An interactive, asymmetric and extensional method for matching conceptual hierarchies

Jérôme David, Fabrice Guillet, Régis Gras, and Henri Briand

LINA CNRS FRE 2729,
Polytechnic School of Nantes University,
3 rue Christian Pauc
44306 Nantes, France
{jerome.david, fabrice.guillet, regis.gras,
henri.briand}@polytech.univ-nantes.fr

Abstract. Our work deals with schema or ontology matching and is driven by the following statements: (1) Most of works only consider intensional description of schemas; (2) They mostly use symmetric similarity measures (and then they match similarity relations betwen concepts); (3) Few prototypes allow an interactive and visual match process. Therefore, we suggest an extensional and asymmetric matching method based on the discovery of significant implication rules between concepts described in textual documents. Our approach relies on the association rules paradigm and use a probabilistic model of deviation from independence, named implication intensity. Our matching method is divided into two consecutive stages: (1) the extraction in documents of relevant terms for each concept; (2) the discovery of significant implications between the concepts. And finally, we enclose this matching approach into an interactive visualization tool in order to facilitate the analyse, the validation and the editing of a mapping set for the knowledge engineer.

Keywords: ontology matching, extensional matching, association rules, implication intensity, matching visualization.

1 Introduction

The hierarchical categorization of data through ontological forms as taxonomies is widely used with the increase of electronic data and knowledge on the Internet or in companies. Web directories such as Yahoo.com and OpenDirectory, the Electronic Document Management, or the Semantic Web with its OWL ontology are examples of such taxonomies.

In the literature, a lot of works deals with schema/ontology matching. The schema or ontology matching aims at finding semantic relations (i.e. equivalence, subsomption, etc) between entities (i.e. concepts, properties) of two schemas/ontologies. These approaches use various techniques such as machine learning [1], FCA-Analysis [2], database schema matching [3], graph matching [4]. These approaches are commonly based on similarity measures for discovering equivalence relations between concepts.

However, the extracted matchings can be enhanced by using asymmetric measures, which deliver more accurate information in the form of implications between concepts. For instance, the use of such measures enables the discovery of equivalence relations between concepts (example : if $car \rightarrow auto$ and $auto \rightarrow car$ then $auto \leftrightarrow car$), and also it can detect if a concept is more specific than another (example : $car \rightarrow vehicle$). In knowledge discovery in databases (KDD), asymmetric measures, called interestingness measures, are widely used for association rules discovery [5]. Association rules are expressions of the type "if *antecedent* then *consequent*" representing implicative tendencies between conjunctions of attributes in databases.

In this paper, we evaluate the use of such asymmetric measures for matching concepts of schemas or ontologies by using the Implication Intensity [6, 7], a probabilistic model of deviation from statistical independence.

Our matching method is both extensional and terminological. It is designed to be used on taxonomies of concepts associated with textual documents. The idea underlying our approach considers that one concept is more specific than another, if the vocabulary used in the documents associated to the first concept tends to be included in the vocabulary of the other one.

Our method is divided into two consecutive stages: (1) The extraction of concept-relevant terms; (2) The discovery of association rules rules between concepts.

The results provided by matching algorithms are not the perfect solution to one matching problem. Thus, the knowledge engineer (i.e. the domain expert) must be able to analyse and edit the produced results. Nervertheless, most prototypes/approaches do not offer a user-friendly interface allowing an interactive match process [8]. We suggest to enclose our method into an anthropocentric step of validation. In this paper, we also present an interactive graphical tool allowing a domain expert to build and validate a matching set between conceptual hierarchies populated with textual documents. For example, this tool is a good decision helper for comparing or merging two products catalogs, two electronic documents bases or two lightweight ontologies.

This paper is organized as follows. In a first section we give an overview of matching approaches. Then, we introduce the Implication Intensity measure, before describing the concept hierarchy model, and the first stage concerning the extraction of concept-relevant terms. Next, we detail the rule extraction stage. Finally, we present our matching visualization tool and show the results obtained on a benchmark dataset.

2 Related works

Many surveys about ontology and schema matching have been proposed in literature [9], [10], [11]. The two last ones propose a classification and a comparative study of matching approaches. The survey [10] focuses on the database schema matching approaches, while [11] reuses this classification for ontology matching. From these surveys we can distinguish: the extensional approaches (or elementbased), and the intentional approaches (or only-schema-based). The matching approaches can be also discriminated regarding the kind of relations that they are based on. Some consider symmetric (equivalence) relations, while other ones also use asymmetric relations such as the subsumption or implication.

