

Improved Relaxation-based Ontology Matching Negotiation

Alda Canito
School of Engineering
Polytechnic of Porto
Rua Dr. António Bernardino de
Almeida, 431
4200-072 Porto
alrfc@isep.ipp.pt

Paulo Maio
School of Engineering
Polytechnic of Porto
Rua Dr. António Bernardino de
Almeida, 431
4200-072 Porto
pam@isep.ipp.pt

Nuno Silva
School of Engineering
Polytechnic of Porto
Rua Dr. António Bernardino de
Almeida, 431
4200-072 Porto
nps@isep.ipp.pt

ABSTRACT

Agent-based Ontology Alignment Negotiation processes aims to generate an alignment through the interaction of two or more agents. When these agents exploit different matching services they can reach incompatible alignments, giving rise to conflicts hence the need to engage in a negotiation to achieve consensus. This consensus is sounder as more conflicts are correctly solved. In this paper, a Relaxation-based approach for Agent-based Ontology Matching Negotiation is explored and compared with the MbA/FDO argument-based approach. The limitations of the original Relaxation approach are addressed and modifications are proposed that increase the circumstances under which conflict resolution may occur, thus generating sounder alignments.

Categories and Subject Descriptors

D.2.12 Interoperability – *Data mapping*

I.2.11 Distributed Artificial Intelligence – *Intelligent agents, Multi-agent systems*

General Terms

Algorithms, Documentation, Experimentation, Verification.

Keywords

Ontology Matching, Negotiation, Relaxation, Argumentation, Agents, Multi-Agent Systems

1. INTRODUCTION

Over the last few years, agents have been introduced as a way to overcome the difficulties associated with managing and sharing the increasing amount of data available in information systems, while ontologies have been used to model this data in a semantically rich way. As the popularity of the Semantic Web increases and more data is being shared in the form of different ontologies, the integration problem escalates. In such a world, it is not reasonable to expect that two agents, interacting on behalf of users and willing to communicate, will be using the same ontology to describe their universe of discourse. Agents may be required to interact without previous knowledge of the ontologies the others are using. In this scenario, it is necessary that agents

can reconcile their ontologies in real-time, in a process usually referred to as Ontology Matching [1].

The ontology matching process is usually available as a service provided to the business agents so it can be requested when an alignment is necessary. However, considering that each agent has its own needs and goals and the subjective nature of ontologies, agents may have different preferences concerning the matching process; they can also exploit the matching services they find more convenient [1]. Different matchers can produce different and even contradictory candidate alignments, giving rise to conflicts between the agents. These conflicts must be addressed and tentatively resolved in some negotiation process, such that the agents may reach an agreement concerning each and every correspondence they will use in the conversation. Conflict resolution can be seen as an important metric in the negotiation process: by taking it into account, it is possible to argue whether an agreed alignment is more or less sounder than other. An alignment with a low number of unresolved conflicts can be taken with more confidence than the same alignment with more remaining conflicts. The process of reaching an agreement is commonly addressed as Ontology Matching Negotiation (OMN) and two approaches have been addressed in literature: (i) argument-based (e.g. [2][3]) and (ii) relaxation-based [4] [5].

This paper proposes improvements to the previously proposed approach by suggesting specific relaxation functions and providing more conflict resolution scenarios (section Proposal). It also demonstrates that agents benefit from engaging in the new relaxation-based approach, which resolves more conflicts and improves the alignment reached by the agents. For that, it is compared with an argument-based approach and with alignments generated by a set of matchers provided by the Ontology Alignment Evaluation Initiative [6] (section Experiments).

1.1 Ontology Matching Overview

Ontology Matching is seen as the process of discovering, (semi-) automatically, the correspondences between semantically related entities of two different but overlapping ontologies. Thus, as stated in [1], the matching process is formally defined as a function $f: (O_s, O_t, p, res, A) \rightarrow A'$ which, from a pair of ontologies to match – O_s and O_t – a set of parameters p , a set of oracles and resources res and an input alignment A , it returns an alignment A' between the matched ontologies. Ontologies O_s and O_t are often denominated as source and target ontologies respectively. An alignment is a set of correspondences expressed according to:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

iiWAS2013, 2-4 December, 2013, Vienna, Austria.

Copyright 2013 ACM 978-1-4503-2113-6/13/12 ...\$15.00.

