The Knowledge Engineering Review, Vol. 18:4, 317–328. © 2004, Cambridge University Press
DOI: 10.1017/S0269888904000128 Printed in the United Kingdom

Context matching for electronic marketplaces: a case study

MATTEO BONIFACIO[1,2], ANTONIA DONÀ[1], ALESSANDRA MOLANI[1,2] and LUCIANO SERAFINI[1]

[1]*ITC-IRST, Centro per la Ricerca Scientifica e Tecnologica, Panté di Povo, Trento, Italy.*
E-mail: antodona@itc.it, serafini@itc.it
[2]*Department of Information Technology, University of Trento, Panté di Povo, Trento, Italy.*
E-mail: bonifacio@dit.unitn.it, molani@dit.unitn.it

Abstract

Matching algorithms automatically discover semantic relations between two autonomously developed conceptual representations of two overlapping domains. Typical examples of such conceptualisations are electronic market catalogues (e.g., UNSPSC and ECL@SS) and Web directories (e.g., GOOGLE and YAHOO). The objective of this paper is the description of a use case in which the matching algorithm CTXMATCH, developed at ITC-IRST and the University of Trento, has been used to re-classify into the Universal Standard Products and Services Classification (UNSPSC) the catalogue of office equipment and accessories used by a worldwide telecommunication company to classify its suppliers. On the basis of this experience we are envisaging new applications of the algorithm in the area of demand aggregation. We conclude the paper by briefly describing a future application in this area.

1 Introduction

In e-business hype, marketplaces have been proposed as the optimal solution to foster efficiency and dynamic business integration, but reality is something different. The assumption that standard catalogues can substitute local ones is contradicted by the existence of different competing standards in similar domains (Agrawal & Srikant, 2001). Moreover, company buyers have difficulty in adopting classification standards that are too complex and generic with respect to those developed in house, which are simpler and more task specific. This is more true when considering the fact that local conceptualisations are not just the result of cultural and historical differences, but a consequence of different ways of doing things. Furthermore, the idea of standardising semantics seems to be conceptually wrong – more than practically unfeasible – if semantic heterogeneity is read in terms of richness to be exploited rather than in terms of noise to be reduced (Bonifacio & Molani, 2003). In very simple words, if people name things in different ways it is because they do different things, have different goals and adopt different perspectives.

In this scenario, the only feasible solution to heterogeneity in e-business is one that admits the existence of a set of heterogeneous and overlapping product catalogues, and supports semantic interoperability between them. Semantic interoperability is obtained through *semantic matching algorithms*, i.e. procedures capable of finding semantic relations between categories of different product catalogues. Two examples of matching algorithms, representing two different approaches, are the CTXMATCH algorithm described in Bouquet *et al.* (2003b) and the GoldenBullet system described in Ding *et al.* (2002). The former is based on Natural Languages Processing (NLP) techniques applied to labels occurring in the classification and the transformation of the matching

problem in a satisfiability problem, while the latter uses techniques for information retrieval and machine learning applied to the content of the classification. While GoldenBullet requires a training set, i.e. a set of examples of mappings between concepts which have been manually checked, CTXMATCH is completely automatic and needs just two classifications as input.

The objective of this paper is to describe an experimental application of the CTXMATCH algorithm to a case where no training set was available. In particular, we applied CTXMATCH to find mappings between a catalogue of office equipment and accessories used to classify company suppliers of a worldwide telecommunication company, and the standard catalogue UNSPSC. In this sense we simulated a situation in which a network actor of a marketplace becomes able to share information about products and services with other actors, without adopting a predefined ontology. This paper can also be consider as a partial answer of the call for a proposal described in Schulten *et al.* (2001), where the authors suggest the following challenge to e-business ''[. . .] *to come up with a generic model and working solution that can semiautomatically map a given product description between two different e-commerce product classification standards*''.

This paper is structured as follows. In Section 2 we describe the CTXMATCH algorithm; Section 3 describes the problem and the solution we have proposed in the use case; Section 4 describes some related work; and Section 5 draws some conclusions and describes future applications.

