Use of Text Patterns for Evaluating Concepts in Corpora of Restricted Domain

Mireya Tovar[1], David Pinto[1], Azucena Montes[2], Gabriel Serna[3]

[1] Benemérita Universidad Autónoma de Puebla,
Faculty Computer Science, Puebla, Mexico

[2] Instituto Tecnológico de Tlalpan, TecNM, Mexico

[3] Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET),
Department of Computer Science, Mexico

{mtovar,dpinto}@cs.buap.mx, ing_tlalpan@tecnm.mx, gabriel@cenidet.edu.mx
http://www.lke.buap.mx/

Abstract. In this paper we present an approach for evaluating concepts in ontologies of restricted domain. The aim is to validate those concepts by discovering the linguistic patterns that have been used in the process of constructing the ontology. For this purpose, we have evaluated the concepts stored in three different ontologies from the following restricted domains: oil recovery, artificial intelligence and e-Learning. Two types of procedures were used for evaluating the ontological concepts: first, an intrinsic process using only the target ontology was carried out, whereas the second one procedure employed an extrinsic evaluation in which a reference corpus supports the validation process. The experimental results show a satisfactory performance when the approach proposed is executed in each ontology evaluated.

Keywords. Ontology domain, concept evaluation, linguistic patterns.

1 Introduction

An ontology can be defined as "an explicit formal specification about a shared conceptualization" [6]. In general, this type of semantic resources are made up of concepts or classes, relations, instances, attributes, axioms, restrictions, rules and events.

Nowadays, there are plenty of computational systems claiming to automatically generate ontologies, however, in the major of the cases, those systems lack of the particular process of automatic evaluation, therefore, the quality of the semantic resources generated is unknown.

The evaluation of ontologies task aims to measure the quality of such linguistic resources. The final aim is to facilitate the work of the ontological engineer or the domain expert when they verify the quality of ontologies with a considerable

high number of items inside it. This analysis is time consuming (hours/person). The evaluation process is not a trivial one, because it is necessary to chose the items of the ontology that should be considered when evaluating it, so as the particular criteria to employ in the evaluation (see [1]).

In this research work, we propose an automatic method based on Natural Language Processing (NLP) for the evaluation of concepts of an ontology of restricted domain. The proposed methodology assumes that the ontology has been constructed in an automatic, semi-automatic or manual way, and that the reference corpus (collection of documents) is semantically associated with the target ontology. The aim is to "validate", en a first stage, the quality of the ontological concepts. The evaluation of these concepts is carried out by means of two independent ways: 1) Using a reference corpus for the target ontology, and 2) Using the same ontology as training and test set, i.e., without any reference corpus.

The remainder of this paper is structured as follows. In Section 2 we present research works related with the automatic extraction of candidate terms. We also discuss the approaches for the automatic identification of restricted domain ontological concepts (see Section 3), including the construction and evaluation of morpho-syntactical patterns (see Section 4). Finally, conclusions and findings are given in Section 5.

2   Related Work

Most of the evaluation approaches for semantic resources in literature are focused on the evaluation of the ontology structure, assuming that the ontological concepts have been correctly defined by the ontological engineer. However, with the aim of provide a much more wide view of the evaluation process, we consider different approaches, even if some of those approaches are not so popular. The reviewed research works are firstly categorized in terms of the type of process employed when the ontology was constructed: automatic, semi-automatic or manual, and secondly, in terms of the practical creation purpose for the ontology. The major of the literature works may be categorized as follows [1]:

Human-Based Evaluation following Criteria, Standards and Requirements This type of evaluation allows to evaluate certain characteristics of the ontology, providing a numeric score [9]. Some features considered in this kind of evaluation are: Completeness, Correctness, Readability and flexibility [5,3].

Application-Based Evaluation This type of evaluation consists of testing the performance of the ontology in a given application. For example, answering user questions using an ontology [12,10].