- Two entity languages Q_{L_1} and Q_{L_2} associated with the ontologies languages L_1 and L_2 of matching ontologies (respectively) defining the matchable entities (e.g. classes, object properties, data properties, individuals);
- A set of relations R that is used to express the relation held between the entities (e.g. equivalence, subsumption, disjoint, concatenation, split);
- A confidence structure φ that is used to assign a degree of confidence in a correspondence. It has a greatest element \top and a smallest element \perp . The most common structure are the real numbers in the interval $[0-1]$, where 0 represents the lowest confidence and 1 represents the highest confidence.

Hence, a correspondence (or a match) is a 4-tuple $c = (s, t, r, conf)$ where $s \in Q_{L_1}(O_s)$ and $t \in Q_{L_2}(O_t)$ are the entities between which a relation $r \in R$ is asserted and $conf \in \varphi$ is the degree of confidence in the correspondence.

Over the recent years, research initiatives in ontology matching have developed many systems that rely on the combination of several basic algorithms yielding different and complementary competencies to achieve better results. A basic algorithm generates correspondences based on a single matching criterion [7]. These algorithms can be multiple classified as proposed in [1] [8] (e.g. terminological, structural, semantic). Yet, systems make use of a variety of functions such as: (i) aggregation functions whose purpose is to aggregate the values of the same match generated by multiple matchers into a single value (e.g. min, max, linear average); (ii) Alignment Extraction functions whose purpose is to select from a set of correspondences those that will be part of the resulting alignment. The selection method may rely on the simplest methods such as the ones based on threshold-values or more complex methods based on, for example, local and global optimizations (e.g.[9] [10]).

The selection of the most suitable algorithms/system is still an open issue as they should not be chosen exclusively with respect to the given data but also adapted to the problem that is to be solved [1]. However, this question has already been dealt with in [11] [12] [13]. Despite all the existing (conceptual and practical) differences between matching systems and algorithms, we will refer to all as matchers as all of them have a set of (candidate) correspondences as output.

2. STATE OF THE ART

This section describes and points out the limitations of both relaxation-based approach and argument-based approaches.

2.1 Relaxation Based approaches

The Relaxation-based approach presented in literature [4] [5] suggests that each of the negotiating parties generates an alignment between O_s and O_t . Each party assigns a confidence value c_n to each match through the use of a utility function (u), normally in the range $[0-1]$. This value is used to classify the match as one of “rejected”, “negotiable”, “proposed” or “mandatory” classes, which are defined by a multi-threshold approach:

- Mandatory Threshold (t_m), such that $t_m < c_n < 1$, determines that the agent is so confident about the match’s relevance such that it is fundamental that the match is accepted by the other agent.
- Proposition Threshold (t_p), such that $t_p < c_n < t_m$, above which the confidence in the match is enough for it to be

proposed to the other agent, but not such that an agreement cannot be reached without it.

- Negotiation Threshold, (t_n), such that $t_n < c_n < t_p$, above which the match is considered negotiable, meaning that the agent is not confident enough to propose the match to the other agent, but is willing to revise/relax its confidence if prompted.
- Rejection Threshold (t_r), above which match is considered rejected and below which eliminated.

The confidence value of a correspondence (c_n) can be updated through the use of a Meta-Utility function (U), allowing the re-categorization of matches from one category to another. This function is responsible for (i) identifying the parameter variation possibilities, (ii) the priorities over parameter variation and (iii) the conditions under which the variation may occur.

The negotiation process unfolds in two main phases: (i) Mandatory Correspondences Processing phase and (ii) Proposed Correspondences Processing phase. During the former phase, each agent shows the other the matches it considers mandatory. If no agreement is achieved in this phase, the negotiation fails and no alignment is generated. Otherwise, it proceeds to the Proposed Correspondences Processing phase.

In this phase, each agent shows their proposed matches to the other agent. Three situations may occur: (i) the match is also proposed by the other agent and is therefore accepted, (ii) the match is considered as “negotiable” by the other agent. In this case, the “negotiable” match is re-evaluated with the Meta-Utility function and it may either be accepted or remain as a conflict, and (iii) the match is considered “not negotiable” by the other agent and is consequently rejected.

