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
[2]Department of Information Technology, University of Trento
Panté di Povo, Trento Italy
{antodona,serafini}@itc.it
{bonifacio,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 conceptualizations 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 matching algorithm has been used to re-classify into UNSPSC the catalogue of the office equipment and accessories used by a worldwide telecommunication company to classify their suppliers. On the basis of this experience we are envisaging new application of the algorithm in the area of demand aggregation, and we will conclude the paper by briefly describing a future application in this area.

1 Introduction

In the e-Business hype, marketplaces have been proposed as the optimal solution to foster efficiency and dynamic business integration, however, reality have shown something different. The assumption that standard catalogues can substitute local ones has been neglected by the simple evidence that, on similar business domains, there are different competing standards [Agrawal and Srikant, 2001]. Moreover, company buyers have difficulties in adopting classification standards that are way complex and generic from their simple and task specific ones. This is more true when considering that local conceptualizations are not just the mere result of cultural/historical differences, but rather the consequence of different, substantial, valuable ways of doing things. Furthermore, the idea of standardizing 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 and Molani, 2003]. In very simple words, if people call things in different ways is because they do different things, have different goals, adopt different perspectives.

In this scenario the only feasible solution to heterogeneity in e-Business is the one that admits the existence of a set of heterogeneous and overlapping products catalogues, and supports semantic interoperability between then. Semantic interoperability is reached by matching algorithms, i.e., procedures capable to find semantic relations between the categories of different products catalogues. Two examples of matching algorithms, representing two different approaches, are the CTXMATCH algorithm described in [Bouquet *et al.*, 2003], and the GoldenBullet system described in [Ding *et al.*, 2002]. The former is based on NLP techniques applied to the 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 demands for a training set, i.e., it needs a set of examples of mappings between concepts which have been checked manually, CTXMATCH is completely automatic and needs as input just two classifications.

The objective of this paper is the description of an experiment of applying the CTXMATCH algorithm in a case where no training set was available. In particular we apply CTXMATCH to find mapping between the 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 be consider also as a partial answer of the call for proposal described in [Schulten *et al.*, 2001], where the author suggest the following challenge in the 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"*.

The 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 use case. Section 4 draw some conclusions and describe a future application.

2 CTXMATCH Algoritm

CTXMATCH enables semantic interoperability between overlapping concept hierarchies through mapping discovery. The algorithm [Magnini *et al.*, 2002] 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 concepts without considering the hierarchical structure of the context; (ii) a logical interpretation of the concept based on the structural relations of the context; (iii) the identification of mapping relations between the logical interpretations of the concepts using SAT.

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 and Vilain, 2000], developed by MITRE Corporation as part of the Alembic extraction system [Aberdeen *et al.*, 1995].

For example, with the label *Globalization and Free Trade*, the chunker first selects a part of speech for each word ('Globalization' and 'Trade' are nouns, 'Free' is an adjective, 'and' is a conjunction); then, it identifies two noun groups (NGs), i.e. 'GLOBALIZATION' and 'Free TRADE' (notice that the syntactic head is marked in small capitals), and a coordinating conjunction between them:
$[(\text{GLOBALIZATION})_{nn}]_{NG}$ $(\text{and})_{cc}$ $[(\text{Free})_{jj}(\text{TRADE})_{nn}]_{NG}$

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. *Globalization and Free Trade* is interpreted as [[Globalization]< | >[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 [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. To make an example, let's 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 labeled 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 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 semantic of this concepts. 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, it 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, we compute some of the logical relations between $w(k)$ and $w(k')$ which are implied by T.

[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 senses 2 and sense 4.

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 $T(k,k')$.

To each relation it is possible to associate also a quantitative measure 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 headquarter of a well known world wide telecommunication company had realized an eProcurement system [2], which all the company branch-quarters have been required to join. In order to join it, each single office was also required to migrate from the product catalogue they used to manage with, 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 the electronic catalogues, search engines, procurement application systems and accounting systems. UNSPSC is a four level hierarchical classification; an extract is reported in the following table:

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

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

The items to be re-classified are mainly labeled with Italian phrases, but labels contain also abbreviations, acronyms, proper names, some English phrases and some typing errors. The English translation of an extract of this list is reported in the following table (the italic parts were contained in the original labels).

