HAL
open science

Explaining Argumentation over Alignment Agreements

Rouaa Wannous, Cassia Trojahn dos Santos

To cite this version:

Rouaa Wannous, Cassia Trojahn dos Santos. Explaining Argumentation over Alignment Agreements. IEEE/WIC/ACM International Joint Conferences on Web Intelligence WI and Intelligent Agent Technologies IAT 2013, Nov 2013, Atlanta, United States. pp. 80-85. hal-01151026

HAL Id: hal-01151026

https://hal.science/hal-01151026v1

Submitted on 12 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 12954

To link to this article : DOI :10.1109/WI-IAT.2013.94
URL : http://dx.doi.org/10.1109/WI-IAT.2013.94

To cite this version : Wannous, Rouaa and Trojahn, Cassia *Explaining Argumentation over Alignment Agreements*. (2013) In: IEEE/WIC/ACM International Joint Conferences on Web Intelligence WI and Intelligent Agent Technologies IAT 2013, 17 November 2013 - 20 November 2013 (Atlanta, United States).

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Explaining Argumentation over Alignment Agreements

Rouaa Wannous
Université de La Rochelle, L3I laboratory
La Rochelle, France
rouaa.wannous@univ-lr.fr

Cássia Trojahn
Université de Toulouse 2 & IRIT
Toulouse, France
cassia.trojahn@irit.fr

***Abstract*—Argumentation frameworks have been used as tools for reconciliating ontology alignments, through a series of proposals and counter-proposals, i.e., arguments. However, argumentation outcomes may not be so obvious to human users. Explaining the reasoning behind the argumentation process may help users to understand its outcome, and influence the user's confidence and acceptance on the results. This paper presents a mechanism for providing explanations on the way agreed alignments are established. Our mechanism is based on tracing each step of the argumentation process. These traces are then interpreted using a set of association rules, built from a decision tree that represents all possible statuses of arguments. From these rules, a multi-level explanation, in natural language, is provided to the users.**

***Index Terms*—argumentation; ontology matching; explanation.**

I. Introduction

Ontology matching is the process of finding relationships between entities of different ontologies and has been seen as the solution for the ontology heterogeneity problem in different scenarios. In agent communication [1], for instance, agents need to reconcile the differences between their vocabularies (ontologies) before starting any kind of cooperation. There are many ways to exploit the features within the ontologies and different matching approaches have been proposed in the literature ([2], [3]). Potentially, matching system may differ in the alignments they generate and argumentation theory has been exploited for conciliating matchers' points of view and iteratively to solve their conflicts ([1], [4]).

Argumentation is a negotiation model based on building and comparing arguments, for defining the most acceptable of them. In the context of ontology matching [4], arguments represent correspondences between entities of two ontologies. Agents exchange their arguments in order to identify the most acceptable correspondences. They may have different preferences, which can be taken into account in the process of evaluating and selecting arguments. However, argumentation is a complex process and its outcome may not be intuitively obvious to human users. Thus, users may not agree on the argumentation outcome because they are not able to understand why or how arguments have been accepted. Explanation of argumentation processes has been exploited in few works [5]. Explanation consists of exposing something in a way that is understandable for the receiver of the explanation [6]. It makes system's results more acceptable for users by increasing their understanding on the reasons for finding them.

This paper presents a mechanism for explaining why and how an alignment agreement, from different matchers that argue on their alignments, has been established. In general, matching systems do not provide explanations on their results and very few approaches have been proposed in the literature [7], [8]. Providing matching explanations is one of the challenges for ontology matching [9]. For instance, explanations are useful when matching thousands of entities, where automatic matching of large ontologies may find many plausible correspondences, and hence user input is required for filtering out the alignment [2]. Our explanation mechanism traces each single step of the argumentation process. These traces are then interpreted using a set of association rules, built from a decision tree that represents all possible statuses of arguments. From these rules, a multi-level explanation, in natural language, is provided to the user. It is able to provide explanations on (a) why it was found by an individual matcher, (b) why a correspondence was accepted by an individual matcher, and (c) why it was included in the agreed alignment.

In the following, we introduce ontology matching and argumentation frameworks (§II) and present our explanation mechanism (§III) and evaluation experiments (§IV). Next, we present related work (§V) and conclude the paper (§VI).