2.1.1 Limitations

Despite the simplicity, this approach has some limitations. It is considered that a match’s confidence value can only be increased. This means that existing disagreements can only be resolved via including the matches in the alignment, and there is no scenario where the parties can agree to exclude a match after it being deemed at least as “negotiable”. When the party holding the “negotiable” position will not relax its match’s confidence value to “proposed”, no further negotiation concerning it occurs, resulting in an unresolved conflict. It gets worse when the disagreement is between “proposed” and “not negotiable” stances. In this case, the disagreement is so pronounced that the protocol does not even consider it worth of relaxation efforts. From this follows that the match is automatically excluded from the alignment, and the “not negotiable” stance automatically wins the conflict, while the “proposed” stance has no say on the matter. The proposal also provides no mechanism to calculate how much the agents are winning/losing under these conditions.

Finally, this approach considers scenarios where human intervention is required, namely when both agents consider a match as “negotiable”. This can be seen as a limitation when compared to some argument-based approaches, which are generally able to run in fully automatic mode [2] [3][14].

2.2 The MbA/FDO argumentation approach

Meaning-based Argumentation (MbA) [15], further improved in [16] into a more flexible approach (FDO), adopts the Value Argumentation Framework (VAF) [2]. Agents can express their matching preferences according to the classification of the matching algorithms:

- Terminological (T): comparing the names, labels and comments related to ontological entities;
- Internal Structural (IS): exploiting the internal features of entities (such as domain and range of their properties, cardinalities of attributes);
- External Structural (ES): exploiting the external relations between the entities in the ontology (such as super-entity, sub-entity or sibling);
- Semantic (S): using theoretical models to determine if there is a match between two entities;
- Extensional (E): comparing the set of instances under evaluation.

Arguments are represented as 3-uples, $t = \{G, c, pos\}$ where c is a match, pos is one of $\{-, +\}$, depending on the agent's belief that the correspondence does or does not hold, and G is the grounds justifying pos . A match is accepted if all agents participating in the negotiation have a positive opinion about it.

2.2.1 Limitations

MbA/FDO has a series of relevant limitations in comparison with the Relaxation-based approach, namely:

- Symmetric attacks. An argument (a) can only be attacked by its negation ($\neg a$), or counter-argument;
- Only rebuttal arguments are explored, i.e. agents must reject the entire argument and not the individual premises. Since the agents cannot argue about the reasons that lead them to a specific opinion, they cannot be argued into changing their stance;
- All agents use the same ontology matching repository. This means that, apart from preferences and thresholds that are unique to each agent, all the agents have the same perception about the set of matches. As a consequence, the outcome of the negotiation process corresponds to the intersection of the alignments proposed by the agents.

While there are more limitations addressed in literature [17], these are the most relevant when in comparison to the previously presented Relaxation-based approach. These show that is very unlikely for agents to be able to revise their initial stances, which is the opposite of what the core point of the Relaxation-based approach.

3. PROPOSAL

The state of the art relaxation approach [4] [5] is improved in order to overcome the identified limitations. Such improvements rely on the premise of increasing the number of scenarios where conflict resolution can occur. For that, agents must obtain an agreement not only concerning the correspondences to include in the final alignment, but also concerning those to remove.

To enable the exclusion of matches, it is necessary to consider both: (i) the possibility of relaxing a "negotiable" correspondence to "proposed" (effort for including) and (ii) a "proposed" correspondence to "negotiable" (effort for rejecting). Thus, two relaxing ways are now possible: (i) rising the match's confidence value of matches deemed "negotiable" and (ii) lowering the confidence value in order to exclude the match from the negotiation. This action can be done over matches deemed "proposed".

Unlike the original relaxation approach, when a match is deemed "negotiable" by both parties it will be rejected. This stems from the observation that none of the agents was interested in proposing

the match. This is devised so user intervention is not required for the completion of the negotiation process, making it completely automatic.

The match's initial category can be reevaluated with the use of a Meta-Utility function.

3.1 Meta-Utility Function

The Meta-Utility function is devised as:

$$relax(c, dir) = \begin{cases} true: & \begin{cases} profit(c) - eff(c, dir) \geq 0 & : dir = 1 \\ -profit(c) - eff(c, dir) \geq -1 & : dir = 0 \end{cases} \\ false: & otherwise \end{cases}$$

where c is a match, $profit(c)$ is the function that yields the agent's gain with the match, $eff(c, dir)$ is the function that yields the effort associated with changing the match's current confidence value, dir is one of $[0,1]$, stating whether the agent is trying to obtain a higher (1) or a lower confidence value (0), and $relax$ is one of $\{true, false\}$, stating whether it is possible or not to relax to the desired value.