2 CtxMatch

CTXMATCH enables semantic interoperability between overlapping concept hierarchies through mapping discovery. The algorithm, as described in Bouquet *et al.* (2003b), takes as input two conceptual hierarchies (i.e. a source hierarchy and a target hierarchy) and returns a set of directed mappings between source and target concepts. The main features of the algorithm are the following: it does not consider concept instances (e.g., documents), so that it can be used in situations where such information is partially available or is not available at all; it returns a semantic evaluation of the mapping between two concepts (i.e. *equivalence*, *more general than*, *less general than*); it is context-based, in the sense that it builds a semantic representation of the meaning of a concept which depends both on the position in which it appears in a concept hierarchy and on world knowledge available in an external resource.

The algorithm performs three main steps: (i) a linguistic analysis of the labels without considering the context hierarchical structure (Magnini *et al.* 2003); (ii) a logical interpretation of the label based on the structural relations between the labels of the context (Bouquet *et al.* 2003a); (iii) the identification of mapping relations between the logical interpretations of the labels using a propositional decision procedure for satisfiability (SAT) (Bouquet *et al.* 2003a).

2.1 Linguistic analysis

The first step of the procedure consists of text chunking, i.e. dividing each label into syntactically correlated parts of words. We run the standard Alembic chunker (Day & Vilain, 2000), developed by MITRE Corporation as part of the Alembic extraction system (Aberdeen *et al.* 1995).

For example, with the label *Globalisation and Free Trade*, the chunker first selects a part of speech for each word (''Globalisation'' and ''Trade'' are nouns, ''Free'' is an adjective, ''and'' is a conjunction); then, it identifies two noun groups (NGs), i.e. ''GLOBALISATION'' and ''Free TRADE'' (notice that the syntactic head is marked in small capitals), and a coordinating conjunction between them:

$$[(\text{GLOBALISATION})_{nn}]_{\text{NG}}\ \text{and}_{cc}\ [(\text{Free})_{jj}(\text{TRADE})_{nn}]_{\text{NG}}$$

The output of the chunker is used to transform each label into a basic logical form. A noun group consisting of more than one word is interpreted as the conjunction of the head and all its modifiers;

for instance, *Iron Trade* is interpreted as [Iron<&>Trade]. The relations between different noun groups are interpreted on the basis of the linguistic material connecting them: coordinating conjunctions and commas are interpreted as a disjunction (e.g., *Globalisation and Free Trade* is interpreted as [[Globalisation]<|>[Free<&>Trade]]), prepositions, like ''in'' or ''of'', are interpreted as a conjunction (e.g., *Iron Trade of Great Britain* is transformed into [[Iron<&>Trade]<&>[Great<&>Britain]]), expressions denoting exclusion, like ''except'' or ''but not'', are interpreted as a negation (e.g., *Garments except Skirts* becomes [[Garments]<&><¬>[Skirts]])[1].

In order to perform the semantic interpretation of the labels, CTXMATCH accesses WORDNET, a freely distributable set of dictionary and thesaurus data compiled by Princeton University (Fellbaum, 1998). When a word is found, all the senses of that word are selected and attached to the basic logical form. When two or more words in a label are contained in WORDNET as a single expression (i.e. a multiword), the corresponding senses are selected and, in the basic logical form, the intersection between the two words is substituted by the multiword. In the case of [[iron*<&>trade*<&>[great*<&>Britain*]], for instance, ''Great Britain'' is provided in WORDNET as a single expression, so the logical interpretation is substituted by the senses of the multiword, thus obtaining [[iron*<&>trade*]<&>[Great_Britain*]].

2.2 Logical interpretation

The full logical form of a label is the conjunction of the basic logical forms of the label and all its ancestors. As an example, take the concept hierarchy whose root is *Soccer*, with a descendant *Leagues* and a further descendant *Clubs*. The full logical form of the root is simply [soccer*], the full logical form of *Leagues* is [[soccer*]<&>[league*]] and the full logical form of *Clubs* is [[soccer*]<&>[league*]<&>[club*]].