II. Foundations

A. Ontology matching

The matching process takes as input two ontologies o and o' and determines as output an alignment A', i.e., a set of correspondences. A correspondence is defined as follows :

Definition 1 (Simple Correspondence [2]) *A simple correspondence can be defined as <e,e′,r,n>, such that: e and e' are entities (e.g., elements, properties, classes) of o and o', respectively; r is a relation holding between two entities e and e' (for instance,* equivalence *($\equiv$),* more general *($\sqsupseteq$),* disjointness *($\perp$),* overlapping *($\sqcap$)); and n is a confidence measure number in the [0;1] range, assigning a degree of trust on the correspondence.*

For instance, $c = \langle article_o, paper_{o'}, \equiv, 1\rangle$ states that *article* in the source ontology is equivalent to *paper* in the target ontology, with a confidence of 1.0.

Different matching approaches have emerged from the literature [2]. The main distinction between each is due to the

type of knowledge encoded within each ontology, and the way it is utilised when identifying correspondences between features or structures within the ontologies. *Terminological* methods lexically compare strings (tokens or n-grams) used in naming entities (or in their labels and comments), whereas *semantic* methods utilise model-theoretic semantics to determine whether or not a correspondence exists between two entities. Approaches may consider the *internal* ontological structure, such as the range of their properties (attributes and relations), their cardinality, transitivity and/or symmetry of their properties, or the *external* ontological structure, such as the position of the two entities within the ontological hierarchy. The instances (or extensions) of classes could also be compared using *extension*-based approaches. Matching systems rely not on a single approach.

B. Argumentation frameworks

In alignment agreement, arguments are positions that support or reject correspondences [10], [11]. Such arguments interact following the notion of attack and are selected according to the notion of acceptability, introduced by Dung [12]. In Dung's model, an argument should be accepted only if every attack on it is attacked by an accepted argument. Dung defines an argumentation framework as follows :

Definition 2 (Argumentation framework [12] (AF)) *An AF is a pair $\langle \mathcal{A}, \ltimes \rangle$, such that $\mathcal{A}$ is a set of arguments and $\ltimes$ (attacks) is a binary relation on $\mathcal{A}$. $a \ltimes b$ 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.*

In Dung's model, all arguments have equal strength and an attack always succeeds (or successfully attacks). Amgoud and Cayrol [13] have 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 local preferences to different audiences, Capon [14] proposes to associate arguments to the values which supports them, leading 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. However, in ontology matching, an objection can still be raised about the lack of complete mechanisms for handling persuasiveness [15]. Matchers output correspondences with a strength that reflects the confidence they have in the fact that the correspondence between the two entities holds, usually derived from similarity assessments made during the matching process. Hence, for associating an argument to a *strength* Trojahn *et al.* [16] introduce the SVAF framework, extending [14]:

Definition 3 (SVAF [16]) *A Strength-based argumentation framework (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 the strength of the argument.*

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* is extended:

Definition 4 (Successful attack [16]) *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*

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

Definition 5 (Acceptable argument [14]) *An argument $a \in \mathcal{A}$ is* acceptable *to an audience α with respect to a set of arguments S, noted $acceptable_\alpha(a, S)$, iff $\forall x \in \mathcal{A}$, $x \dagger_\alpha a \Rightarrow \exists y \in S; y \dagger_\alpha x$.*

A preferred extension represents a consistent position within a framework, which defends itself against all attacks and cannot be extended without raising conflicts [14]. An argument is *subjectively acceptable* iff it appears in some preferred extension for some specific audience, and an argument is *objectively acceptable* iff it appears in all preferred extensions for every specific audience.

C. Arguments on correspondences

Representing correspondences as arguments within an AF and attacks between them are as follows:

Definition 6 (Argument [10], [11]) *An argument $a \in \mathcal{A}$ is a triple $a = \langle c, v, h \rangle$, such that c is a correspondence, $\langle e, e', r, n \rangle$; $v \in V$ is the value of the argument and h is one of +,- depending on whether the argument is that c does or does not hold.*

Definition 7 (Attack [10], [11]) *An argument $\langle c, v, h \rangle \in \mathcal{A}$ attacks an argument $\langle c', v', h' \rangle \in \mathcal{A}$ iff $c = c'$ and $h \neq h'$.*

The way arguments are generated differs in each scenario. The strategy in [16], *negative arguments as failure*, relies on the assumption that matchers return complete results. Each possible pair of ontology entities which is not returned by a matcher is considered to be at risk, and a negative argument is generated $h = -$. An alignment associated with an extension contains only the arguments with $h = +$.