Code	Description
ENT.21.13	cartridge *hp desk jet 2000c*
ENR.00.20	magnetic tape cassette *exatape 160m xl 7,0gb*
ESA.11.52	*hybrid roller pentel* red
EVM.00.40	safety scissors, length 25 cm

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).

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

The linguistic analysis of the labels is reported in the following table:

	Item list	UNSPSC (44,14)
# 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
# Multiwords	11	15

3.1 Methodology

We started with the linguistic analysis (normalization phase) of the company item catalogue and of the UNSPSC segments we took into account. The linguistic analysis inovolves first 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 associate to them all the available synonyms. The output was two files with the semantic explicitations of the catalogue items on one hand, and of the UNSPSC nodes on the other, both in terms of IDs of WordNet. Then we went on with the matching phase by running the alghoritm on the two files. The result of the matching can be cleary 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 made up of two couple of numbers the first referring to a set of more general categories (for example, in 21.13 - cartridge hp desk jet 2000c- 21 corresponds to *printer tapes, cartridge and toner*). We first normalized and matched against UNSPSC such plain list. This did not lead us to a satisfactory result. The algorithm performed much better when we made explicit the hierarchical classification contained in the item codes. This has been done by substituting the items first numerical codes with their textual description, provided us by the company. The validation phase of our results has been made by comparing them with the results of a simple keyword based alghoritm. Obviously, in order to set the correctness, in terms of precision and recall, of such results we needed a correct matching list to be used as point of reference for the validation. Then we ask a domain expert, Alessandro Cederle, Managing Director of Kompass Italia[3] to validate a possible correct matching list we provided him with.

3.2 Results

This sections presents the results of the re-classification. Consider first the baseline macthing process. The baseline has been performed by a simple keyword based matching which worked according to the following rule: for each item description (made up of one or more words) gives back the set of nodes, and their paths, which maximize the occurrences of the item words.

The following tables summarizes the results for baseline matching:

[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.

	Baseline classification	
Total items	194	100%
Rightly classified	75	39%
Wrongly classified	91	50%
Non classified	27	14%

Given the 194 items to be re-classified, the baseline process found 1945 possible nodes, that means that for each item it found a set of 6 possible maximazing nodes by average. What is crucial is that only 75 out of the 1945 proposed nodes are correct. The baseline, being a mere simple string matching, is able to capture a certain number of re-classifications, but the percentage of error is quite high (50%), with respect to the one of correctness (39%). Such parameter shows different values for the matching algorithm, that is able not only to compare strings but also to interpret their meaning. The results of the matching algorithm are reported in the following table:

	Matching classification	
Total items	194	100%
Rightly classified	134	70%
Wrongly classified	16	8%
Non classified	42	22%

In this case the percentage of success is sensibly higher (70%) and, even most relevant, the percentage of error is minimal (8%). [4] This is confirmed also by the values of precision and recall, computed with respect to the validated list:

	Founded Match	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 be done also for the recall values.

Table 1 reports some examples where the algorithm found out a correct item for re-classification, while the baseline did not[5].

The ability of the alghoritm to reason both on linguistical data and on structural ones, accounts for its good performance with respect to the baseline one. ¿From the linguistical point of view, two considerations are to be done. First the possibility for our algorithm to manage with synonimous allows it, for example, to recognize 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* so it suggested to re-classify it under that node. Second, the fact that during the normalization 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 the structural point of view, the possibility to reason on some topologichal properties allows the alghoritm to point out a semantic relation such as *More General than* between the company catalogue item *nastro per stampante, toner, cartuccia, testina di stampa* and the UNSPSC node *Cartucce d'inchiostro*. 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 one just of one. These two consideration are enough to set that the meaning of the first should be more general then the second's. If there are not enough structural data to suggest a semantic relation, the alghoritm gives back at least a compatibility between the two elements, computed on some linguistic occurrences measures (see the last four lines of the table). For the Non Classified items, it should be noticed the following:

- in some cases the item to be re-classified has a incorrect position within the company catalogue, so the matching alghoritm couldn't compute in the right way the relations with the node and its father node. Examples are the *ashtray* which has been classified under *tape dispenser*, the *wrapping paper* which has been classified under *adhesive labels*.
- in some cases in order to understand the meaning of the item to be re-classified, more domain knowledge should be requested and then embedded within the system. An example is the case of *paper for hp*: in order to understand that it is printer paper, it's necessary to know that hp stands for Helwett Packard and that this is a company which produces printers.

A final step of the use case was having a second matching between the company catalogue's IDs and the output of the normalization of the English version of UNSPSC with the English version of WordNet. This procedure, viable because the matching computation performing on the matrix takes into account just the concepts' IDs, allows us to find many more matching than using just one language.

More in general, this way allows us to approach and manage multilanguage environments and to exploit the richness which typically charaterizes the English version of any linguistic resources. [8]

Two lessons have been learned from this experiment. First, our algorithm is good for *re*-classification rather than for simple classification of a plain list of items. As previuosly explained, the alghoritm exploits two kinds of data, linguistical and structural ones: in the case of a plain list, the second set of data are missed and this impacts on the results' goodness. This is the reason why results definitely improve when we run the alghoritm 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 are written, the better is the matching obtained. In presence of meaningless labels such as shortcuts ("num" for number, "cart." for cartridge, etc.), proper names (Duracell, Hewlett Packard, etc.), this version of the

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

[5] Since the whole matching work strongly depends on languages, we will present results in Italian but provided with a very literally English translation. We apologise if in some cases such translation will not completely support the reader in the comprehension of the example.

[6] In English: *drilling machine*.

[7] In English: *highlight*.

[8] We do not report here the results of this last step, since we cannot compare them with the results of the baseline that, being a mere keyword based algorithm, could run only on homogeneus linguistical situation.

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

Table 1: Reclassifications found by CTXMATCH and not found by the baseline ($\supseteq$ stands for "more general than", and "c" stands for "compatible")

algorithm is not capable of assigning a proper semantic to the labels, decreasing the performances. Possible improvements in this direction could be reached through two ways. First, the linguistic analysis of labels can be improved by using domain-oriented linguistical resources. Several domains are developing thesaura, ontologies, standard classifications which specifically deal with their lexicons. [9] These kind of resources involve also relevant proper names, company names, abbreviations and acronyms, which right now are still problematic data for the algorithm.[10] The second way to be investigated is the possibility of supporting funcionalities of spell checking, able to detect spelling errors and to suggest the right alternatives.

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 support 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 indeed very well when it is used in a supervised way. This approach can therefore be useful (and maybe it could perform better than CTXMATCH) in presence of a representative set of pre-classified examples. CTXMATCH, instead provides good results also without such a training set.

A relevant approach to ontology matching has been proposed in [Doan *et al.*, 2002] and [Madhavan *et al.*, 2002]. Althought the aim of the work (i.e. establishing mappings among concepts of overlapping ontologies) is in many respects similar to our goals, the methodologies differ significantly. A major difference is that the GLUE system builds mappings taking advantage of information contained in instances, while our current version of the CTXMATCH algorithm completely ignores them. This makes CTXMATCH more appealing, since most of the ontologies currently available on the Semantic Web still do not contain significant amount of instances. A second difference concerns the use of domain-dependent constraints, which, in case of the GLUE system, need to be provided manually by domain experts, while in CTXMATCH they are automatically extracted from an already existing resource (i.e. WordNet). Finally, CTXMATCH attempts to provide a qualitative caracterization of the mapping in terms of the relation involved among two concepts, a feature which is not considered in GLUE. Although a strict comparison with the performances reported in [Doan *et al.*, 2002] is rather difficult, the accuracy achieved by CTXMATCH could be roughly compared with the accuracy of the GLUE module which uses less information (i.e. the "name learner").