Gold Standard-Based Evaluation In an evaluation based on Gold Standard, the quality of the ontology is expressed by means of the similarity of it

with respect to another ontology built manually, i.e., a gold standard ontology [10,11,15]. The comparison of both ontologies can be on two levels: lexical (similarity between concepts), and conceptual (similarity between relationships and taxonomies) [7]. In [4] different evaluation measures for lexical and semantic levels of ontologies are presented.

Reference Corpus-Based Evaluation In this case, the quality of the ontology is represented by the degree of a corpus topic covered by the ontology. For example, precision and recall metrics were used in [14] to evaluate the degree of lexical similarity of the ontology triplets with respect to elements extracted from the reference corpus. Furthermore, in [2] a probabilistic approach is used to compare the labels of an ontology with respect to a set of important terms identified in the reference corpus (extended by adding two levels of hyperonyms from WordNet).

We are interested in the automatic extraction of terms in a corpus domain from morpho-syntactical patterns identified in the words that form the concepts defined in the ontology. The morpho-syntactic patterns are built automatically from the words that integrated the concept using clustering. These patterns are used to extract the terms in the corpus. Next, we validate these terms with the concepts of the ontology.

3   The Proposed Approach for Evaluation of Ontological Concepts

In this section, we present an approach for the evaluation of restricted domain ontological concepts. First, we present the target ontologies to be evaluated (see Table 1), together with the metrics employed for the evaluation. Thereafter, we introduce an approach based on morpho-syntactical pattern for evaluating the concepts by means of a v-fold validation process, using the same ontology. These patterns are employed in another stage of the complete process for extracting and evaluating the candidate concepts by using a restricted domain corpus associated to the ontology.

In Table 1 we show the total number of concepts (C) and hierarchical relations (R) of three restricted domain ontologies that were evaluated following the proposed approach of this research work. In the same table is also shown the total number of documents (D) of each reference corpus, the number of tokens (T) and the vocabulary size (V). The restricted domains considered in this paper are: oil improved recovery methods (OIL), artificial intelligence (AI), and standard e-Learning SCORM (SCORM)[1] [16].

The evaluation of the proposed approach is carried out by means of metrics traditionally employed in information retrieval, such as precision (P), recall (R) and F-measure (F). Precision measures the proportion of candidate concepts that really belong to the target ontology between the number of terms identified as

---

[1] AI and SCORM ontologies are freely available at http://azouaq.athabascau.ca/

Table 1. The restricted domain ontologies and reference corpora employed in the experiments.

Dominio	Ontology		Corpora		
	C	R	D	T	V
OIL	48	37	575	9,727,092	188,047
AI	276	205	8	10,805	2,180
SCORM	1461	1038	36	32,644	2,154

candidate concepts by the system. Recall measures the proportion of candidate concepts identified by the system as ontological concepts between the number of all the real ontological concepts of the ontology. F-Measure is an harmonic measure that combines precision and recall.

The evaluation method for ontological concepts considers firstly the automatic identification of morpho-syntactical patterns employed by the ontological engineer (a human being or a computational system) in the original construction of the ontological concepts. Thus, we present the method used in the construction of linguistic patterns and the validation of it.

3.1 Construction of Morpho-Syntactical Patterns

The method employed for constructing the patterns and its use for extracting candidate terms, either from the ontology or from the reference corpus, is presented as follows:

1. To apply a Part-Of-Speech (PoS) tagger to the ontological concepts. In this case, we use TreeTagger [13].
2. To identify the morpho-syntactic tags for each ontological concept.
3. To cluster the morpho-syntactical PoS tags. We intercepted the morpho-syntactic PoS tags for forming clusters.
4. To construct regular expressions for the cluster of morpho-syntactical tags.
5. To extract candidate terms by applying the mentioned regular expressions to the reference corpus or ontology.

The results obtained by the aforementioned approach when it was applied to each ontology and corpus are shown in Tables 2, 3, and 4. The first column indicates the number of patterns identified in the ontology, the second column indicates the frequency of the pattern in the ontology, the third column shows the identified pattern itself, whereas the fourth column indicates the number of repeated terms in the corpus that match the pattern. The PoS tags are indicated by the following letters: P for preposition, A for adverb, N for noun, J for adjective, V for verb, C for number, F as a foreign word, and S is the symbol ".".