III. Explanation on argumentation

Our explanation mechanism attempts to fill the gap between users and argumentation outcomes. Basically, it is based on (a) tracing the argumentation process in order to register the statuses of arguments during the argumentation process, and (b) using a rule-based mechanism for interpreting the trace entries. The explanations are given to the user in natural language, with different levels of technical details. These levels allow users to navigate through the explanations from general explanations to more specific contextual ones. More specifically, the following explanations are provided: (a) why a correspondence was found by an individual matcher, (b) why it was accepted by an individual matcher, and (c) why it was included in the agreed alignment. In the following, we detail

the justification of correspondences and our explanation on how and why agreements were obtained from argumentation.

A. Justifying correspondences

We assume that matchers can provide one or more justifications about the reasons that support finding their correspondences. In order to consider this information, we extend the definition in §II-A to include the *provenance* information used for finding a specific correspondence :

Definition 8 (Extended correspondence (EC)) *An extended correspondence is a tuple <c,J>, such that: c is a correspondence; J is a set of justifications j_i for finding c.*

From the ontology matching literature, possible justification categories can be established: *terminological*, such as lexical similarity and synonyms; *structural*, such as super-class, sub-class, sibling and property; *semantic*, such as semantics model-based; and *extensional*, such as instance-based. For instance, in the example provided in §II-A, the extended correspondence (EC) between the *article* and *paper* entities is $c = \langle article_o, paper_{o'}, \equiv, 1, \{terminological\}\rangle$ since these terms can be considered synonymous. Hence, an extended argument contains an extended correspondence :

Definition 9 (Extended argument (EA)) *An extended argument (EA) is a triple a = <EC,v,h> such that EC is an extended correspondence <c,J>; $v \in V$ is the value of the argument and h is one of {+,-} depending on whether the argument is that EC does or does not hold.*

B. Explanation mechanism

Overall, the explanation process can be summarised as follows : (1) trace the argumentation process and log the outcomes of each single step of it; (2) a query for a correspondence is made by a user; (3) filter the arguments to a list which contains the queried correspondence; (4) apply the rules on the log entries that give the status of the filtered arguments; (5) provide multi-level explanations in natural language to the user about the queried correspondence.

The explanation mechanism has two main modules: argumentation logger and rule-based interpreter. The argumentation logger module traces each single step of the argumentation process. Each log entry contains the status of each argument, for each matcher. The log entries for each matcher contain the statuses of each single argument $\langle A\rangle$. The argument $\langle A\rangle$ may have many statuses during the argumentation process. The output of this module is a set of log entries.

The rule-based interpreter module contains the rules to interpret the log entries and associates to each entry an explanation in natural language. We assume that the argumentation process includes the instantiation of several argumentation frameworks, from several matchers. For defining all possible statuses of arguments within an argumentation process, we have manually draw a decision tree (Figure 1) based on the analysis of the argumentation process and how argumentation frameworks establish their preferred extensions. From this decision tree, we extracted 9 rules to explain all possible

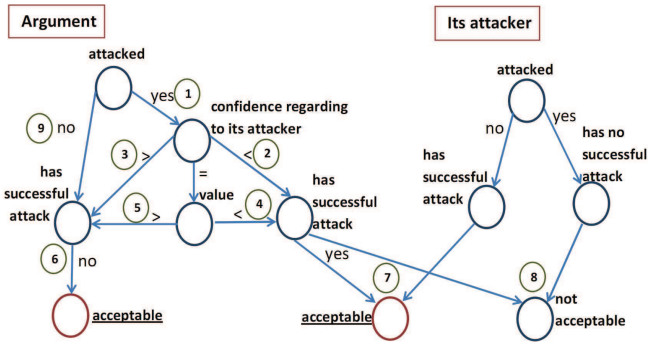

Figure 1. Decision tree represents the statuses of an argument, the numbers in the circles refer to the rule number.