The main part of these works propose to process the concept name by using string-similarities (Anchor-PROMPT [12], Cupid [3], Coma [8], S-MATCH [13]) or/and external oracles such as Wordnet (H-MATCH [14], [13]). They can also use the schema or ontology structure (Similarity Flooding [4], Artemis [15], [8], [12], [3], [14]).

Most of these approaches are intensional and symmetric. None of them are both asymmetric and extensional. Among extensional approaches, we can cite GLUE [16]. This symmetric approach uses Bayesian learners in order to classify instances of the first ontology into the other and vice-versa in order to estimate the joint probability distribution and then predict concept similarities.

We can also notice that there is only one intensional method distinguishing asymmetric relations. The method S-MATCH [13] search equivalence ($=$) relation between concepts but also the more general ($\sqsupseteq$), less general ($\sqsubseteq$), mistmatch ($\perp$) and overlapping ($\sqcap$) relations. This method use a lot of single matchers: 13 linguistic-based matchers and 3 logic-based matchers.

3 The definition of the Implication Intensity

Let us now consider a finite set T of n individuals described by a set I of p items. Each transaction t can be considered as an itemset so that $t \subseteq I$. We denote by $A = \{t \in T; a \subseteq t\}$ the extension of itemset a and we denote by $\overline{B} = T - \{t' \in T; b \subseteq t'\}$ the complementary of the extension of b (i.e. the extension of $\bar{b}$). Then, we introduce the quantities $n_a = card(A)$, $n_{\bar{b}} = card(\overline{B})$ and $n_{a\wedge\bar{b}} = card(A \cap \overline{B})$.

An association rule [5] is an implication of the form $a \rightarrow b$, where a and b are disjoined itemsets. In practice, it is quite common to observe a few transactions which contain a and not b without having the general trend to have b when a is present contested. Therefore, the number $n_{a\wedge\bar{b}}$ of counter-examples must be taken into account to statistically accept to retain or not the rule $a \rightarrow b$.

More precisely, we compare the observed number of counter-examples $n_{a\wedge\bar{b}}$ to a probabilistic model noted $N_{a\wedge\bar{b}}$. Let us assume that we randomly draw two subsets X and Y in T which respectively contain n_a and n_b transactions, i.e. $N_{a\wedge\bar{b}} = card(X \cup \overline{Y})$.

The implication intensity of the association rule $a \rightarrow b$ is defined by:

$$\varphi(a \rightarrow b) = 1 - \Pr(N_{a\wedge\bar{b}} \leq n_{a\wedge\bar{b}}) \tag{1}$$

The distribution of the random variable $N_{a\wedge\bar{b}}$ depends on the drawing mode [6]. Here, we use a Poisson distribution with $\lambda = n_a n_{\bar{b}}/n$.

4 The concept hierarchy model and the extraction of concept-relevant terms

Our approach (figure 1) is designed for conceptual hierarchies of concepts organized by a partial order relation, connected to a set of textual documents.

We define a conceptual hierarchy $\mathcal{H}$ as a quadruplet:

$$\mathcal{H} = (C, \leq, D, \sigma_0) \tag{2}$$

where C is a set of concepts, $\leq$ represents the partial order, D is the set of documents, and σ_0 is the relation which associates a set of documents to each concept (i.e. for a concept $c \in C$, $\sigma_0(c)$ represents the documents associated to c). From the partial order $\leq$, we extend the relation σ_0 to σ, where:

$$\sigma(c) = \bigcup_{c' \leq c} \sigma_0(c') \tag{3}$$

In a first stage, we transform the hierarchy $\mathcal{H}$ defined on documents in a hierarchy $\mathcal{H}'$ defined on terms as follows:

$$\mathcal{H}' = (C, \leq, T, \gamma_0) \tag{4}$$

where T is the set of relevant terms extracted from D, and $\gamma_0 \subseteq C \times T$ is the relation associating terms to concepts (i.e. $\gamma_0(c)$ represents the set of relevant terms selected for the concept c). From σ and the relation δ linking terms to documents (i.e. $\delta(t)$ is the set of documents in which the term t appears), we can deduce the relation γ_0. Technically, this is done by evaluating association rules $t \rightarrow c$ (between a term t and a concept c) with the implication intensity measure. A such rule means that the term t tends to appear in documents associated the concept c. The relevant term set of the concept c, noted $\gamma_0(c)$, is defined as follows:

$$\gamma_0(c) = \{t \in T_0 | \varphi(t \rightarrow c) > \varphi_t\} \tag{5}$$

where T_0 represents the set of the binary terms (terms composed of two meaningful words) and of the verbs contained in the documents. Binary terms have the advantage to be more informative and less ambiguous than simple words: they permits to avoid the problem of polysemy. The acquisition of binary terms is performed with the software program ACABIT [17] on previously POS-tagged and stemmed textual documents. φ_t is the implication intensity threshold value and $\varphi(t \rightarrow c)$ is the implication intensity value of the rule $t \rightarrow c$ defined by:

$$\varphi(t \rightarrow c) = 1 - Pr\left(N_{t \wedge \overline{c}} \leq n_{t \wedge \overline{c}}\right) \tag{6}$$

where $n_{t \wedge \overline{c}} = card(\delta(t) - \sigma(c))$ is the observed number of counter-examples, that is to say documents which contain the term t and which are not associated with the concept c. And $N_{t \wedge \overline{c}}$ is the expected number of counter-examples under independence hypothesis.

From the partial order $\leq$, we extend the relation γ_0 to γ, where:

$$\gamma(c) = \bigcup_{c' \leq c} \gamma_0(c') \tag{7}$$

The common term set of two hierarchies $\mathcal{H}'_1 = (C_1, \leq_1, T_1, \gamma_1)$ and $\mathcal{H}'_2 = (C_2, \leq_2, T_2, \gamma_2)$ is noted $T_{1\cap 2} = T_1 \cap T_2$. Next, we define the relation $\gamma_{1\cap 2}$ which associates a subset of $T_{1\cap 2}$ for each concept $c \in C_1 \cup C_2$:

$$\gamma_{1\cap 2}(c) = \begin{cases} \gamma_1(c) \cap T_2 \; if \; c \in C_1 \\ \gamma_2(c) \cap T_1 \; if \; c \in C_2 \end{cases} \tag{8}$$

An implicative match set between two hierarchies $\mathcal{H}'_1$ and $\mathcal{H}'_2$ is a set of implicative rules. A rule $a \rightarrow b$ between the concepts $a \in C_1$ and $b \in C_2$ represents a quasi-implication (i.e. an implication tendancy) from the set of terms $\gamma_{1\cap 2}(a)$ to the set of terms $\gamma_{1\cap 2}(b)$.

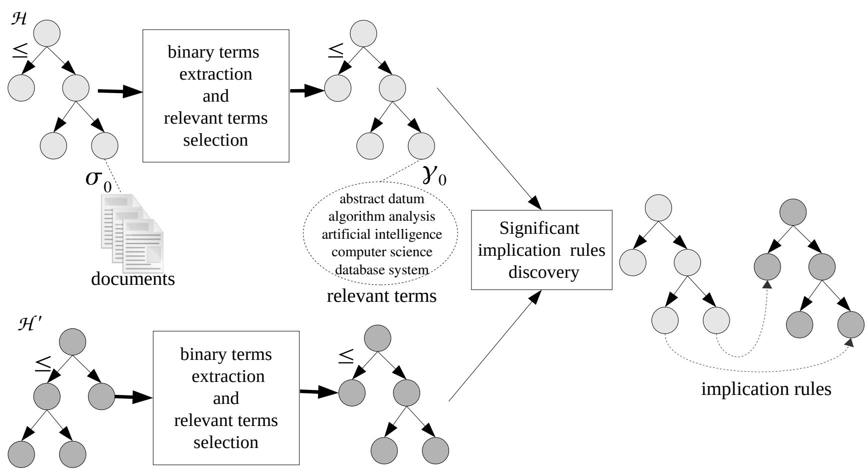

Fig. 1. Methodology scheme

5 Discovery of significant rules between concepts

5.1 Selection criteria of significant rules

In section 4, we have defined a match result as a set of implication rules between concepts issued from two hierarchies $\mathcal{H}_1$ and $\mathcal{H}_2$. Nevertheless, a lot of rules can be discovered. In this section, we define the implication intensity of a rule between concepts, and then we give two criteria defining the notion of significant rule.