The most frequent morpho-syntactic pattern is the one starting with a noun, followed by nouns and adjectives (pattern 1). This result is expected because

the ontological concepts mostly will have this behavior. Actually, the number of multi-word expressions matching this pattern in the reference corpus is extremely high, thus they should be considered only as "candidate" concepts and filtered through some kind of term reduction technique.

Interestingly, some unexpected patterns such as pattern 10 of Table 3 have appeared in the AI ontology. The frequency of this pattern is 1 in the ontology, but it was not possible to find multi-word terms in the reference corpus. We suppose this problem is caused by the morphological tagger, because it does not have contextual information for correctly assigning the PoS tag to the concept. In order to avoid having incorrect morpho-syntactical patterns, this problem needs to be solved.

Table 2. Morpho-syntactical patterns identified in the OIL ontology.

N	Fr Ont	Pattern	Fr corpus
1	21	$N^+J?$	1,823,294
2	11	$(NV)((J?|(N^+)?)$	125,308
3	11	$J(N^+)?$	646,029
4	5	$A(N?|V?)$	223,301
5	4	$V(J|N)$	71,358
6	1	$P(N^+)$	192,879
		Candidate terms (without repetition):	378,465

Table 3. Morpho-syntactical patterns identified in the AI ontology.

N	Fr Ont	Pattern	Fr corpus
1	243	$(N^+)((VN)?|(V^+)?)$	2,693
2	84	$(J)^+(N^+)?$	1,000
3	35	$N((JN)|(PJN)|(C)|(PV)|(VJN)\ |(PN^+))$	400
4	17	$(V^+)(N^+)?$	1,582
5	10	$(AN)((PJN?)|(VN?)|(JN))$	135
6	6	$J((NPJN?)|(VN?))$	43
7	3	$A((VJN)|(JN))$	27
8	2	$(VP)(N|(JN))$	105
9	1	PJN	151
10	1	$FNPJN$	0
		Candidate terms (without repetition):	3,581

Table 4. Morpho-syntactical patterns identified in the SCORM ontology.

N	Fr Ont	Pattern	Fr corpus
1	9,238	$(N^+)((PJ^+(N^+?))?|(PJ^+N^+)?|(JN^+)?|$ $(VN^+)?|(J(N^+)?)?|(PN^+)?|(PV^+N^+)?|$ $(PNJ)?|(V^+)?)$	7,422
2	500	$V((JN^+)?|(PJN^+)?|(NPJN)?|(PJ^+)?|$ $(NPN^+)?|(NJN)?|(NPVN^+)?|(NPJ^+)?|$ $(J^+)?|(PJVN)?|(N^+V)?|(PV^+J)?|\ (PVN)?)$	5,560
3	327	$(J^+)((N^+PN^+)?|(N^+(PN^+)?)?)$	2,108
4	123	$V^+(PN^+|N^+)$	1,105
5	84	$(J)(V(PN)?|NVN|NPJ^+(N^+)?|NPNPJ$ $N|V^+N^+|NPVN|N^+J|V(PN)?|NPJVN$ $|NPNPN^+|N^+V)$	336
6	34	$(NP)(V|NPJN?|JNPNPN|JV^+|NVN^+$ $|NVJN|NPN^+|JN^+V|JVN^+|NCSCS\ CN^+)$	230
7	17	$(N)(PV|VJN?|JN^+PJ^+N^+|CJN|VPN^+$ $|CSCSCJ|JN^+V|VPVN)$	313
8	9	$(J)(JVN|V^+N|VJ)$	95
9	7	$(JV)((PN)?|(NJ)?)$	74
10	5	$(AJ)(VN|NV?)$	27
11	1	PJN	273
		Candidate terms (without repetition):	5,552

4 Evaluation of the Morpho-Syntactical Patterns

In a first phase of the evaluation of the ontological concepts, the regularity of the morpho-syntactical patterns employed by the ontological engineer (either being a human or a computer program) is considered. We evaluate the concepts using a v-fold cross validation process under the same target ontology. The explanation of this procedure together with the obtained results is given in the following section. Thereafter, we show how to validate the ontology by a second type of procedure which uses the morpho-syntactical patterns for findings concepts in a reference corpus with the aim of comparing those with the ones already stored in the ontology to be evaluated.