statuses for an argument and the inputs of our rules are the outputs from the argumentation logger module (log entries of a given argument and the log entries of its attackers). For instance, rule (1), according to Figure 1, has the log entry $attacked$ for the considered argument A, while the rule (7) has the log entry $has\ successful\ attack$ for the argument A and for its attacker B. We have also defined two categories of rules: *local rule*, which considers log entries for one single matcher as an input to a rule; and *global rule*, which considers log entries for all matchers as an input to a rule. These categories of rules are transparent to user. Below, we present the rules extracted from the decision tree depicted in Figure 1 (the numbers represent the rule number). $\langle A\rangle$ is an argument and $\langle B\rangle$ is its attacker; $\langle s(A)\rangle$ represents the strength of $\langle A\rangle$ and $\langle \nu(A)\rangle$ is the value of $\langle A\rangle$, which represents the preferred value of the matcher generating it; $\langle c\rangle$ is the $queriedcorrespondence$ the user asked about. In the explanation, EXP is the output from each rule with multi-level explanations in natural language, representing their technical details (i.e, different levels of explanation): *L1* (level 1), where the sentences are presented in an understandable way, without any technical details; *L2* (level 2), where sentences are presented with technical details, supporting the explanation provided in *level 1*. The $acceptable\ conflict$ represents a conflict which is not solved. The $attackrelation$ represents the successful attack between the considered arguments, such as $\langle A\rangle \rightarrow \langle B\rangle$ that means $\langle A\rangle$ successfully attacks $\langle B\rangle$. The argument $\langle C\rangle$ represents the argument that attacks $\langle B\rangle$. Hence, for an argument $\langle A\rangle$, that contains the queried correspondence $\langle c\rangle$, and a matcher $\langle m\rangle$, we define the following local rules :

(1) IF $\langle A\rangle$ is attacked
EXP L1 not all matchers have agreed on aligning $\langle c\rangle$
$\langle m\rangle$ has found it because $\langle J\rangle$
L2 $\langle A\rangle$ has an attack $\langle B\rangle$

(2) IF $\langle s(A)\rangle$ is lower than $\langle s(B)\rangle$
EXP L1 $\langle A\rangle$ is not an acceptable correspondence
L2 $\langle attackrelation\rangle$, there is a conflict
$\langle A\rangle$ has lower strength than $\langle B\rangle$, $\langle A\rangle$ has a successful attack

(3) IF $\langle s(A)\rangle$ is greater than $\langle s(B)\rangle$
EXP L1 $\langle A\rangle$ is an acceptable correspondence
L2 $\langle attackrelation\rangle$, there is no conflict
$\langle A\rangle$ has greater strength than $\langle B\rangle$, it has no successful attack

(4) IF $\langle s(A)\rangle$ is equal to $\langle s(B)\rangle$ AND
$\langle \nu(A)\rangle$ has lower preference than $\langle \nu(B)\rangle$
EXP L1 there is an acceptable conflict
L2 arguments have the same strength $\langle s(A)\rangle$
however, $\langle m\rangle$ has preferred $\langle \nu(B)\rangle$ over $\langle \nu(A)\rangle$
$\langle A\rangle$ has a successful attack

(5) IF $\langle s(A)\rangle$ is equal to $\langle s(B)\rangle$ AND
$\langle \nu(A)\rangle$ has greater preference than $\langle \nu(B)\rangle$
EXP L1 there is no acceptable conflict
L2 arguments have the same strength $\langle s(A)\rangle$
however, $\langle m\rangle$ has preferred $\langle \nu(A)\rangle$ over $\langle \nu(B)\rangle$
$\langle A\rangle$ has no successful attack

(6) IF $\langle A\rangle$ has no successful attack
EXP L1 $\langle A\rangle$ is an acceptable correspondence
L2 $\langle A\rangle$ has no successful attack

(7) IF $\langle A\rangle$ has successful attack AND
$\langle B\rangle$ has successful attack
EXP L1 $\langle A\rangle$ is an acceptable correspondence
L2 $\langle attack\text{-}relation\rangle$, the conflict is solved by $\langle C\rangle$
the argument and its attacker satisfy the notion of acceptability
respect to a set (S), which contains the arguments that
represent the same correspondences
acceptability: the argument has a successful attack, and its
attacker has successful attack, so the argument is acceptable