As explained before, the disjunction between noun groups can be made explicit by the presence of a coordinating conjunction, but we can also have implicit disjunction between elements placed at different levels of the hierarchy. In the example above, at a deeper level of analysis, there are two conflicting interpretations: from the point of view of the hierarchical structure, *clubs* denotes a subset of *leagues*; on the other hand, from the point of view of the world knowledge provided in WORDNET, [club#2] and [league#1] are disjoint because they have the same hypernym, i.e. association#1. In order to combine the two information sources, *leagues* has to be reinterpreted as if it were *leagues and clubs*, i.e. [[league#1]<|>[club#2]].

Similarly, also the negation is not always marked by expressions like ''but not'' or ''except''. For instance, we can have *Sociology* and *Science* as sibling nodes classified under *Academic Study of Soccer*. From the point of view of world knowledge, sociology is a science (and in fact in WORDNET sociology#1 is a second level hyponym of science#2). As a consequence, the node labelled with *Science* has to be interpreted as if it were *Science except Sociology*.

The recognition of multiwords can also be performed on different contiguous levels. For instance, in WORDNET there is a multiword ''billiard player'', so in a hierarchy where *Sport* has *Billiards* as a child and *Player* as a further descendant, the conjunction of [billiard*] and [player*] can be substituted with the multiword, giving as a result the logical form [[sport*]<&>[billiard_player*]].

CTXMATCH performs word sense disambiguation by taking into consideration both structural relations between labels and conceptual relations between words belonging to different labels.

Let L be a generic label and L_1 either an ancestor label or a descendant label of L and let s^* and s_1^* be respectively the sets of WORDNET senses of a word in L and a word in L_1. If one of the senses

[1] We use the following notation: ''Trade'' indicates a simple word; *Trade* indicates a label of a concept; Trade indicates a predicate in a logical form; trade* indicates the disjunction of all the senses of ''trade'' in WORDNET; trade#3 indicates sense 3 of ''trade'', while trade#[2,4] indicates the disjunction of sense 2 and sense 4.

Table 1

Level 1	**Furniture and Furnishings**
Level 2	**Accommodation furniture**
Level 3	**Furniture**
Level 4	**Stands**
Level 4	**Sofas**
Level 4	**Coat racks**

belonging to s* is either a synonym, a hypernym, a holonym, a hyponym or a meronym of one of the senses belonging to s_1^*, these two senses are retained and all other senses are discarded.

As an example, imagine *Apple* (which can denote either a tree or a fruit) and *Food* as its ancestor; since there exists a hyponymy relation between apple#1 (denoting a fruit) and food#1, we retain apple#1 and discard apple#2.

2.3 Computing concepts relations via SAT

In the first two steps, CtxMatch associates a formula $w(k)$ (expressed in a simple description logic) to each concept k of a hierarchy. This formula is supposed to capture the semantics of this concept. The last phase of CtxMatch addresses the problem of discovering the semantic relationship between two concepts k and k' by reducing it to the problem of checking, via SAT, a set of logical relations between the formulas $w(k)$ and $w(k'')$. The SAT problem is built in two steps. First, CtxMatch selects the portion T of the background theory relevant to $w(k)$ and $w(k')$, namely the WordNet relations involving the senses that appear in $w(k)$ and $w(k')$. In the second phase, it computes some of the logical relations between $w(k)$ and $w(k')$ which are implied by T.

The background theory $T(k,k')$ relevant for computing the relation between k and k' is obtained by translating the WordNet hierarchical relations on senses appearing in $w(k)$ and $w(k')$ into a set of subsumptions in description logic.

The equivalence between k and k' is checked by verifying that $w(k) \sqsubseteq w(k')$ and $w(k') \sqsubseteq w(k)$ are both implied by $T(k,k')$. Similarly, the fact that k is more specific [general] than k' is checked by verifying that $w(k) \sqsubseteq w(k')$ [$w(k') \sqsubseteq w(k)$] is implied by $T(k,k')$; the fact that k is compatible with k' is checked by verifying that $w(k) \sqcap w(k')$ is satisfiable in $T(k,k')$; finally the fact that k is disjoint from k' is checked by verifying that $w(k) \sqcap w(k')$ is not satisfiable in $T(k,k')$.