4.1 Intrinsic Evaluation of Ontological Patterns for Restricted Domain Ontologies

The intrinsic evaluation of ontological concepts considers only the target ontology, i.e., no other external resources such as a reference corpus is used. The procedure can be described as follows:

1. The concepts of the ontology to be evaluated are considered into two sets: training and test.
2. We identify regular morpho-syntactical patterns on the training set

3. The patterns identified are then used to validate the test set

We execute this procedure k times by changing the training and test set in a leave-one-out process, thus we execute a k-fold cross validation process [8] over the same ontology, in which $k-1$ sets are used as a training set, and the remaining one as the test set. The average of the obtained results in each iteration is given as a final result.

In a strict sense, the morpho-syntactical patterns should be employed for matching complete strings of words, however, in some cases, a sub-string may match with a particular pattern. For example, the concept "long term planning system" is already stored in the AI ontology, however, there are some patterns which permit to determine that the string "planning system" may be also a candidate concept, even if this multiword term is not stored in the target ontology.

In Table 5 we show the obtained results when this particular method of sub-string matching is employed. Precision (P), recall (R) and F-Measure (F) evaluate the quality of the ontological concepts (OC) stored in the ontology. As can be seen, we were able to find every concept of the ontology ($R = 1$) besider a number of new concepts (candidates to be concepts (CC)) that also match with the morpho-syntactical patterns discovered. From our particular point of view, this process may be interesting, since it will allow to suggest the inclusion of new concepts to the ontology. Taking into account the total number of candidate concepts found, we obtain an F-measure greater than 0.8 in all the ontologies evaluated.

Table 5. Results obtained with the intrinsic evaluation of ontological concepts.

Ontology	$|OC|$	$|CC|$	$|OC \cap CC|$	P	R	F
OIL	48	68	48	0.70588	1	0.82759
AI	276	346	276	0.79769	1	0.88746
SCORM	1461	1874	1461	0.77962	1	0.87616

Table 6 shows the results obtained by averaging the different v folds of cross-validation when it was applied to the different restricted domain ontologies evaluated in this research work. We have also evaluated the ontologies by taking into account the total number of candidate concepts found. The F-Measure obtained is about 70%, a value that indicates the degree of regularity when constructing the ontological concepts of the target ontologies. The obtained results show that the SCORM ontology presents a more stable construction schema than the other two ontologies evaluated, i.e., the SCORM ontology has been constructed employing a uniform set of morpho-syntactical patterns when the ontological concepts were identified by the ontological engineer.

The following research question arises when we obtain this kind of results: Are there universal morpho-syntactical linguistic structures employed by the ontological engineers when constructing restricted domain ontologies?. In order to answer this important question, we have executed an experiment in which

Table 6. V-fold cross-validation for intrinsic evaluation of ontologies.

Ontology	P	R	F
OIL	0.58881	0.84893	0.69103
AI	0.57950	0.92216	0.71113
SCORM	0.60909	0.96251	0.74575

the k-fold cross-validation process has been carried out among different domains, i.e., we have used one domain as a training set and another one as the test set (k-domain cross-validation). Table 7 presents the evaluation results obtained when we employed the morpho-syntactical patterns discovered using the ontology of the first column (Ont) for discovering candidate terms using the ontologies Ont_1 and Ont_2. Thus, in the first row the OIL ontology has been used as training set, whereas the other two ontologies (AI and SCORM) were used as test set. The second row shows the results obtained when the AI ontology is used as a training set and the other two ontologies are used as the test set (Ont_1=OIL and Ont_2=SCORM). Finally, in the third row we present the evaluation values obtained when the SCORM ontology is used as the training set, whereas Ont_1=OIL and Ont_2=AI are used as a test set.