The function $profit$ returns a value in the range $[0-1]$ and is given by one of the Gain Functions (cf. below). When relaxing for exclusion the agent is actually losing the value associated with the match, thus resulting in a negative profit.

The function eff follows the notion that the relaxation attempts do not come for free, meaning that agents will not always be willing to change their initial positions. It is important to compute how much effort the relaxation entails and how much the agent is willing to relax. For that, the following formula is proposed:

$$eff(c, dir) = \left| (1 + (step(dir) - c_n))^{effp} - 1 \right|$$

where c_n is the confidence value ($[0-1]$) of c , $step(dir)$ is the function giving the minimum value the match should have if changing category in the intended direction (given by dir). It corresponds to one of the thresholds previously introduced, and $effp$ is the power applied to determine how fast the effort grows with the increasing distance between $step(dir)$ and c_n .

3.2 Gain Functions

Concerning the computation of the profit associated with each match, two different functions are proposed: (i) Ontological Type and (ii) Ontology Usage.

The Ontological Type function allows an agent to assign different gain values to different kinds of matches depending on the type of the related entities. As a result, for instance, an agent can decide if a match between two Object Properties is more or less valuable than one between two Classes.

The Ontology Usage function pertains to relate the value of an entity presented in an ontology with the number of times it is referred in the ontology. Having a $triple = \{S, P, O\}$, the number of references corresponds to the number of times an entity appears as either S, P, or O, both in the ABox and TBox. In the context of this function, it is considered that an entity is more relevant to the ontology as more times it is used. Relevance is given through the comparison of how often the entity is referred with the most mentioned entity. This is computed as follows:

$$profit(c) = \frac{ref(e) + ref(e')}{2 \cdot somf + tomf}$$

where $ref(e)$ yields the number of references of a certain entity (e) in an ontology, $somf$ is the number of references of the most referred entity in the source ontology and corresponds to $somf = Max(ref(e)) : e \in O_s$, and $tomf$ is the number of references of

the most referred entity in the target ontology and corresponds to $tomf = \text{Max}(\text{ref}(e')) : e' \in O_t$.

4. EXPERIMENTS

In the described experiments we compare the relaxation-based approach results with those obtained with the MbA/FDO method. The latter results were obtained using an Extensible Argumentation Model [17] which is able to mimic the MbA/FDO process. With these experiments, we want to assess not only the benefits in terms of conflict resolution but also the consequences in terms of the accuracy of the alignments.

4.1 Setup

Concerning the dataset, several ontologies of overlapping domains were used, taken from the OAEI 2011 Conference Track repository, along with 21 reference alignments. These are grouped and treated as one big alignment, with 305 correspondences, as suggested by the provided gold standard alignments.

The matches are generated using the GECAD Ontology Alignment System (GOALS) [13] and the experiments were run between three agents, in two different pairs – A&B and A&C – in order to acknowledge if the same results could be attained in both scenarios. The agents' setup is:

Table 1 - Relaxation Thresholds

Agent	tr	tn	tp	tm
Agent A	0.60	0.85	0.93	1.0
Agent B	0.60	0.85	0.99	1.0
Agent C	0.60	0.85	0.95	1.0

Table 2 - Effort and Gain calculation parameters

Agent	effp	Function
Agent A	8	Ontology Usage
Agent B	8	Ontological Type
Agent C	8	Ontology Usage

Agent B is using the Ontological Type function, with the following parameters: matches between two Classes or two Properties have both a profit of 0.20. All other combinations were given a profit of 0.15.

4.2 Results

In order to know if agents are taking benefit from the negotiation process, it is relevant to know the accuracy of the alignment they have initially proposed and compare it to the accuracy of the final agreement. Figure 1 compares the agents' initial proposals with the alignment they reach after consensus.

Agents clearly benefit from engaging in a relaxation-based negotiation process since the agreement's accuracy is higher than their initial proposals', both in terms of precision and f-measure. Recall is shown to decrease slightly since some matches are known only to one of the parties and in these cases, they are not considered and are excluded from the negotiation process and thus from the reached alignment.

