) are applied. Such measures are often useful to compute the average performance of a system over multiple test sets, where the results of *n* binary tasks can be averaged to get a single performance value. *Macro-averaging* corresponds to the standard way of computing an (arithmetic) average. The performance (i.e. precision or recall) is computed

[4]Here, we use the name of the matcher to indicate the value it promotes. However, in a more comprehensive way, we could have values representing the different matching approaches (i.e., lexical, structural, etc.) and then different matchers using such approaches could promote such values.

separately for each of the *n* tests. The average is computed as the arithmetic mean of the performance measure over all tests. *Micro-averaging* averages the contingency tables of the various tests. For each cell of the table, the arithmetic mean is computed – m^{avg}_{++}, m^{avg}_{+-}, m^{avg}_{-+}, m^{avg}_{--} – and the performance is computed from this averaged contingency table. For the precision, *macro-averaging* and *micro-averaging* imply:

$$P^{macro} = \frac{1}{n}\sum_{i=1}^{n} P_i, \qquad P^{micro} = \frac{m^{avg}_{++}}{(m^{avg}_{++} + m^{avg}_{+-})}$$

Macro-averaging gives equal weight to each test whereas micro-averaging gives equal weight to each correspondence (example). For all comparative results, a significance test is applied, considering a confidence degree of 95%. Such comparison is indicated in bold face in the tables below.

4.4 Results and discussion

In this section we first present the individual results for each matcher and next we present the results of our argumentation model. We discuss the use of different values of strength for the arguments rejecting correspondences (*h* = -) and then the argumentation results are compared with the baseline – which is composed of the union of all individual matcher results – and individual matcher results. When evaluating our argumentation model itself, we consider three sets of alignments: the alignments in every preferred extensions for every matcher (here we call "objectively acceptable"), the alignments in the preferred extension of some matcher ("subjectively acceptable") and the arguments in the majority of preferred extensions ("majority"). Note that we take into account only the positive arguments (correspondences) in the preferred extensions (Section 3.2).

Firstly, table 4 shows the results for each matcher. These results out a group of systems, ASMOV, Lily, Falcon, OLA, PriorPlus and RiMOM which perform the tests at the highest level of quality (Euzenat, Isaac, Meilicke, Shvaiko, Stuckenschmidt, Sváb, Svátek, van Hage and Yatskevich, 2007). Of these, ASMOV, Lily and RiMOM have slightly better results than the three others.

Table 4: Individual matcher results

	ASMOV	DSSim	Falcon	Lily	Ola	OntoDNA
P^{macro}	0.93	0.97	0.93	0.94	0.88	0.54
P^{micro}	0.95	0.98	0.92	0.96	0.89	0.83
R^{macro}	0.84	0.64	0.81	0.85	0.81	0.42
R^{micro}	0.90	0.64	0.86	0.89	0.87	0.49
F^{macro}	*0.87*	0.71	0.84	*0.88*	0.83	0.54
F^{micro}	*0.92*	0.77	0.89	0.92	0.88	0.62

	PriorPlus	RiMON	Sambo	SEMA	TaxoMap	XSom
P^{macro}	0.89	0.95	0.89	0.89	0.93	0.72
P^{micro}	0.93	0.95	0.98	0.90	0.92	0.76
R^{macro}	0.79	0.83	0.55	0.72	0.27	0.66
R^{micro}	0.81	0.86	0.56	0.74	0.21	0.70
F^{macro}	0.82	*0.86*	0.69	0.76	0.58	0.73
F^{micro}	0.86	*0.91*	0.71	0.81	0.34	0.73

Second, we have used these matchers to evaluate our argumentation model. Our model aims at reaching a consensus between the matchers, improving or balancing the individual results. Table 5 shows the results of the three sets of alignments. Using $str = 0.5$ or $str = 1.0$ for the strength of arguments rejecting correspondences is a trade-off between precision and recall. We have the same behaviour for both strengths ($str = 0.5$ and $str = 1.0$), for all sets of selected arguments ("subjectively", "objectively" and "majority" sets).

Basically, the matchers produce arguments supporting correspondences (positive arguments) with confidence between 0.80 and 1.0. Considering $str = 0.5$ for arguments rejecting such correspondences, good values of recall are achieved, while precision is lower. It is due to the fact that the arguments with such strength do not represent attacks for the arguments representing true and false positive correspondences. As the matchers have good performance, this results in average better values of recall (the majority of the true positive correspondences are selected).