As expected, the performance of the approach is lower than the previous experiments, because in this case we are dealing with a knowledge transfer process from a domain X to a different domain Y. Again, we observed that the concepts of the OIL ontology do not have homogeneous linguistic patterns, while the SCORM ontology seems to have a set of linguistic patterns much more generic than the other two ontologies. Since the F-Measure is not greater than 70% for all the evaluations, we consider that more investigation need to be done in future work with the aim of deeply understand the manner the ontological engineers employ the morpho-syntactical patterns when they construct ontologies of restricted domain.

Table 7. v-domain cross-validation evaluation results.

Ont	Ont$_1$			Ont$_2$		
	P	R	F	P	R	F
OIL	0.50795	0.69335	**0.58581**	0.43975	0.56912	**0.49594**
AI	0.53304	0.76214	**0.62455**	0.51982	0.82250	**0.63700**
SCORM	0.61797	0.96233	**0.75121**	0.57237	0.89571	**0.69452**

4.2 Extrinsic Ontological Concepts Evaluation (using a Reference Corpus)

The extrinsic evaluation is carried out when we have a corpus associated to the same domain of the target ontology to be evaluated (reference corpus). In

this case, we use the ontology as a source for identifying the morpho-syntactical patterns employed by the ontological engineer when constructed the ontological concepts. These patterns are then used for extracting candidate concepts from the reference corpus and compared them against the original concepts in order to evaluate the quality of the original ontological concepts.

The methodology employed in this case consists of the following steps:

1. To apply a Part-Of-Speech (PoS) tagger to the ontological concepts of the target ontology.
2. To identify the morpho-syntactic tags for each ontological concept.
3. To apply a Part-Of-Speech (PoS) tagger to the reference corpus.
4. To extract candidate concepts in the reference corpus by using the morpho-syntactic patterns previously discovered.
5. To compare the candidate concepts against the original ontological concepts and provide a measure of the ontology quality.

Table 8 shows the results obtained with this kind of ontology validation. We observe that by applying this methodology, we are able to discover 87.5% of the OIL ontological concepts, 74.27% of the AI ontological concepts, and 59% of the SCORM ontological concepts. The proposed approach allows then to evaluate the quality of the methodology when a reference corpus is available. Of course, the better the reference corpus (amount and quality of the documents), the higher the reliability of this procedure of evaluation.

Some patterns obtaines a extremely generic, thus generating in some cases huge numbers of candidate concepts (CC). It is then important to implement a candidate concept reduction schema that allows to improve the evaluation procedure, limiting the candidate concepts to those associated to the ontology domain. Again, the new concepts discovered may be used for suggesting to the ontological engineer their inclusion into the ontology.

Table 8. Extrinsic evaluation of ontological concepts.

Ontology	$|C|$	$|CC|$	Enc	P
OIL	48	364,033	42	0.87500
AI	276	3,581	205	0.74275
SCORM	1,461	5,552	864	0.59138

5   Conclusions

In this research work we have presented an approach based on morpho-syntactical patterns for evaluating ontological concepts stored in ontologies of restricted domain. Intrinsic and extrinsic evaluation procedures were carried out, showing that it is possible to evaluate the quality of the ontology when a reference corpus is available, but also when this collection of documents is not available.

By evaluating the target ontologies employing the intrinsic procedure we were able to analyze the regularity of the morpho-syntactical patterns employed by the ontological engineers when constructing the ontological concepts. According to the results obtained, the SCORM ontology is the one that has been constructed with more care, despite of the high number of ontological concepts that this ontology has, because the linguistic patterns discovered in this ontology are homogeneous and less specific than the ones used in the other two ontologies.

On the other hand, when we employed the extrinsic procedure, we were able to validate the original ontological concepts by using candidate concepts discovered in a reference corpus. The approach proposed obtained a minimum of 59% of accuracy and a maximum of 87.5%, whereas the recall was 100% in all the cases.

Even if the purpose of the method proposed is to validate ontological concepts stored in ontologies of restricted domain, it is also possible to use it for suggesting or recommending the inclusion of new concepts in the target ontology which we consider an important contribution of the experiment carried out.