(8) IF $\langle A\rangle$ has successful attack AND
$\langle B\rangle$ has no successful attack
EXP L1 $\langle A\rangle$ is not an acceptable correspondence
L2 $\langle attack\text{-}relation\rangle$, there is an acceptable conflict
$\langle A\rangle$ has successful attack and $\langle B\rangle$ has not successful attack

For the argument $\langle A\rangle$, that contains the queried correspondence $\langle c\rangle$ and a set of matchers $\langle m_1, .., m_n\rangle$, the global rules have been defined :

(9) IF $\langle A\rangle$ is not attacked for all matchers
EXP L1 all matchers have agreed on aligning $\langle c\rangle$
L2 m_i has found it because $\langle J_i\rangle$
$\langle A\rangle$ is an acceptable correspondence for all matchers

(10) IF $\langle A\rangle$ is objectively acceptable
EXP L1 $\langle A\rangle$ is an acceptable correspondence for all matchers
L2 $\langle A\rangle$ it is in the preferred extension of all matchers

(11) IF $\langle A\rangle$ is subjectively acceptable
EXP L1 $\langle A\rangle$ is accepted from some matchers
L2 $\langle A\rangle$ is in the same preferred extensions of some matchers

For combining the rules and providing an explanation, we consider the notion of objectively acceptable sets to determine the agreed alignment. So, we always apply the *rule (10)*. Furthermore, we combine the *L2* of the *rule (10)* at the end of the *L2* of the rules (3), (6) and (9). We did not apply the explanations of *L1* of these rules because all of them already have the same *L1*. For example, in the *rule (3)*, we did not change its *L1*. However, after combining them, its *L2* will be *L2*: $\langle A\rangle$ has no successful attack and it is in the preferred extensions of all matchers.

IV. Experimental evaluation

A. Data set and matchers

The OAEI benchmark[1] is a reference data set for evaluating matching systems. The tests in this data set are based on one particular ontology dedicated to the domain of bibliography and a number of alternative ontologies of the same domain, for which alignments are provided. For our study case, we have chosen the test 301 of Benchmark. This test is based on aligning the reference ontology to a real one, which is reminiscent to BibTeX. We have choose this test because it is one of the real cases available in Benchmark and it is very known in the ontology matching community. We based our experiments on three basic matchers implemented in the Alignment API[2] : (a) string similarity (*str*): uses an editing (Levenshtein [17]) distance between ontology entity names; (b) synonymous-based similarity (*syn*): looks for synonymous in the WordNet[3] database; (c) property-based similarity (*pro*): establishes an alignment based on the comparison of the properties that classes have in common.

B. Explanation system

We have used the available implementation of S-VAF as basis for implementing our system. The implementation of this framework is available in Java and we have extended it in order to include our explanation mechanism. We implemented our two modules: the argumentation logger and the rule-based interpreter. In the argumentation logger module, each step of generating the preferred extensions in the argumentation process has been logged. The rule-based interpreter module uses these logs to provide the explanations.

The user can run the system on command line by passing all the required input parameters: (a) name of each matcher used for finding the alignments; (b) RDF files which contain the alignments that matchers have generated; (c) confidence for the negative arguments ($h = -$); and (d) HTML output file path and name, which will contain the explanations. Once the user has run the system, the agreed alignment is presented. Then he/she can query about a specific correspondence to see the explanations generated for it. In this initial version of our system, the output of the system is an HTML file containing the explanations on the queried correspondence, which are provided by our rules. This file contains both the explanations for the correspondences in the agreed alignment (for all matchers) and the explanations for the correspondences for each individual matcher's point of view. The first level of explanation (L1) is presented by default to the users (Figure 2). If the user is interested in more details, he/she can navigate on the second explanation level (L2) by clicking on 'More explanation'. Figure 2 shows an explanation interface for the queried correspondence "Institution Incollection".

Our system is able to provide the following explanations (for sake of brevity, we provide examples for few cases):

- Why a correspondence was determined between ontology entities: the matchers reasons for finding this correspondence are given in Figure 3 (1).
- Why a correspondence is in the agreed alignment: the explanations for accepting a correspondence as an agreement are shown in Figure 3 (2).