On the other hand, for $str = 1.0$, good values of precision are obtained, while recall is lower. It happens because the false positive correspondences and some true positive correspondences are successfully attacked. In such a way, the precision is better.

Looking for each set, they are more or less selective, depending on the selected strength. For 0.5 (Table 5), as expected, the sets of objectively and majority alignments are the most selective in terms of precision, while subjectively and majority sets are more selective in terms of recall.

Specifically, using 1.0 objectively sets have a low recall because one argument rejecting a correspondence can attack a number of arguments accepting the correspondence, while subjectively sets output low precision because they include the correspondences from the preferred extensions of all matchers. The majority sets represent a compromise between the two other sets of alignments, selecting the set of correspondences being defensible by the majority of matchers.

For 0.5, objectively sets improve the results because the lack of information of one matcher is not enough to attack the false positive correspondences. In this way, the behaviour of objectively sets are similar to the subjectively ones. Majority sets decrease their performance for the same reasons: lack of counter-arguments to attack false positive correspondences and then the majority of them are acceptable (similar behaviour with subjectively sets).

Table 5: Subjectively, objectively and majority sets – all cases.

	Subjectively		Objectively		Majority	
	0.5	1.0	0.5	1.0	0.5	1.0
P^{macro}	0.69	0.73	0.75	**1.0**	0.70	**0.98**
P^{micro}	0.67	0.74	0.73	**1.0**	0.67	**1.0**
R^{macro}	**0.89**	**0.87**	**0.92**	0.13	**0.89**	0.78
R^{micro}	**0.93**	**0.91**	**0.95**	0.13	**0.93**	0.83
F^{macro}	**0.77**	**0.79**	**0.82**	0.40	**0.77**	**0.84**
F^{micro}	0.78	0.81	**0.82**	0.22	0.78	**0.90**

Looking for the most interesting subset of the benchmark – real cases – majority sets outperform significantly subjectively and objectively sets in terms of f–measure (Table 6).

Table 6: Objectively, subjectively, majority – real cases.

	Subjectively		Objectively		Majority	
	0.5	1.0	0.5	1.0	0.5	1.0
P^{macro}	0.53	0.60	0.55	**1.0**	0.53	0.96
P^{micro}	0.53	0.61	0.55	**1.0**	0.54	**0.97**
R^{macro}	**0.84**	**0.84**	**0.86**	0.16	**0.84**	0.73
R^{micro}	**0.86**	**0.85**	**0.87**	0.18	**0.86**	0.76
F^{macro}	0.64	0.70	0.66	0.35	0.65	**0.82**
F^{micro}	0.66	0.71	0.68	0.30	0.66	**0.85**

Third, comparing argumentation and baseline (Tables 5 and 7, respectively), as expected, baseline has results closer to the less selective configuration of argumentation – subjectively and majority sets with $str = 0.5$ for arguments rejecting correspondences – specially in terms of recall. In terms of precision, all argumentation sets are able to filter out false positive correspondences, resulting in better values of precision than the baseline. We give special attention to precision because usually task operations, processes or problem-solving issues are based on ontologies, thus higher precision allow for systems to operate more on the safe side when interoperability is treated through ontology matching.

Table 7: Baseline and best matchers.

	Baseline	Best Matchers
P^{macro}	0.61	ASMOV (0.93), Lily (0.94), RiMON (0.95)
P^{micro}	0.57	ASMOV (0.95), Lily (0.96), RiMON (0.95)
R^{macro}	0.90	ASMOV (0.84), Lily (0.85), RiMON (0.83)
R^{micro}	0.94	ASMOV (0.90), Lily (0.89), RiMON (0.86)
F^{macro}	0.71	ASMOV (0.87), Lily (0.88), RiMON (0.86)
F^{micro}	0.71	ASMOV (0.92), Lily (0.92), RiMON (0.91)

Fourth, when comparing the results of each matcher (Table 4) with the results of argumentation (Table 5), specially for the best matchers (Table 7), the most selective sets (objectively and majority with $str = 1.0$) improve the precision of all matchers. The best f–measure for argumentation (majority with $str = 1.0$) is similar to the best matchers (ASMOV, Lily, and RiMON). In this way, consensus achieved by argumentation is a balancing between the individual results. By consensus there is not exactly an improvement of all individual results, but intermediary values near the best matchers and an improvement of the matchers with low performance.