One of the methods proposed for discrimination of candidate terms, as future work, is the identification of taxonomic relationships and not taxonomic relationships associated with these terms in the corpus.

The resuts has been applied to suport the semiautomatic build of ontological conceptual model. The next step is to integrate this approach in order to obtain relevant term extraction on specific domain.

References

1. Brank, J., Grobelnik, M., Mladenić, D.: A survey of ontology evaluation techniques. In: Proc. of 8th Int. multi-conf. Information Society. pp. 166–169 (2005)
2. Brewster, C., Alani, H., Dasmahapatra, S., Wilks, Y.: Data driven ontology evaluation. In: Proceedings of International Conference on Language Resources and Evaluation (2004)
3. Cantador, I., Ferández, M., Castells, P.: A collaborative recommendation framework for ontology evaluation and reuse. In: Actas de International Workshop on Recommender Systems, en la 17th European Conference on Artificial Intelligence (ECAI 2006), Riva del Garda, Italia. pp. 67–71 (2006)
4. Dellschaft, K., Staab, S.: Strategies for the evaluation of ontology learning. In: Buitelaar, P., Cimiano, P. (eds.) Bridging the Gap between Text and Knowledge Selected Contributions to Ontology Learning and Population from Text. IOS Press, Amstedam (2008)
5. Gómez-Pérez, A.: Ontology Evaluation. International Handbooks on Information Systems, Springer (2004)

6. Gruber, T.R.: Towards Principles for the Design of Ontologies Used for Knowledge Sharing. In: Guarino, N., Poli, R. (eds.) Formal Ontology in Conceptual Analysis and Knowledge Representation. Kluwer Academic Publishers, Deventer, The Netherlands (1993)

7. Maedche, A., Staab, S.: Measuring similarity between ontologies. In: Proceedings of European Knoeledge Ackquisition Workshop (EKAW) (2002)

8. Murty, M., Devi, V.: Pattern Recognition: An Algorithmic Approach. Undergraduate topics in computer science, Springer-Verlag London (2011)

9. Pak, J., Zhou, L.: A framework for ontology evaluation. In: Sharman, R., Rao, H.R., Raghu, T.S. (eds.) WEB. Lecture Notes in Business Information Processing, vol. 52, pp. 10–18. Springer (2009)

10. Reyes Ortiz, J.A.: Creación automática de Ontologías a partir de Textos con un Enfoque Lingüístico. Ph.D. thesis, Dept Ciencias Computacionales, Cenidet, Cuernavaca, Mor., Mex. (2013)

11. Sabou, M., Lopez, V., Motta, E., Uren, V.: Ontology selection: Ontology evaluation on the real semantic web. In: Proceedings The 4th International EON Workshop, Evaluation of Ontologies for the Web (2006)

12. Salem, S., AbdelRahman, S.: A multiple-domain ontology builder. In: Proceedings of the 23rd International Conference on Computational Linguistics. pp. 967–975. COLING '10, Association for Computational Linguistics, Stroudsburg, PA, USA (2010), `http://dl.acm.org/citation.cfm?id=1873781.1873890`

13. Schmid, H.: Probabilistic part-of-speech tagging using decision trees. In: International Conference on New Methods in Language Processing. pp. 44–49. Manchester, UK (1994)

14. Spyns, P., Reinberger, M.L.: Lexically evaluating ontology triples generated automatically from texts. In: Gómez-Pérez, A., Euzenat, J. (eds.) ESWC. Lecture Notes in Computer Science, vol. 3532, pp. 563–577. Springer (2005)

15. Zavitsanos, E., Paliouras, G., Vouros, G.A.: Gold standard evaluation of ontology learning methods through ontology transformation and alignment. IEEE Trans. Knowl. Data Eng. 23(11), 1635–1648 (2011)

16. Zouaq, A., Gasevic, D., Hatala, M.: Linguistic patterns for information extraction in ontocmaps. In: Blomqvist, E., Gangemi, A., Hammar, K., del Carmen Suárez-Figueroa, M. (eds.) WOP. CEUR Workshop Proceedings, vol. 929. CEUR-WS.org (2012)