A mapping procedure based on lexical information has been 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 and Musen, 2001] has been conducted on two ontologies and the mappings identified by the algorithm have been manually checked. Results are presented in term of the achieved precision.

5 Conclusions and Further Application

We focused on the evaluation of 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, so that 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 neighborhood 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

[9] for example in the Healthcare domain a good linguistical resources is MESH (http://www.nlm.nih.gov/mesh/), the National Library of Medicine's controlled vocabulary thesaurus. It consists of sets of terms naming descriptors in a hierarchical structure that permits searching at various levels of specificity.

[10] see http://www.acronymfinder.com/ for an example of acronym database.

to test different evaluation methodologies. In particular, we have experimented CTXMATCH on Web directories and market place catalogues.

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

Another application we are now working on is a system to manage the automatical aggegation of buyers' demands. This is aimed to be embedded in those technological platforms (such as eProcurements system, or also marketplaces) where the possibility for buyers to aggregate their product demands could give them some advantages in terms of forniture 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 some requested features in order to get to 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 preferes mobiles with the features A and B, the other office is more interested in mobile with the features C and D. Let's 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 item at the product description level, but also at the attribute level. Attributes are used to specify product features such as color, lenght, size, etc, and typically they are the elements on which a negotiation process could be done. The first idea we are investigating on is to develop different attribute contexts (one for colors, one for lenght, one for size,...), and a specific version of the alghoritm aimed to match this kind of structures. This way would allow us to split the matching between product structures, and the one between attribute structures and to combine afterwards the two sets of results.

A step further will be to develop the capability for the system to support buyers in the negotiation on attributes, for example by providing users with simulations of different combinations of attributes and displaying the related advantages.

References

[Aberdeen *et al.*, 1995] J. Aberdeen, J. Burger, D. Day, L. Hirschman, P. Robinson, and M. Vilain. Mitre: Description of the Alembic system as used for MUC-6. In *Proc. of the Sixth Message Understanding Conference (MUC-6)*, Columbia, Maryland, November 1995.

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

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

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

[Bouquet *et al.*, 2003] P. Bouquet, B. Magnini, L. Serafini, and S. Zanobini. A SAT–based algorithm for context matching. In P. Blackburn, C. Ghidini, R.M. Turner, and F. Giunchiglia, editors, *Proceedings of the 4th International and Interdisciplinary COnference on Modeling and Using Context (CONTEXT-03)*, volume 2680 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, 2003. To appear.

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

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

[Doan *et al.*, 2002] A. Doan, J. Madhavan, P. Domingos, and A. Halevy. Learning to map between ontologies on the semantic web. In *Proc. of WWW-2002, 11th International World Wide Web Conference*, May, Honolulu, Hawaii 2002.

[Fellbaum, 1998] Christiane Fellbaum, editor. *WordNet: An Electronic Lexical Database*. The MIT Press, Cambridge, US, 1998.

[Madhavan *et al.*, 2002] J. Madhavan, P. Bernstein, P. Domingos, and A. Halevy. Representing and reasoning about mappings between domain models. In *Proc. of AAAI-2002*, Edmonton, Alberta, Canada, July-August 2002.

[Magnini *et al.*, 2002] B. Magnini, L. Serafini, and M. Speranza. Linguistic based matching of local ontologies. In P. Bouquet, editor, *Working Notes of the AAAI-02 workshop on Meaning Negotiation. Edmonton (Canada)*, Edmonton, Alberta, Canada, July 2002. AAAI, AAAI Press.

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

[Schulten *et al.*, 2001] E. Schulten, H. Akkermans, G. Botquin, M. Drr, N. Guarino, N. Lopes, and N Sadeh. Call for participants: The e-commerce product classification challenge. *IEEE Intelligent Systems*, 16(4), 2001.