Finally, regarding the subset of real cases, PriorPlus, Falcon, ASMOV, Lily, OntoDNA and XSOM are in the group of the best matchers in terms of f–measure: PriorPlus (0.84), Falcon (0.82), ASMOV (0.81), Lily (0.79), OntoDNA and XSOM (0.77). In terms of f–measure$_{micro}$ (Table 6) majority set ($str = 1$) slightly outperforms all the best matchers, while f–measure$_{macro}$ is similar to the best results. Moreover, best matchers vary depending on specific differences of multiple data sets. For instance, OntoDNA is the best matcher for the test 303, while it is the last one for the test 304. In such a way, argumentation approaches are an advantage.

5 Related Work

Most of the ontology negotiation-based systems aim at arriving to a common ontology which agents can use in their interactions, as proposed by (Bailin and Truszkowski, 2002), (Beun, van Eijk and Prust, 2004) and (van Diggelen, Beun, Dignum, van Eijk and Meyer, 2006). Differently from these proposals, (Tamma, Wooldridge, Blacoe and Dickinson, 2002) present an ontology that describes the basic concepts of a negotiation process. Such ontology is not the object being negotiated. While (Bailin and Truszkowski, 2002) propose a protocol where incremental interpretation, clarification and explanation steps are used to establish a common basis for communication between the agent, (Beun et al., 2004) define sets of discrepancies and feedbacks to solve ontological discrepancies and (van Diggelen et al., 2006) present a layered communication protocol, where each layer is able to solve a kind of ontological mismatch. In such systems, the ontologies that agents use are incrementally modified during the negotiation step.

On other perspective, (Silva, Maio and Rocha, 2005) describe a protocol for negotiating ontology correspondences, which is based on a utility function. Agents are able to achieve consensus about correspondence rules defined between two different ontologies. Each agent keep its ontologies unaltered. However, the correspondences are specified by services outside the negotiation framework, the system is highly dependent on the MAFRA framework and cannot be flexibly applied in other environments.

The use of argumentation in ontology matching was initially proposed by (Laera, Tamma, Euzenat, Bench-Capon and Payne, 2006b) and (Laera et al., 2007), where two agents argue on the correspondences provided by an ontology correspondence repository (OMR). The value-based argumentation framework (VAF) is used, where audiences represent different agent's preference ($Pref_1$ and $Pref_2$) and are chosen on the basis of the ontological information. The argumentation process takes four main steps: (i) for each agent, an argumentation framework is constructed, by specifying the set of arguments (according agents's preferences and thresholds); (ii) the individual frameworks are merged, by forming the union of the individual argument sets of the multiple agents, and then the attack relations are computed; (iii) for each VAF, the set of arguments that are undefeated by attacks is determined, given a value ordering – the global view is considered by taking the union of these preferred extensions for each audience; and (iv) the arguments in every preferred extension of every audience are considered – the correspondences that have only arguments supporting a correspondence are included in the a set called *agreed alignments*, the correspondences that have only arguments against are rejected, and the correspondences which are in some preferred extension of every audience are part of the set called *agreeable alignments*.

In (Morge and Routier, 2007), the authors propose an argumentation framework for agents to reach an agreement on their terminologies. The framework formalises a debate between the agents, where the conflicting representations (claims) are discussed. Agents exchanges claims with each other and argumentation is used for internal reasoning on the acceptability of that claims. The argumentation framework is build on description logics, where (conflicting) claims (concept definitions and assertions) can have different relevance according to different audiences. Arguments are formed by two elements, a claim (conclusion) and a set of claims (from which the conclusion can be inferred). Each audience is associated with a particular priority relation and individually evaluates the relevance of arguments (i.e., the most important value promoted by one claim in the premise). What the authors define as 'relevance' in their work is similar to what Bench-Capon defines as 'value'. Moreover, they share a similar definition of 'successful attack': an argument defeats another argument if they attack each other and the second argument is not more relevant than the first one. However, it is not clear how audiences and their preferences on arguments are defined and applied during the debate.

The closer proposal to ours is that of Laera and colleagues. However, we can outline many differences from our work and Laera's approach: (i) the way the arguments are generated differs (Section 3.3); (ii) in this paper, the correspondences are computed by several specialised matchers (the evaluation of Laera's model is carried out using only two agents that argue in the argumentation process), which argument to solve conflicts between their individual results; (iii) in Laera's model, if the agents select directly opposing preferences, it leads to an inability to reach agreement on many of the correspondences (specifically for tests 205 and 206 of benchmark, where the argumentation produces F-measure of 0). This situation is avoided in the proposed argumentation model, through relations with strength (S-VAF); (iv) different argumentation frameworks are used (VAF in Laera vs. S-VAF); (iv) Laeras's model considers *agreeable alignments* while we exploit different sets of acceptable arguments, namely subjectively, objectively and majority sets.