It is also possible to associate a quantitative measure to each semantic relation that considers the relation on the cardinality of models satisfying $w(k)$ and $w(k')$.

3 Use case product re-classification

In order to centrally manage all the company acquisition processes, the headquarters of a well-known worldwide telecommunications company had realised an e-procurement system[2], which all the company branch-quarters were required to join. In order to join it, each single office was also required to migrate from the product catalogue they used to manage to the new one managed within the platform. This catalogue is extracted from the Universal Standard Products and Services Classification (UNSPSC), which is an open global coding system that classifies products and services. The UNSPSC is used extensively around the world in electronic catalogues, search engines, procurement application systems and accounting systems. UNSPSC is a four-level hierarchical classification; an extract is reported in Table 1.

[2] An e-procurement system is a technological platform which supports a company in managing its procurement processes and, more in general, the re-organisation of the value chain on the supply side.

Table 2

Code	Description
ENT.21.13	cartridge *hp desk jet 2000c*
ENR.00.20	magnetic tape cassette *exatape 160m* × *l 7.0gb*
ESA.11.52	*hybrid roller pentel* red
EVM.00.40	safety scissors, length 25 cm

Table 3

	Item list	UNSPSC (44, 14)
No of concepts	194	272
Average label repetition	1.0	1.0
Average label length	5.5 words	3.8 words
WordNet's coverage	33.6%	21.2%
Average polysemy	2.3	2.3
No of multiwords	11	15

The Italian office asked us to apply the matching algorithm to re-classify into UNSPSC (version 5.0.2) the catalogue of office equipment and accessories used to classify company suppliers.

The items to be re-classified are mainly labelled with Italian phrases, but labels also contain abbreviations, acronyms, proper names, some English phrases and some typing errors. The English translation of an extract of this list is reported in Table 2. The italic parts were contained in the original labels. The item list was matched with two UNSPSC's-segments, namely: *Office Equipment and Accessories and Supplies* (segment 44) and *Paper Materials and Products* (segment 14).

The linguistic analysis of the labels is reported in Table 3.

3.1 Methodology

We started with the linguistic analysis (normalisation phase) of the company item catalogue and the UNSPSC segments that we took into account. The linguistic analysis first involves a morphological cleaning process, then a disambiguation and enrichment process. This is performed by accessing WordNet and, for each given term, finding out all the instances of its semantic meanings (corresponding to WordNet numeric IDs) and associating to them all the available synonyms. The output was two files with the semantic explicitations of the catalogue items on the one hand and the UNSPSC nodes on the other, both in terms of IDs of WordNet. Then we went on with the matching phase by running the algorithm on the two files. The result of the matching can be clearly interpreted in terms of re-classification: if the algorithm returns that the item i is equivalent to, or more specific than, the node c_{UNSPSC} of UNSPSC, then i can be classified under c_{UNSPSC} of UNSPSC.

Notice that the company item catalogue we had to deal with was a plain list of items, each identified with a numerical code composed of two numbers, the first referring to a set of more general categories. For example, the number 21 at the beginning of *21.13-cartridge hp desk jet 2000c* corresponds to *printer tapes, cartridge and toner*. We first normalised and matched the plain list against UNSPSC. This did not lead us to a satisfactory result. The algorithm performed much better when we made the hierarchical classification contained in the item codes explicit. This was done by

Table 4

	Baseline classification	
Total items	194	100%
Rightly classified	75	39%
Wrongly classified	92	47%
Non classified	27	14%

Table 5

	Matching classification	
Total items	194	100%
Rightly classified	136	70%
Wrongly classified	16	8%
Non classified	42	22%

substituting the first numerical code of each item with their textual description provided by experts of the company. The validation phase of our results was made by comparing them with the results of a simple keyword-based algorithm. Obviously, in order to establish the correctness of results in terms of precision and recall we have to compare them with a correct and complete matching list. Not having such a list, we asked a domain expert, Alessandro Cederle, Managing Director of Kompass Italia[3], to manually validate them.