[1] http://oaei.ontologymatching.org/2011/benchmarks

[2] http://alignapi.gforge.inria.fr/

[3] http://wordnet.princeton.edu/

Query: Institution, Incollection

Correspondences:

1. matcher pro: ‹pro,Institution,Incollection,0.6739130434782609,+›
2. matcher syn: ‹syn,Institution,Incollection,0.5,-›
3. matcher str: ‹str,Institution,Incollection,0.5,+›

Explanation:

(the blue correspondence is the acceptable one, the acceptable conflict represents a conflict which is not solved, A ---> B: means A successfully attacks B)

- not all matchers have agreed on aligning ‹Institution,Incollection›
  - Matchers reasons for finding a correspondence
  - More explanation
- ‹pro,Institution,Incollection,0.6739130434782609,+› is an acceptable correspondence for all matchers
  - Attack relation
  - More explanation
- Matchers point of view

Figure 2. Level 1 explanation for "Institution Incollection" query.

Query: Institution, Incollection

Correspondences:

1. matcher pro: ‹pro,Institution,Incollection,0.6739130434782609,+›
2. matcher syn: ‹syn,Institution,Incollection,0.5,-›
3. matcher str: ‹str,Institution,Incollection,0.5,+›

Explanation:

(the blue correspondence is the acceptable one, the acceptable conflict represents a conflict which is not solved, A ---> B: means A successfully attacks B)

- not all matchers have agreed on aligning ‹Institution,Incollection› (1)
  - Matchers reasons for finding a correspondence
    - matcher syn: did not find it
    - matcher pro: found it because the entities have a property similarity
    - matcher str: found it because the entities have a lexical similarity
  - More explanation
    - ‹pro,Institution,Incollection,0.6739130434782609,+› has an attack
    - ‹syn,Institution,Incollection,0.5,-›
- ‹pro,Institution,Incollection,0.6739130434782609,+› is an acceptable correspondence for all matchers (2)
  - Attack relation
    - ‹pro,Institution,Incollection,0.6739130434782609,+› ---> ‹syn,Institution,Incollection,0.5,-›
  - More explanation
    - there is no conflict, ‹pro,Institution,Incollection,0.6739130434782609,+› has greater strength than ‹syn,Institution,Incollection,0.5,-›
    - it has no successful attack and it is in the preferred extensions of all matchers
- Matchers point of view

Figure 3. Navigation in the explanations for "Institution Incollection" query.

- How a conflict was solved between two correspondences: the explanations show the conflict and the responsible for solving it.
- Why a correspondence was acceptable by a particular matcher: the explanations, show the acceptability of a correspondence by a matcher.
- Why a correspondence was not in the agreed alignment: the explanations for not accepting a correspondence as an agreement.
- When a correspondence was not found between ontology entities (for all matchers): if the user asks about two entities, such as "title Entry", the explanation is "None of these matchers find any correspondence between them".

C. Evaluation and discussion

Evaluating explanation requires human user feedback [18] [19]. Here, we evaluate the explanation in a qualitative way, focusing on criteria such as clearness and usefulness. To that extent, we have prepared a questionnaire (8 questions) that has been used to evaluate each case our system provides an explanation. Each question in the questionnaire had four options to select from: No, Maybe, Yes and Absolutely. Also the user could add any comments they may have. These four options were interpreted as numeric values from 0-3, respectively (no $\rightarrow$ 0, Maybe $\rightarrow$ 1, etc.). We have asked two users to fill in the questionnaire, after using our system. Both of them are computer scientists. One is an expert on the ontology matching domain and the other has good notions of ontology matching. Figure 4 shows the evaluation results, for each question in the questionnaire. Regarding to the clearness of the explanations on the agreed alignment, the expert evaluator found the explanations clear while the second evaluator found the explanations not fully clear. With respect to the acceptability of a particular correspondence by a matcher, both of reviewers were not completely convinced with the explanations. It is the same case for the matchers justifications on finding their correspondences. In this point, as we may have no control on the matching process itself, we can not provide more complete or better explanations. Regarding to the clearness of the explanations on solving a conflict, they are also convinced but not fully.