Regarding our previous work (Trojahn, Quaresma, Vieira and Moraes, 2008) and (Trojahn, Quaresma and Vieira, 2008a), we can outline the main differences and improvements in relation to this paper. First, different matchers were used in those works, i.e., representing different matcher approaches (namely syntactic, semantic, and structural). Second, the argumentation models considered only discrete values of strength (i.e., *certainty* and *uncertainty*) for arguments. Third, only objectively acceptable arguments have been considered to evaluate the quality of the correspondences generated by the argumentation process. Fourth, the evaluation was carried out using few examples of matching between ontologies[5] or subsets of OAEI benchmark. Furthermore, in this paper, we have significantly improved the evaluation of our

[5]Examples obtained from http://dit.unitn.it/ accord/Experimentaldesign.html

approach with more complete sets of alignments as well as we have introduced and evaluated the notion of majority acceptable arguments.

6 Final Remarks and Future Work

In this paper, the problem of combining different matching approaches was formalised using an argumentation model based on strengths. We have focused on the evaluation of the notion of majority acceptable arguments, where arguments in the majority of preferred extensions, from a group of matchers, are considered as consensus on the different alignments. Such evaluation has shown that such notion seems to be promising, specially when compared with the subsets being acceptable for every matcher or for some matcher. The notion of majority is based on the fact that the more often a correspondence is agreed on, the more chances for it to be valid, as outlined by (Euzenat, Mochol, Shvaiko, Stuckenschmidt, Sváb, Svátek, van Hage and Yatskevich, 2006). For combining argumentation frameworks that do not share the same set of arguments, this notion is not valuable. However, in our specific setting, all matchers have the same set of arguments, what makes the notion of majority acceptable.

In a more general setting, where the performances of matchers are known before hand, a complete preference order can be defined (i.e., $A > B > C$). When this information is not available, a partial preference order may be specified (i.e., $A > B$; and $A > C$). The possibility of adjusting preferences between matchers, what can be done in order to fit ontologies and scenarios, relates a potential advantage of our argumentation model.

Together with preference ordering, our model takes into account the strengths of arguments and the performance of argumentation is directly related with the strength associated to arguments rejecting a correspondence.

The importance of using arguments with strengths is recognised, reflecting the confidence a matcher has in the similarity between the two entities (the matching tools actually output correspondences with a confidence measure). Such confidence levels are usually derived from similarity assessments made during the matching process, e.g., from edit distance measures between labels, or overlap measures between instance sets. However, there is no objective theory nor even informal guidelines for determining such confidence levels. Using them to compare results from different matchers is therefore questionable especially because of potential scale mismatches. For example, a same strength of 0.8 may not correspond to the same level of confidence for two different matchers.

According to our evaluation, using different values to represent the strength of arguments rejecting a correspondence is a trade-off between precision and recall. Such evidence points to the necessity of a more comprehensive study in order to specify strengths that could balance the results.

A potential weakness of our argumentation model is associated to the fact that an argument against a correspondence can successfully attack all the arguments in favour of it, even if there are dozens of these, if the strength of counter-arguments for true positive correspondences is higher. For example, three arguments in favour of a true positive correspondence may be successfully attacked by one argument representing a false negative correspondence.

Another potential weakness of our model is related to the generation of complete alignments. At first sight it seems to be quite unrealistic, but it can nevertheless be supported by the observation that most matchers try to provide as much correspondences as possible. However, dealing with a large number of arguments can become costly. The number of arguments depends on the size of the ontologies.

Regarding the different sets of alignments, sets containing the majority of acceptable arguments perform, in average, better than the other sets (regarding F–measure values for both settings of strength). The majority sets represent a compromise between the two other kinds of sets, selecting the correspondences from the majority of matchers. The arguments selected by voting are ”theoretical” defensible because they appear in at least the majority of preferred extensions.

Finally, it is hard to improve the best matcher, specially when there is a great intersection in the individual sets. In this case, the performance of argumentation is similar to the individual matchers. On the other hand, as best matchers vary depending on specific differences of multiple data sets, argumentation approaches are an advantage.