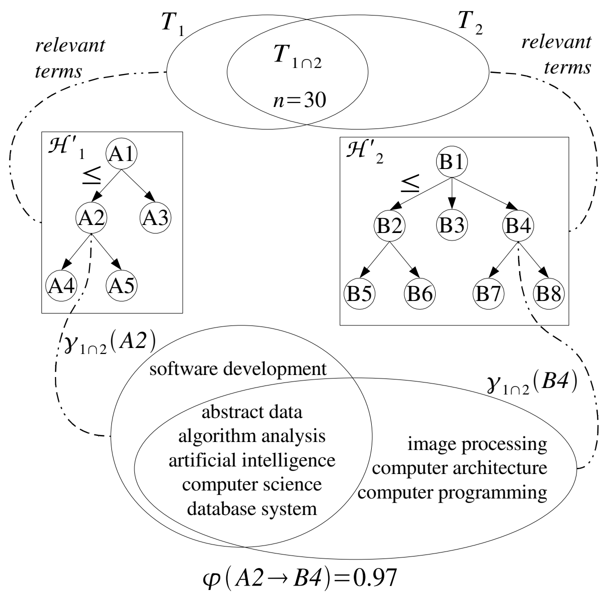

Fig. 2. Evaluation of significant rules

The implication intensity of a rule $a \rightarrow b$ (with $a \in C_1$ and $b \in C_2$) is defined by:

$$\varphi(a \rightarrow b) = 1 - Pr\left(N_{a \wedge \bar{b}} \leq n_{a \wedge \bar{b}}\right) \tag{9}$$

where $n_{a \wedge \bar{b}} = card(\gamma_{1\cap 2}(a) - \gamma_{1\cap 2}(b))$ is the number of relevant terms for concept a which are not relevant for concept b. $N_{a \wedge \bar{b}}$ is the expected number of relevant terms for concept a which are not relevant for concept b. On figure 2, the rule $A2 \rightarrow B4$ has $n_{A2 \wedge \overline{B4}} = 1$ counter-examples. Its implication intensity value is:

$$\varphi(A2 \rightarrow B4) = \sum_{k=0}^{n_{A2 \wedge \overline{B4}}} e^{-\lambda}.\frac{\lambda^k}{k!} = 0,97$$

where $\lambda = n_{A2}.n_{\overline{B4}}/n = 6.(30-8)/30$ (see figure 2).

Thus, the two criteria defining a significant rule are, first, its implication intensity value and, second, the specificity of its consequent combined with the generality of its antecedent. A rule $a \rightarrow b$ (with $a \in C_1$ and $b \in C_2$) will be significant if:

$$\varphi(a \rightarrow b) \leq \varphi_r \tag{10}$$

$$and\ \forall x \geq a,\ \forall y \leq b,\ \varphi(x \rightarrow y) \leq \varphi(a \rightarrow b) \tag{11}$$

The second criterion (equation 11) selects only generative rules and then permits to reduce redundancy in the extracted rules set. Indeed, from the rule $a \rightarrow b$, we can deduce all the rules of the form $x \rightarrow y$ because at the term level: $\gamma_{1\cap 2}(b) \subseteq \gamma_{1\cap 2}(y)$ and $\gamma_{1\cap 2}(x) \subseteq \gamma_{1\cap 2}(b)$. We say that the rule $a \rightarrow b$ is generative of the rules set $x \to y$. For example (figure 2), the rule $A2 \to B4$ is generative of the rules set $\{A2 \to B1, A4 \to B4, A5 \to B4, A4 \to B1, A5 \to B1\}$.

5.2 Algorithms for rule extraction

During the rule extraction step, we can reduce the computation time with the help of the partial order. A top-down search phase enables us to avoid the evaluation of rules having too specific antecedents. This section presents our selection strategy divided into two algorithms.