3.2 Results

This section presents the results of the re-classification phase. Consider first the baseline matching process. The baseline was performed by a simple keyword-based matching that worked according to the following rule:

> for each item description (made up of one or more words) return the set of nodes, and their paths, which maximise the occurrences of the item words

Table 4 summarises the results of the baseline matching.

Given the 193 items to be re-classified, the baseline process found 1945 possible nodes in UNSPSC. This means that for each item the baseline found an average of six possible classifications. What is crucial is that only 75 out of the 1945 proposed nodes are correct. The baseline, being a simple string matching, is able to capture a certain number of re-classifications. However, the percentage of error is quite high (47%) with respect to that of correctness (39%). The results of the matching algorithm are reported in Table 5. In this case, the percentage of success is sensibly higher (70%) and, even more relevant, the percentage of error is minimal (8%)[4] This is also confirmed by the values of precision and recall, computed with respect to the validated list shown in Table 6.

[3] Kompass (`www.kompass.com`) is a company which provides product information, contacts and other information about 1.8 million companies worldwide. All companies are classified under the Kompass Product Classification with more than 52,000 products and services.

[4] Notice that the algorithm did not take into account just the UNSPSC level 4 category, since in some cases catalogues items can be matched with UNSPSC level 3 category nodes.

Table 6

	Total matches	Precision	Recall
Baseline	1945	4%	39%
Matching	641	21%	70%

The baseline precision level is quite small, while the matching one is not excellent, but definitely better. The same observations can also be made for the recall values.

Table 7 reports some examples where the algorithm found a correct item for re-classification, while the baseline did not[5]. The translation of Table 7 is reported in Table 8.

From the linguistical point of view, two considerations are needed. First, the possibility for our algorithm to manage with synonyms allows it to recognise that the item *perforatrice*[6], and all its variants (*perforatrice 2–4 fori*, *perforatrice universale*, etc.) have the same meaning of the UNSPSC node *punzonatrice*. This suggests its re-classification under that node. Second, the fact that during the normalisation phase a stemming cleaning is performed on words allows the matching phase to deal with lemmas in their basic form, without any morphological modification (e.g., the singular or plural word). This means that in the example of the item *evidenziatore*[7] (singular form) the algorithm is able to suggest the matching with the UNSPSC node *evidenziatori*, which is the plural form.

From a structural point of view, the possibility of reasoning on topological properties allows the algorithm to point out a semantic relation such as *more general than* between the company catalogue item *nastro per stampante*, *toner*, *cartuccia*, *testina di stampa*[8] and the UNSPSC node *Cartucce d'inchiostro*[9]. In this case, the properties taken into account are two: the fact that the catalogue item has just one level higher, while the UNSPSC one has two, and the fact that the catalogue item is made up of several single items, while the UNSPSC has just one. These two considerations are enough to state that the former should be more general than the latter. If there is not enough structural information to infer a semantic relation, CTXMATCH returns a percentage, which is intended to represent the degree of compatibility between the two elements. Degree of compatibility is computed on the basis of linguistic co-occurrence measures. Examples of compatibility relations are contained in the last four rows of Table 7.

As far as the non classified items are concerned, notice the following.

- In some cases, the item to be re-classified was not correctly classified in the company catalogue. Therefore, CTXMATCH could not compute the relations with the node and its father node in the right way. For example, *ashtray* was classified under *tape dispenser* and *wrapping paper* was classified under *adhesive labels*.
- In other cases, semantic matching was not discovered due to a lack of domain knowledge. For instance, to match *paper for hp* with the UNSPSC class of printer paper it would have been necessary to know that *hp* stands for Hewlett Packard and that they are a company which produce printers.

[5] Since the whole matching work strongly depends on languages, we will present results in Italian, with literal English translations. We apologise if in some cases such a translation does not completely support the reader in the comprehension of the example.
[6] In English: drilling machine.
[7] In English: highlight.
[8] In English: printer tape, toner, cartridge, print head.
[9] In English: ink cartridges.