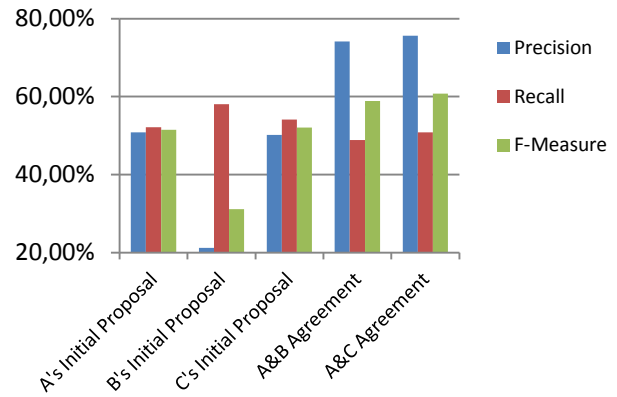


Figure 1 - Comparing Initial Proposal's and Agreement's alignment accuracies

4.2.1 Alignment Accuracy

Agents generating benefit after the engagement in a relaxation-based approach on itself is not enough to prove that this method produces good results. For that, the relaxation approach will be compared to the argument-based approach MbA/FDO and the average results of the OAEI 2011 participants for the same dataset.

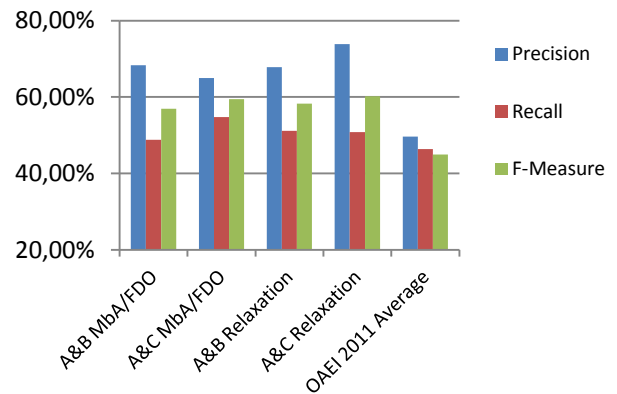


Figure 2 - Comparing MbA/FDO, Relaxation and the OAEI 2011 Average results

Relaxation-based results are shown to be better than the average results of the OAEI 2011 participants for this dataset and to those of the MbA/FDO approach, even if only for a very small margin (in the best scenario, the increase was only slightly above 1% in F-Measure).

4.2.2 Conflict Resolution

We consider that the more conflicts are correctly resolved the better and more significant the agreement is. The ideal scenario would be having all conflicts correctly resolved, i.e. when agents agree about including all the correct matches under negotiation and excluding all the incorrect ones. If some conflicts are wrongly resolved, they will impact the agreement's alignment by decreasing its F-Measure.

Table 3 below shows the percentage of resolved conflicts and how many of the resolved conflicts were correct.

Table 3 - Resolved conflicts and their accuracy

Agents	# Matches		% Conflicts	
	Initial	Remain	Resolved	Correctly
Relaxation				
A vs. B	748	35	95,32%	94,53%
A vs. C	234	44	81,20%	92,63%
MbA/FDO				
A vs. B	1319	1319	0,00%	0,00%
A vs. C	493	493	0,00%	0,00%

According to the MbA/FDO's limitations previously identified (cf. section 2.1.1), the alignment generated by this approach corresponds to the intersection of the agents' initial proposals and therefore all conflicts remain unresolved. As for the relaxation-based approach, although Table 3 shows that not all conflicts are resolved, it is evident that the number of resolutions is very high and the percentage of those which are correct is very high as well.

Summarily, the relaxation-based approach generates better alignments than the MbA/FDO approach even when not resolving all its conflict correctly.

5. CONCLUSIONS

Two approaches for the Ontology Matching Negotiation problem have been addressed in the literature: relaxation-based and argument-based. This paper focuses firstly on the relaxation-based proposal, addressing its limitations and providing improvements in order to increase the resolution of conflicts that arise during the negotiation process. For that, it proposes an improvement on the relaxation approach analyzed from literature, such that the agents can negotiate not only about the matches they want to include in their alignment, but also about the matches they wish to exclude. This provides more conflict resolution opportunities, consequently allowing agents to reach a sounder alignment, while still improving on their initial proposal's intersection.