Inputs :
  A : a concept of $\mathcal{H}_1$.
  $\mathcal{B}_{current}$: a set of concepts taking from $\mathcal{H}_2$.
Procedure specializeAntecedent(A, B)
Begin
  ForEach $B_x \in \mathcal{B}_{current}$ **Do**
    $specializeConsequent(A, B_x, \mathcal{B}_{current}, 0.0)$
  End Do
  ForEach $child \in children(A)$ **do**
    $\mathcal{B}'_{current} := \mathcal{B}_{current}$
    $specializeAntecedent(child, \mathcal{B}'_{current})$
  End Do
End

Fig. 3. Algorithm specializing the antecedent

Our first algorithm (figure 3) takes in a concept a from the hierarchy $\mathcal{H}_1$ and a set of concepts $\mathcal{B}_{current} \subset C_2$ from $\mathcal{H}_2$. For each concept of $\mathcal{B}_{current}$, the second algorithm (figure 4) searches and selects valid consequents. It also updates the set $\mathcal{B}_{current}$. And then, this first procedure is recursively launched over the children of a and with a copy of the set $\mathcal{B}_{current}$. The set $\mathcal{B}_{current}$ contains the subtrees of $\mathcal{H}_2$ with concepts that were selected during the previous recursion steps.

The second algorithm (figure 4) searches a set of valid consequents for the current antecedent a. The search is performed over the set candidate consequents $\{B_x | B_x \leq_2 B\}$. A consequent b_s will be selected if the rule $a \to b_s$ satisfies the two criteria 10 and 11.

This algorithm provides a top-down search of rules in $\mathcal{H}_2$, and then explores all branches of the hierarchy. We choose to stop the descent in a branch if $\forall b'_x \leq_2 b_x, \varphi(a \to b'_x) < \varphi_r$. For a rule $x \to y$, a property of implication intensity defines $x \cup y$ as the best specialization of the consequent. We exploit this property in order to avoid the evaluation of all rules $a \to b'_x$.

The describing search method does not consider the roots of hierarchies because all selected terms are associated to root-concepts. The implication intensity value of such rules (i.e. rules which contain root-concepts) is either undefined or equal to 0.

```
Global variable :
  φ_r : The Implication Intensity threshold
Inputs :
  A : a concept of H_1.
  B : a concept of H_2.
  φ_max : The best value ϕ(A → B_p) with B ≤ B_p
Input/Output variables :
  B_current : the list of "current" concepts taking from H_2.
  ruleList : the list of selected rules.
return value :
  The value φ of the best rule A → B_x with B_x ≤ B

Function specializeConsequent(A, B, B_current, φ_max)
Begin
  bestChild := FALSE
  φ_current := φ(A, B)
  returnVal := φ_current
  If (φ_current < φ_r) then
    φ' := φ(A, A ∩ B)
    If (φ' < φ_r) then
      return φ_current
    EndIf
  EndIf
  ForEach child ∈ children(B) do
    φ_child := specializeConsequent(A, child, B_current)
    If (φ_child ≥ φ_current) then
      bestChild := TRUE
      B_current := B_current − {B}
      If (returnVal < φ_child) then
        returnVal := φchild
      EndIf
    EndIf
    If (φ_current > φ_r) and ⌉(bestChild) and (φ_current ≥ φ_max) then
      ruleList := ruleList ∪ (A → B)
      B_current := B_current ∪ {B}
      φ_max := φ_current
    EndIf
  EndDo
  return returnVal
End
```

Fig. 4. Algorithm specializing consequent

6 Experiments and interactive visualization

6.1 The analysed data

We experimented our algorithms and our interactive visualization tool on a benchmark proposed in [1]. The benchmark "Course catalog" describes courses which are proposed at the Cornell and Washington universities. The courses descriptions are hierarchically organised. These two hierarchies contain respectively 166 and 176 concepts to which are associated 4360 and 6957 textual course descriptions.

6.2 Visualization of the mapping set and interaction

Our maching algorithms are enclosed into a decision helper system which enables the visualization of a matching. This tool aims to help a domain expert to study, edit and validate a matching set pairs between hierarchies. The tool provides a visual metaphor based on graphs: it represents the two hierarchies by trees where nodes are concepts and edges represent the partial-order relations. The implicative matching pair set is represented by directed edges valued by the implication intensity value. Therefore, the global view of the matching graph is intricate and confused. From this general view, the user can intuit the general structure of the matching and find concepts having a lot of implications connected to.

In order to lighten the graph display, the tool proposes a zoom and a φ_r threshold chooser which permits to filter the displayed relations. The user can also choose to keep only the more generative significant rules. Finally, the concepts which do not match other concepts can be hidden.