As future work we plan to compare the results reported in this paper with the negotiation model based on voting we have proposed in previous work; introduce weighted strength based on the individual performance of matchers; and analyse the impact of using an instantiation of SVAF without complete preference order, considering situations where individual performances of matchers are not available for computing a complete order.

7 *Acknowledgement

We would like to thank Jérôme Euzenat for his suggestions on the argumentation formalism we are using in this paper.

References

Amgoud, L. and Cayrol, C. 1998. On the acceptability of arguments in preference-based argumentation, *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, San Francisco (California), pp. 1–7.

Amgoud, L. and Cayrol, C. 2002. A reasoning model based on the production of acceptable arguments, *Annals of Mathematics and Artificial Intelligence* **34**(1-3): 197–215.

Bailin, S. C. and Truszkowski, W. 2002. Ontology negotiation between intelligent information agents, *Knowledge Engineering Review* **17**(1): 7–19.

Bench-Capon, T. 2003. Persuasion in practical argument using value-based argumentation frameworks, *Journal of Logic and Computation* **13**(3): 429–448.

Bench-Capon, T. and Dunne, P. 2002. Value-based argumentation frameworks, *Technical report*, Department of Computer Science, University of Liverpool.

Beun, R.-J., van Eijk, R. M. and Prust, H. 2004. Ontological feedback in multiagent systems, *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems*, IEEE Computer Society, Washington DC (USA), pp. 110–117.

Coste-Marquis, S., Devred, C., Konieczny, S., Lagasquie-Schiex, M.-C. and Marquis, P. 2007. On the merging of dung’s argumentation systems, *Artificial Intelligence* **171**(10–15): 730–753.

Curino, C., Orsi, G. and Tanca, L. 2007a. X-SOM: A flexible ontology mapper, *Proceedings of the 18th International Conference on Database and Expert Systems Applications*, IEEE Computer Society, Washington DC (USA), pp. 424–428.

Curino, C., Orsi, G. and Tanca, L. 2007b. X-SOM results for OAEI 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 276–285.

Dung, P. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n–person games, *Artificial Intelligence* **77**(2): 321–357.

Euzenat, J., Isaac, A., Meilicke, C., Shvaiko, P., Stuckenschmidt, H., Sváb, O., Svátek, V., van Hage, W. R. and Yatskevich, M. 2007. Results of the ontology alignment evaluation initiative 2007, *in* P. Shvaiko, J. Euzenat, F. Giunchiglia and B. He (eds), *Proceedings of the 2nd International Workshop on Ontology Matching*, Vol. 304 of *CEUR Workshop Proceedings*, CEUR-WS.org.

Euzenat, J., Mochol, M., Shvaiko, P., Stuckenschmidt, H., Sváb, O., Svátek, V., van Hage, W. R. and Yatskevich, M. 2006. Results of the ontology alignment evaluation initiative 2006, *in* P. Shvaiko, J. Euzenat, N. F. Noy, H. Stuckenschmidt, V. R. Benjamins and M. Uschold (eds), *Proceedings of the 1st International Workshop on Ontology Matching*, Vol. 225 of *CEUR Workshop Proceedings*, CEUR-WS.org.

Euzenat, J. and Shvaiko, P. 2007. *Ontology Matching*, Springer-Verlag, Heidelberg (DE).

Euzenat, J. and Valtchev, P. 2004. Similarity-based ontology alignment in OWL-Lite, *Proceedings of the 16th European Conference on Artificial Intelligence*, Vol. 16, IOS Press, pp. 333–337.

Fensel, D. 2003. *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*, 2nd edn, Springer-Verlag, New York (USA).

Hu, W. and Qu, Y. 2008. Falcon–AO: A practical ontology matching system, *Journal of Web Semantics* **6**(3): 237–239.

Jean-Mary, Y. and Kabuka, M. 2007a. ASMOV: Ontology alignment with semantic validation, *Joint SWDB-ODBIS Workshop on Semantics, Ontologies, Databases*, Vienna (Austria), pp. 15–20.

Jean-Mary, Y. and Kabuka, M. 2007b. ASMOV results for OAEI 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 141–150.

Joachims, T. 2002. *Learning to Classify Text Using Support Vector Machines – Methods, Theory and Algorithms*, Vol. 668 of *The Springer International Series in Engineering and Computer Science*, Kluwer Academic Publishers/Springer.

Kiu, C.-C. and Lee, C.-S. 2006. Ontology mapping and merging through OntoDNA for learning object reusability, *Journal of Educational Technology and Society* **9**(3): 27–42.