Table 7 Some examples of matching found by CTXMATCH and not found by the baseline. Items (depicted in columns) are matched against UNSPSC classes (depicted in rows). $\subseteq$ stands for ''more general than'' and percentages represent the compatability degree

	perforatrice/perforatrice a 2–4 fori	nastro per stampante, toner, cartuccia, testina di stampa	penna lampostil, pennarello, evidenziatore/evidenziatore	penna a sfera, penna biro	decimetro doppio/decimetro doppio in plastica bianco	kit pulizia/kit pulizia PC	kit pulizia/kit pulizia testine nastro exatape 4 mm
Macchine da ufficio, materiali e accessori/							
Forniture per ufficio/							
Forniture per scrivanie/	$\subseteq$						
Punzonatrici per carta							
Macchine da ufficio, materiali e accessori/							
Forniture di stampanti, telecopiatrici e copiatrici/		$\subseteq$					
Cartucce d'inchiostro							
Macchine da ufficio, materiali e accessori/							
Forniture di stampanti, telecopiatrici e copiatrici/		$\subseteq$					
Toner							
Forniture per ufficio/							
Strumenti per scrittura/			$\subseteq$				
Evidenziatori							
Forniture per ufficio/							
Strumenti per scrittura/				70%			
Assortimento di penne e matite							
Accessori per l'ufficio e la scrivania/							
Accessori per disegno					71%		
Macchine da ufficio, materiali e accessori/							
Accessori per maccine d'ufficio/						80%	
Kit per pulitura di computer							
Materiali per ufficio/							
Macchine da ufficio, materiali e accessori/							
Acessori per macchine d'ufficio/							72%
Pulitori di nastri							

In a further experiment, we ran CTXMATCH between the company catalogue (in Italian) and the English version of UNSPSC. This was possible because the matching is computed on the basis of the WORDNET sense IDs and, in the version of WORDNET we used, the WORDNET sense IDs of Italian

Table 8 For the sake of comprehension, we report the English translation of Table 7

	drill/drill with 2–4 holes	printing tape, toner, cartridge, printing head	lampostil pen, marker, highlighter/highlighter	ball-point pen, pen	double ruler/double ruler in white plastic	cleaning kit/PC cleaning kit	cleaning kit/cleaning kit for exatape 4 mm tape heads
Office, materials and accessories/							
Supplies for office/							
Supplies for writing-desks/	⊆						
Paper staplers							
Office machines, materials and accessories/							
Supplies of printing, telecopying-machine and copying-machine/		⊆					
Ink cartridges							
Office machines, materials and accessories							
Printing furniture, telecopying-machine and copying-machine/		⊆					
Toner							
Supplies for office/							
Tools for writing/			⊆				
Highlighters							
Supplies for office/							
Tools for writing/				70%			
Assortment of pens and pencils							
Accessories for the office and the writing desk/							
Accessories for drawing					71%		
Office machines, materials and accessories/							
Accessories for office machines/						80%	
Computer cleaning kit							
Materials for office/							
Office machines, materials and accessories/							
Accessories for office machines/							72%
Cleaners of tapes							

and English words are aligned (i.e. the WordNet sense ID associated to the word and its translation in the other language is the same). This experiment allowed us to find more semantic matches.

More in general, this method allows us to approach and manage multilanguage environments and to exploit the richness which typically characterises the English version of linguistic resources[10].

Two lessons were learned from this experiment. First, CTXMATCH is good for *re*-classification rather than for classification from scratch of an unstructured set of items. As previously explained, the algorithm exploits two types of information: linguistic and structural. A plain set of items does not contain any structural information and this impacts the quality of the results. Performance definitely improved when we ran CTXMATCH on the company items catalogue enriched by the category structure extracted from the numerical codes.

The second lesson concerns the quality of the labels. The better the labels, the better the semantic matching. In the presence of meaningless labels such as shortcuts (''num'' for number, ''cart.'' for cartridge, etc.) or proper names (Duracell, Hewlett Packard, etc.), this version of the algorithm is not capable of assigning proper semantics to the labels. This negatively affects the performance. Possible improvements in this direction could be reached in two ways. First, the linguistic analysis of labels can be improved by using domain-oriented linguistical resources, by integrating CTXMATCH with thesaura, ontologies and standard classifications on specific domain lexicons[11]. These kinds of resources contain relevant information such as proper names, company names, abbreviations and acronyms. A second improvement could be obtained by running a spell checker that proposes automatic corrections on the labels.