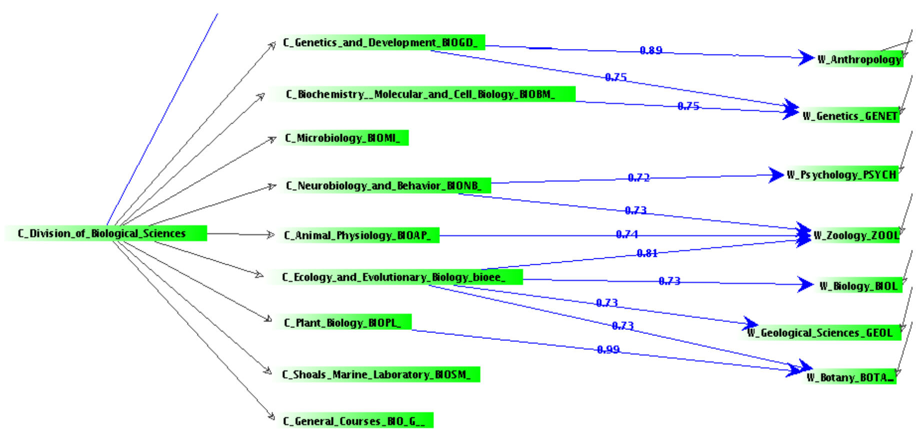

Fig. 5. Lightened and zoomed representation of an interesting part of a matching graph

The user can select concepts and then visualizes its related information: its number of documents attached, its list of significant terms, and the list of concepts that it matches. In the same way, information about matching relations can be displayed: the source and target concepts, the relevant terms shared by the two concepts (examples), the relevant terms only associated with one concept and the more specific relations that it generates. Therefore, this information helps the user to decide if a matching relation must be kept or removed.

On the figure 5, we kept only generative rules having an intensity implication value greater than 0.7. For example, if the expert is interested by the rule "PLANT_BIOLOGY → BOTANY", he can select this and study the underlying terms (figure 6).

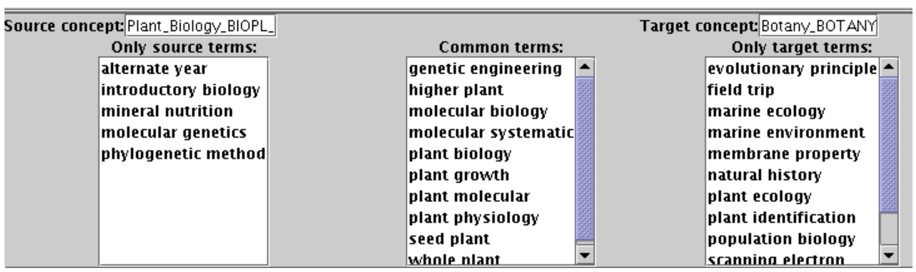

Fig. 6. Details of an implication rule

This tool can load several kind of hierarchies: OWL ontologies, filesystem directory structure with her textual contents. The matching result can be exported into an RDF format for ontology alignment.

6.3 Experimental results

Here, we present results provided by a qualitative test of precision and recall. We compare the results produced by our approach with a reference matching pair set provided by [1]. However, the reference relations are symmetric while ours are asymmetric. In order to retain only equivalence relations, we symmetrized our results by following this rule: *If* $a \rightarrow b$ *and* $b \rightarrow a$ *then* $a \leftrightarrow b$.

Then, we varied the two thresholds φ_t and φ_r (from 0.8 to 1) and computed the precision and the recall values according to the "reference" matching set. These two measures issued from information retrieval are defined as follows: let us considers F the set of matching pairs found using our approache and R the set of "reference" matching pairs. The precision ($precision = card(F \cap R)/card(F)$) measures the ratio of the number of good matching pairs (i.e. matching pairs that are both in our result set and in the reference matching set) over the number of matching pairs found by our method. The recall ($recall = card(F \cap R)/card(R)$) measures the ratio of good matching pairs over the number of reference matching pairs.