Kiu, C.-C. and Lee, C.-S. 2007. OntoDNA: Ontology alignment results for OAEI 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 196–204.

Laera, L., Blacoe, I., Tamma, V., Payne, T., Euzenat, J. and Bench-Capon, T. 2007. Argumentation over ontology correspondences in MAS, *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, ACM, New York (USA), pp. 1–8.

Laera, L., Tamma, V., Euzenat, J., Bench-Capon, T. and Payne, T. 2006a. Reaching agreement over ontology alignments, *Proceedings of the 5th International Semantic Web Conference*, Vol. 4273/2006 of *Lecture Notes in Computer Science*, Springer, Berlin (DE), pp. 371–384.

Laera, L., Tamma, V., Euzenat, J., Bench-Capon, T. and Payne, T. 2006b. Reaching agreement over ontology alignments, *Proceedings of the 5th International Semantic Web Conference*, Vol. 4273/2006 of *Lecture Notes in Computer Science*, Springer, Berlin (DE), pp. 371–384.

Lambrix, P. and Tan, H. 2006. SAMBO: A system for aligning and merging biomedical ontologies, *Journal of Web Semantics* **4**(3): 196–206.

Mao, M. and Peng, Y. 2007. The PRIOR+: Results for OAEI campaign 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 219–226.

Morge, M. and Routier, J.-C. 2007. Debating over heterogeneous descriptions, *Applied Ontology, Special issue on Formal Ontology for Communicating Agents* **2**(3–4): 333–349.

Nagy, M., Vargas-Vera, M. and Motta, E. 2007. DSSim – managing uncertainty on the semantic web, *Proceedings of the 2nd International Workshop on Ontology Matching*, Busan (Korea).

Qu, Y., Hu, W. and Cheng, G. 2006. Constructing virtual documents for ontology matching, *Proceedings of the 15th International Conference on World Wide Web*, ACM, New York (USA), pp. 23–31.

Silva, N., Maio, P. and Rocha, J. 2005. An approach to ontology mapping negotiation, *Proceedings of the 3rd International Conference on Knowledge Capture Workshop on Integrating Ontologies*, ACM, Banf (Canada), pp. 1–7.

Spiliopoulos, V., Valarakos, A., Vouros, G. and Karkaletsis, V. 2007. SEMA: Results for the ontology alignment contest OAEI 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 244–254.

Tamma, V., Wooldridge, M., Blacoe, I. and Dickinson, I. 2002. An ontology based approach to automated negotiation, *Revised Papers from the Workshop on Agent Mediated Electronic Commerce on Agent-Mediated Electronic Commerce IV, Designing Mechanisms and Systems*, Springer-Verlag, London (UK), pp. 219–237.

Tan, H. and Lambrix, P. 2007. SAMBO results for the ontology alignment evaluation initiative 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 236–243.

Tang, J., Liang, B., Li, J.-Z. and Wang, K. 2004. Risk minimization based ontology mapping, *Advanced Workshop on Content Computing*, Vol. 3309 of *Lecture Notes in Computer Science*, Springer Verlag, pp. 469–480.

Trojahn, C., Quaresma, P. and Vieira, R. 2008a. An argumentation framework based on confidence degrees to combine ontology mapping approaches, *International Journal of Metadata, Semantics and Ontologies* **2**(3): 142–150.

Trojahn, C., Quaresma, P. and Vieira, R. 2008b. An argumentation framework based on strength for ontology mapping, *Proceedings of the 5th International Workshop on Argumentation in Multi-Agent Systems*, Vol. 4, Springer-Verlag, Berlin (DE), pp. 57–71.

Trojahn, C., Quaresma, P., Vieira, R. and Moraes, M. 2008. A cooperative approach for composite ontology mapping, *LNCS Journal on Data Semantic X (JoDS)* **4900**(0): 237–263.

van Diggelen, J., Beun, R.-J., Dignum, F., van Eijk, R. M. and Meyer, J.-J. 2006. ANEMONE: An effective minimal ontology negotiation environment, *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, ACM, New York (USA), pp. 899–906.

Wang, P. and Xu, B. 2007. LILY: The results for the ontology alignment contest OAEI 2007, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 179–185.

Zargayouna, H., Safar, B. and Reynaud, C. 2007. TaxoMap in the OAEI 2007 alignment contest, *Proceedings of the 2nd International Ontology Matching Workshop*, Busan (Korea), pp. 268–275.