4 Related work

An alternative approach to CTXMATCH for product classification in UNSPSC, called GoldenBullet, is described in Ding *et al.* (2002). GoldenBullet is an environment that supports product classification according to content standards. It applies techniques of information retrieval and machine learning. The classification is based on a training set. The classification algorithm implemented in GoldenBullet performs very well when used in a supervised way. This approach can be useful (and might perform better than CTXMATCH) in the presence of a representative set of pre-classified examples. CTXMATCH, however, provides good results without a training set.

A relevant approach to ontology matching was proposed in Doan *et al.* (2002) and Madhaven *et al.* (2002). Although the aim of the work – matching ontologies – is in many respects similar to our goals, the methodology differs significantly. The major difference is that their system builds mappings taking advantage of information contained in instances, while in the current version CTXMATCH completely ignores them. This makes CTXMATCH more appealing since most of the ontologies currently available on the Web still do not contain a significant set of instances. A second difference concerns the use of domain-dependent constraints; in the system developed in Doan *et al.* (2002) and Madhaven *et al.* (2002) these need to be provided manually by domain experts, while in CTXMATCH domain information is automatically extracted from an already existing resource (i.e. WORDNET). Finally, CTXMATCH attempts to provide a qualitative characterisation of the mapping in terms of the relation involved between two concepts, a feature which is not considered in the work of Doan *et al.* (2002) and Madhaven *et al.* (2002).

A mapping procedure based on lexical information was proposed in Bergamaschi *et al.* (2002). No quantitative evaluation is reported. Only a qualitative exemplification, based on the task proposed in Schulten *et al.* (2001), is described to show the algorithm capabilities.

Finally, the evaluation of the Anchor-PROMPT System (Noy & Musen 2001) was conducted on two ontologies and the mappings identified by the algorithm were manually checked. Results are presented in terms of the precision achieved.

[10] For instance, in the health-care domain, linguistic resources to be taken into account are MESH (`http://www.nlm.nih.gov/mesh/`) and the controlled vocabulary thesaurus of the National Library of Medicine.

[11] See `http://www.acronymfinder.com/` for an example of a database of acronyms.

5 Conclusions and further application

We have focused on the evaluation of CTXMATCH, a context matching algorithm, which automatically generates mappings among the concepts of two overlapping hierarchies. The main features of the algorithm are the following: it does not consider concept instances, therefore it can be used in situations where such information is partially available or not available at all; it returns a qualitative estimation of the mapping between two concepts (i.e. *equivalence*, *more general than*, *less general than*); it is content-based, in the sense that it builds a semantic representation of the meaning of a concept given both the context of its neighbourhood and the world knowledge available in an external resource (i.e. WORDNET).

We have presented three empirical experiments with a twofold aim: first, we wanted to evaluate the CTXMATCH algorithm in real, large-scale scenarios; second, we wanted to test different evaluation methodologies. In particular, we have experimented with CTXMATCH on Web directories and marketplace catalogues.

A number of data have been collected, which are to be considered as a first contribution towards common evaluation practices and the possibility of sharing resources for context matching algorithms.

Another application we are now working on is a system for managing the automatic aggregation of buyers' demands. The aim is to embed these in technological platforms (such as e-procurement systems or marketplaces) where the possibility for buyers to aggregate their product demand could give them some advantage in terms of supply conditions or buying power. In order to support the aggregation process, the system should be able first to point out groups of buyers interested in a similar category of product, then to suggest if and how each buyer should modify requested features in order to gain more advantages. As an example, consider the following case: the acquisition offices of two public universities are interested in buying 200 mobile phones, but one office prefers mobiles with the features A and B, and the other office is more interested in mobiles with the features C and D. Let us suppose that a mobile seller proposes a strong discount for 400 mobiles with features A, D and E. If the two offices converged on this last kind of mobile, they would get to the discount. In order to support this process, the system should be able not only to match items at the product description level, but also at the attribute level. Attributes are used to specify product features such as colour, length, size, etc., and typically they are objects of negotiations.