Results show that the term selection threshold φ_t has a greater influence than the rule selection threshold φ_r. We obtain good precision values (from 0.71 to 1). Nevertheless, the recall values are quite bad: the best value is equal to 0.54. Our method seems to be too selective. These results can be firstly explained by the the lack of textual data associated with the dataset. As [1] remarks, a lot course descriptions contain only vacuous phrases such as ”3 credits” and then a lot of leave-concepts do not have relevant terms selected. The second explanation concerns the kind of relations that we use. Indeed, the symmetrization of our results introduces a strong bias: a lot of matching pairs contained in the reference matching set are considered as simple implications and not as equivalences.

7 Conclusion

In this paper, we propose an extensional matching method based on the discovery of significant implication rules between concepts. Our approach takes in two hierarchies of concepts to be matched and the textual corpus indexed to these concepts. The matching task is divided into two stages: (1) the extraction and selection of relevant terms for each concepts; (2) the discovery of significant rules between concepts by using their relevant terms set. The main advantages of this method are the consideration of semantics by using binary terms contained in the corpus and the discovery of rules allowing to enhance the produced matching results only regarding similarity-based matching systems. We implemented our algorithms and an interactive visualization tool. This tool helps a domain expert to edit and validate a matching between conceptual hierarchies populated with textual documents. We illustrated our tool on a real conceptual hierarchy related to university courses descriptions and we compared our results to a manual matching reference.

Currently, we propose an extensional individual matcher. In the near future, we will propose a schema based matcher and combine the two approaches in order to enhance the matching task. We will also enhance the mapping visualization by using more improved graph drawing algorithms and add more editing features for facilitate the validation step.

References

1. Doan, A., Madhavan, J., Domingos, P., Halevy, A. In: Ontology Matching : a machine learning approach. Springer-Velag (2004) 397–416
2. Stumme, G., Maedche, A.: FCA-MERGE: Bottom-up merging of ontologies. In: IJCAI. (2001) 225–234
3. Madhavan, J., Bernstein, P.A., Rahm, E.: Generic schema matching with cupid. In: the International Conference on Very Large Data Bases (VLDB'01). (2001) 49–58
4. Melnik, S., Garcia-Molina, H., Rahm, E.: Similarity flooding: A versatile graph matching algorithm and its application to schema matching. In: the 18th International Conference on Data Engineering (ICDE'02), IEEE Computer Society (2002) 117–128

5. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: the 1993 ACM SIGMOD international conference on Management of data, ACM Press (1993) 207–216
6. Gras et al., R.: L'implication statistique, une nouvelle méthode exploratoire de données. La pensée sauvage (1996)
7. Blanchard, J., Kuntz, P., Guillet, F., Gras, R.: 28. In: Implication intensity: from the basic statistical definition to the entropic version. CRC Press (2003) 473–485
8. Do, H., Rahm, E.: Coma - a system for flexible combination of schema matching approaches. In: the International Conference on Very Large Data Bases (VLDB '02). (2002) 610–621
9. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art. Knowedge Engineering Review **18**(1) (2003) 1–31
10. Rahm, E., Bernstein, P.A.: A survey of approaches to automatic schema matching. The VLDB Journal **10**(4) (2001) 334–350
11. Shvaiko, P., Euzenat, J.: A survey of schema-based matching approaches. Journal on Data Semantics IV **4**(LNCS 3730) (2005) 146–171
12. Noy, N., Musen, M.: Anchor-prompt: Using non-local context for semantic matching. In: the Workshop on Ontologies and Information Sharing at the International Joint Conference on Artificial Intelligence (IJCAI). (2001) 63–70
13. Giunchiglia, F., Shvaiko, P., Yatskevich, M.: S-match: an algorithm and an implementation of semantic matching. In: European Semantic Web Symposium. LNCS 3053 (2004) 61–75
14. Castano, S., Ferrara, A., Montanelli, S.: Matching ontologies in open networked systems: Techniques and applications. Journal on Data Semantics **3870**(V) (2006) 25–63
15. Castano, S., Antonellis, V.D., Vimercati, S.D.C.D.: Global viewing of heterogeneous data sources. IEEE Transactions on Knowledge and Data Engineering **13**(2) (2001)
16. Doan, A., Madhavan, J., Domingos, P., Halevy, A.: Learning to map between ontologies on the semantic web. In: The Eleventh International WWW Conference, ACM Press (2002) 662–673
17. Daille, B.: Conceptual structuring through term variations. In Bond, F., Korhonen, A., MacCarthy, D., Villacicencio, A., eds.: ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment. (2003) 9–16