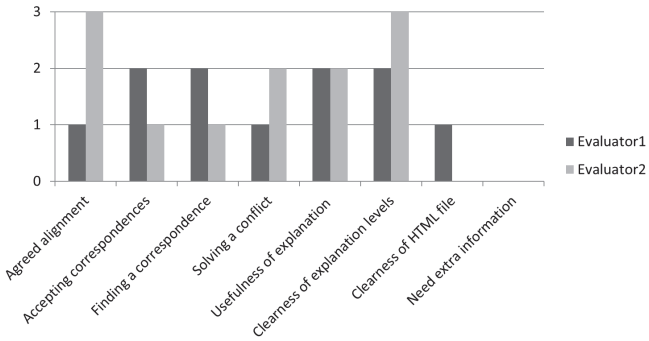

Figure 4. Evaluating the explanations provided by our system.

With respect to the usefulness of our explanations for understanding the argumentation outcomes, they agree on their usefulness. The evaluators almost fully agree on the clearness of the explanation levels for supporting the understanding. The drawback of our system, however, as stated by the evaluators, is that the way we have provided the explanations (i.e., explanations as long natural language sentences) in HTML file is not clear enough for fully understanding the interactions within the argumentation process. They stated that exploiting more graphical visualisation options could be a solution to this drawback.

V. Related Work

Since many years, explanation has been exploited in different areas of artificial intelligence, such as proof [20], expert systems [21] and belief revision [22]. In the proof domain, provers are based on providing reasoning details, usually as reasoning trees, what is easily understandable to experts. In expert systems, different types of explanation, such as trace, justification and strategy, are exploited to provide to human experts the explanations on how the system has taken its decisions. Rule-based and case-based systems exploit differently the knowledge for explaining their decisions. In the specific context of ontology matching, few systems are able

to provide explanations on their alignments [7] [8]. S-Match system [7] uses the Inference Web (IW) infrastructure to provide proofs and explanations to the results from matching process steps. Following a different strategy, IMAP system [8] explains its correspondences by retrieving the justifications from a dependency graph, which represents the steps the system followed to establish its correspondences. Scarlet[4] presents how two entities can be related, using information from online available ontologies.

In argumentation frameworks and explanation of its argumentation reasoning process, the closest proposal is from Garcia *et al.* [5]. They proposed a dialectical proof based explanation, based on Dung's framework, for explaining the acceptability of arguments. They provide dialectical trees which represent the interaction between arguments, as an explanation to the user. More recently, these authors formalise dialectical explanation for argumentation in [23]. Similar to provers, this kind of dialectical explanation may be not so intuitive. While the explanations in deductive and dialectical systems are based on reasoning and dialectical trees, respectively, we focus on natural language explanations. In that way, similar to what is done in expert systems, we trace each step of the argumentation process and make a decision tree, which represents the status of each argument within an argumentation framework. The multi-levels of explanations in our system represent different levels of technical details, as in IMAP. Although our focus is on explaining argumentation, if available, we provide justifications from matchers.

VI. Conclusion and future work

This paper has proposed a first attempt to provide multi-level natural language explanations on agreed alignments, using a specific framework (SVAF). Although few users have participated in the evaluation, their feedbacks were useful for evaluating the clearness and the usefulness of our explanations. The overall evaluation is positive, however, there is a room for improvements and our work can be improved in several ways. First, alternative ways for visualising our explanations could be exploited, specially providing more graphics, using coloured graphs of attacks between arguments, and colouring attacks that succeeds, as well as allowing the user to dynamically choose different argumentation settings (agents preferences, strength of attacks, etc). Second, we plan to provide explanations on alternative argumentation frameworks. Third, we can work on extending the justification on finding correspondences that come from matchers. Fourth, we have focused on explanations for users. However, we could think on explanations between matchers, in a similar way of what is done in belief revision. A matcher can explain the way it finds its correspondences for another matchers, in an explanation-based dialogue. Finally, we would like to propose an approach for learning decision trees from the agents interaction.

[4]http://scarlet.open.ac.uk/

References

[1] L. Laera, I. Blacoe, V. Tamma, T. Payne, J. Euzenat, and T. Bench-Capon, "Argumentation over Ontology Correspondences in MAS," in *Proceedings of 6th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS)*, 2007, pp. 1285–1292.