Acknowledgements

We thank Tiziana Favretto and Brian Martin for translations and proof checking. We also thank Alessandro Cederle for manually verifying the experimental results.

References

Aberdeen, J., Burger, J., Day, D., Hirschman, L., Robinson, P. & Vilain, M. 1995 Mitre: description of the Alembic system as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6), Columbia, MA, November 1995*.

Agrawal, R. & Srikant, R. 2001 On integrating catalogs. In *Proceedings of the Tenth International World Wide Web Conference (WWW-2001), Hong Kong, China, May 2001*.

Bergamaschi, S., Guerra, F. & Vincini, M. 2002 Product classification integration for e-commerce. In *Proceedings of WEBH-2002, Second International Workshop on Electronic Business Hubs, Aix En Provence, France, September 2002*.

Bonifacio, M. & Molani, A. 2003 The richness of diversity in knowledge creation: an interdisciplinary overview. In *I-KNOW '03 – 3rd International Conference on Knowledge Management, 2003*. `edamok.itc.it/dissemination/pep.htm`.

Bouquet, P., Magnini, B., Serafini, L. & Zanobini, S. 2003 A SAT-based algorithm for context matching. In Blackburn, P., Ghidini, C., Turner, R. M. & Giunchiglia, F. (eds.), *Proceedings of the 4th International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT-03) (Lecture Notes in Artificial Intelligence, Vol. 2680)*. Springer-Verlag.

Bouquet, P., Serafini, L. & Zanobini, S. 2003 Semantic coordination: a new approach and an application. In *Second International Semantic Web Conference (Lecture Notes in Computer Science, Vol. 2870)*. Springer-Verlag. pp. 130–145.

Day, D. S. & Vilain, M. B. 2000 Phrase parsing with rule sequence processors: an application to the shared CoNLL task. In *Proceedings of CoNLL-2000 and LLL-2000, Lisbon, Portugal, September 2000*.

Ding, Y., Korotkiy, M., Omelayenko, B., Kartseva, V., Zykov, V., Klein, M., Schulten, E. & Fensel, D. 2002 Goldenbullet: automated classification of product data in e-commerce. In *BIS-2002: 5th International Conference on Business Information Systems, 2002*.

Doan, A., Madhavan, J., Domingos, P. & Halevy, A. 2002 Learning to map between ontologies on the semantic Web. In *Proceedings of WWW-2002, 11th International World Wide Web Conference, May, Honolulu, Hawaii 2002*.

Christiane Fellbaum (ed.) 1998 *WordNet: An Electronic Lexical Database*. The MIT Press, Cambridge, MA.

Madhavan, J., Bernstein, P., Domingos, P. & Halevy, A. 2002 Representing and reasoning about mappings between domain models. In *Proceedings of AAAI-2002, Edmonton, Alberta, Canada, July–August 2002*.

Magnini, B., Serafini, L. & Speranza, M. 2003 Making explicit the semantics hidden in schema models. In Cappelli, A. & Turini, F. (eds.), *AI*IA 2003: Advances in Artificial Intelligence. Proceedings of the 8th Congress of the Italian Association for Artificial Intelligence, Pisa, Italy, September 2003 (Lecture Notes in Artificial Intelligence, Vol. 2829)*. Springer-Verlag, pp. 436–-448. Presented at the *Human Language Technology for the Semantic Web and Web Services, Workshop at ISWC 2003 International Semantic Web Conference*.

Noy, N. F. & Musen, M. A. 2001 Anchor-PROMPT: using non-local context for semantic matching. In *Proceedings of the IJCAI-2001 Workshop on Ontologies and Information Sharing, Seattle, WA, August 2001*.

Schulten, E., Akkermans, H., Botquin, G., Dšrr, M., Guarino, N., Lopes, N. & Sadeh, N. 2001 Call for participants: the e-commerce product classification challenge. *IEEE Intelligent Systems* **16**(4). The results of this experiment are not reported as they are not comparable with our simple keyword-based baseline, which makes no sense with multiple languages.