Results not only show that agents benefit from engaging in a relaxation-based negotiation process, but also that the generated (agreed) alignments' accuracy is not distant from that achieved through an argument-based approach (i.e. MbA/FDO) and thus proving that both are similarly capable of achieving an agreement. The relaxation approach, however, has the advantage of achieving that while solving most of the conflicts, therefore adding more certainty to its results.

Since the relaxation approach was modified in order to be completely automatic, it is relevant to compare the alignments it generates with those generated by other automatic matchers. The results were thus compared to the average of the OAEI 2011 participants, showing a clear advantage for the relaxation approach in all the considered measures.

6. ACKNOWLEDGMENT

This work is supported by FEDER Funds through the "Programa Operacional Factores de Competitividade - COMPETE" program and by National Funds through FCT "Fundação para a Ciência e Tecnologia" under the projects: FCOMP-01-0124-FEDER-PEst-OE/EEI/UI0760/2011;PTDC/EIA-EIA/104752/2008.

7. REFERENCES

- [1] J. Euzenat and P. Shvaiko, "Ontology Matching," Springer-Verlag, Heidelberg, Germany, 2007.
- [2] T. Bench-Capon, "Persuasion in Pratical Argument Using Value-Based Argumentation Frameworks," *Journal of Logic and Computation*, vol. 13, pp. 429-448, 2003.
- [3] C. Trojahn, M. Moraes, P. Quaresma and R. Viera, "A Negotiation Model for Ontology Mapping," in *Conference on Intelligent Agent Technology (IAT'06)*, 2006.
- [4] P. Maio, N. Bettencourt, N. Silva and J. Rocha, "Ontology Mapping Negotiation Based on Categorization of Semantic Bridges," in *ICKEDS'06*, Lisbon, Portugal, 2006.
- [5] N. Silva, P. Maio and J. Rocha, "An Approach to Ontology Mapping Negotiation," in *Workshop on Integrating Ontologies on the Third International Conference on Knowledge Capture*, Banff, Canada, 2005.
- [6] "Ontology Alignment Evaluation Initiative," [Online]. Available: <http://oaei.ontologymatching.org/>.
- [7] E. Rahm and P. Bernstein, "A survey of approaches to automatic schema matching," *The VLDB Journal*, no. 10, pp. 334-350, 2001.
- [8] P. Shvaiko and J. Euzenat, "A survey on schema-based matching approaches," *Journal on Data Semantics*, vol. IV, pp. 146-171, 2005.
- [9] D. Gale and L. Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, no. 69, pp. 5-15, 1962.
- [10] J. Munkres, "Algorithms for the Assignment and Transportation Problems," *Journal of the Society for Industrial and Applied Mathematics*, no. 5, pp. 32-38, 1957.
- [11] D. Ngo, Z. Bellahsene, R. Colleta and others, "A flexible system for ontology matching," 2011.
- [12] K. Saruladha, G. Aghila and B. Sathiya, "A Comparative Analysis of Ontology and Schema Matching Systems," *International Journal of Computer Applications*, no. 34, pp. 14-21, 2011.
- [13] P. Maio and N. Silva, "GOALS - A test-bed for Ontology Matching," in *Knowledge Engineering and Ontology Development (KEOD)*, Funchal, Portugal, 2009.
- [14] L. Laera, V. Tamma, J. Euzenat and T. Payne, "Reaching Agreement over Ontology Alignments," in *5th International Semantic Web Conference (ISWC)*, Athens, GA, US, 2006.
- [15] L. Laera, I. Blacoe, V. Tamma, T. Payne, J. Euzenat and T. Bench-Capon, "Argumentation over Ontology Correspondences in MAS," in *AAMAS'07*, Honolulu, Hawaii, USA, 2007.
- [16] P. Doran, T. Payne, V. Tamma and I. Palmisiano, "Deciding Agent Orientation on Ontology Mappings," in *9th International Semantic Web Conference (ISWC)*, 2010.
- [17] P. Maio and N. Silva, "A Three-Layer Argumentation Framework," in *Theorie and Applications of Formal Argumentation*, Springer Berlin, 2012, pp. 163-180.