[2] J. Euzenat and P. Shvaiko, *Ontology matching*. Springer-Verlag, 2007.

[3] Y. Kalfoglou and M. Schorlemmer, "Ontology Mapping: The State of the Art," *The Knowledge Engineering Review*, no. 04391, pp. 1–31, 2005.

[4] C. Trojahn, J. Euzenat, V. Tamma, and T. Payne, "Argumentation for reconciling agent ontologies," in *Semantic Agent Systems*, ser. Studies in Computational Intelligence, 2011, vol. 344, pp. 89–111.

[5] A. J. García, C. I. Chesnevar, N. D. Rotstein, and G. R. Simari, "An abstract presentation of dialectical explanations in defeasible argumentation," in *Proceedings of Workshop on Argumentation and Non-Monotonic Reasoning (ArgNMR)*, 2007, pp. 12–32.

[6] L. R. Ye and P. E. Johnson, "The impact of explanation facilities on user acceptance of expert systems advice," *MIS*, vol. 19, no. 2, pp. 157–172, 1995.

[7] P. Shvaiko, F. Giunchiglia, P. P. D. Silva, and D. L. McGuinness, "Web explanations for semantic heterogeneity discovery," Tech. Rep., 2004.

[8] Y. Qian, Y. Li, J. Song, and L. Yue, "IMAP: Discovering Complex Semantic Matches between Database Schemas," in *Proceedings of the 2009 International Conference on Web Information Systems and Mining (WISM)*, 2009, pp. 756–760.

[9] P. Shvaiko and J. Euzenat, "Ten challenges for ontology matching," in *Proceedings of the 7th Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE)*, 2008, pp. 1163–1181.

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

[11] C. Trojahn, P. Quaresma, and R. Vieira, "Conjunctive queries for ontology based agent communication in mas," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2008, pp. 829–836.

[12] P. Dung, "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n–person games," *Artificial Intelligence*, vol. 77, no. 2, pp. 321–357, 1995.

[13] L. Amgoud and C. Cayrol, "On the acceptability of arguments in preference-based argumentation," in *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 1998, pp. 1–7.

[14] T. Bench-Capon, "Persuasion in practical argument using value-based argumentation frameworks," *Journal of Logic and Computation*, vol. 13, no. 3, pp. 429–448, 2003.

[15] A. Isaac, C. Trojahn, S. Wang, and P. Quaresma, "Using quantitative aspects of alignment generation for argumentation on mappings," in *Proceedings of the 3rd Workshop on Ontology Matching*, 2008.

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

[17] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals." *Soviet Physics Doklady*, pp. 707–710, 1966.

[18] D. Doyle, P. Cunningham, and P. Walsh, "An Evaluation of the Usefulness of Explanation in a CBR System for Decision Support in Bronchiolitis Treatment," in *Proceedings of the Workshop on Case-Based Reasoning in the Health Sciences, Workshop Programme at the 6th International Conference on CaseBased Reasoning*, 2005, pp. 32–41.

[19] P. Cunningham, D. Doyle, and J. Loughrey, "An Evaluation of the Usefulness of Case-Based Explanation," in *Proceedings of the 5th International Conference on Case-Based Reasoning*, 2003, pp. 122–130.

[20] D. L. Mcguinness and A. T. Borgida, "Explaining subsumption in description logics," *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pp. 816–821, 1995.

[21] W. Swartout, C. Paris, and J. Moore, "Explanations in knowledge systems: Design for explainable expert systems," *IEEE Expert: Intelligent Systems and Their Applications*, vol. 6, pp. 58–64, 1991.

[22] M. A. Falappa, G. Kern-Isberner, and G. R. Simari, "Explanations, belief revision and defeasible reasoning," *Artificial Intelligence*, vol. 141, pp. 1–28, 2002.

[23] A. J. García, C. I. Chesnevar, N. D. Rotstein, and G. R. Simari, "Formalizing dialectical explanation support for argument-based reasoning in knowledge-based systems," *Expert Syst. Appl.*, vol. 40, no. 8, pp. 3233–3247, 2013.