

Multilingual Ontology Alignment Based on Visual Representations of Ontology Concepts

Srdan Mihic and Dragan Ivetic
 Computing and Control Department,
 Faculty of Technical Sciences, University of Novi Sad
 Novi Sad, Serbia
 e-mail: {smihic, ivetic}@uns.ac.rs

Abstract— Image search represents one of the most frequent user actions on the Internet. Existing image search engines do not understand the images they return, nor do they support multilingualism. These issues can be addressed with the introduction of a semantic layer. The semantics is encoded in ontologies, which contain structured information about a domain of application. In order to provide semantic interoperability between (multilingual) ontologies, it is necessary to obtain semantic correspondences – ontology alignments. Various strategies have been proposed for multilingual ontology alignment. In this concept paper, the idea of alignment discovery based on semantic similarity of visual representations of ontology concept is explored.

Keywords-multilingual ontologies; ontology alignment; image retrieval; multimedia semantics

I. INTRODUCTION

People often use Internet for querying images. Existing image search engines are syntax-based, and thus do not understand the images they return. The result sets are mostly large but lack precision. Namely, a good part of the result set is irrelevant to the formulated query. Introduction of a semantic layer in image retrieval improves the precision of results obtained [1].

Current image search engines have very limited support for multilingualism. Although they provide users with ability to narrow the region and/or language (used for image tagging/description) but that does not provide satisfactory results. In the following, some of the pressing issues will be presented.

Distribution of images in relation to languages is non-uniform on the Internet. Usually, the higher the presence of a language, the bigger result set is retrieved. For example, a query for 'шак' (Serbian for 'sack') produces no semantically valid results. Yet issuing equivalent query in English produces a vast number of semantically valid results with high precision.

In linguistics, homographs are group of words that share the same spelling but have different meanings, regardless of how they are pronounced. Homonyms are homographs that have the same pronunciation as well. Word in one language is often a homograph/homonym for an unrelated word in another language – a cross-lingual homograph/homonym. Therefore, it is possible for one word (in language with higher presence) to mask the other (in language with lower presence). For example, a query for 'fog' (Hungarian for 'tooth') yields in images of misty weather (because higher presence of English than Hungarian).

User queries are often composed of a few words (generally two or three words), and are too imprecise to express the query that the user had in mind [2]. Especially, it is hard to formulate proper queries in image search [2]. In addition, studies have shown that users tend to look only at the first answers pages [3]. Thus, it is necessary to obtain and rank the “right” answers first based on a short fuzzy description.

Images, that are relevant to the formulated query, are retrieved if a user queries in the “right” language. Thus, users have to issue queries in various natural languages in order to obtain satisfactory results. Not all users have necessary linguistic skills to adequately translate queries in a foreign language. Even translation tools fail to provide adequate translations. This results in imprecise translations that can lead to even poorer set of results. For example, the aforementioned term 'шак' can be translated as 'bag'. This translation is more imprecise than one with the term 'sack', but more common for users who do not know English language well. As expected, this translation yields no satisfactory results. Therefore, automatic inclusion of translation of terms in a semantically meaningful way would provide a richer set of retrieved images and would lead to an enhanced user experience.

By addressing the aforementioned issues users would be able to state queries in the language of their choice and to get the most appropriate image results regardless of the language used.

Ontologies represent an economic and efficient way to address aforementioned issues and to model semantic layer. Thus, in recent years, they have gained a large amount of attention and many have been developed and are available on-line.

With the expansion of ontologies in terms of application domains, the number of natural languages in which they were written grew. Thus, reasoning and mapping of these multilingual ontologies has become an important issue [4].

The process of linking related ontology elements is called ontology alignment (or mapping) [5][6]. Ontology alignment enables semantic interoperability between distributed information systems. The resulting alignments can be used for agent communication (interoperability between distributed information systems), query answering (executing query in all available natural languages), ontology merging, or for navigation on the Semantic Web [7].

There are many ontology alignment techniques (see [6] for an exhaustive review) and various multilingual ontology alignment strategies have been proposed (see Section V for detailed review). Common to all these solutions is that the

generation of alignments is based on comparison of (multilingual) ontology labels.

In addition to these approaches, we propose to use images as visual representations of ontology concepts for alignment discovery between two multilingual ontologies. This approach complements the aforementioned approaches.

The outline of this paper is as follows. In the following section, the image-based multilingual ontology alignment approach for building indirect mapping between multilingual ontologies is described as the main contribution of this work. In Section 3, the initial proposal for image-based alignment discovery is presented. In Section 4, some previous studies related to this work are introduced. Finally, Section 5 concludes this paper and presents further research directions.

II. IMAGE-BASED MULTILINGUAL ONTOLOGY ALIGNMENT

We draw our inspiration from the natural way the humans learn new languages. One can learn a foreign language visually by establishing pictorial inter-language mappings between visual representations of corresponding terms/concepts. These pictorial inter-language mappings have proven quite useful in a number of commercial language learning applications, as for instance Rosetta Stone [16], and therefore applying it more formally to ontology alignment might be a promising idea.

For example, let us consider a situation in which two speakers want to communicate with each other (Fig.1). The first speaker is from Serbia and speaks only Serbian, and the other one is from Japan and speaks only Japanese. Unfortunately, neither of them knows the language spoken by the other one, nor they speak the common language. If they want to communicate with each other, they will have to teach each other their respective languages. The most natural way to this is to use real life objects, more precisely their visual representations (images), and to exchange their labels (in Fig.1 using image of a dog the speakers learn its label in foreign language). This way, the speakers will most likely learn the most common and the most adequate word meaning.

In many natural languages, entities are described by nouns, which are, in majority, picturable entities [1]. The number of nouns in natural language is usually significantly higher than

the number of verbs, adjectives, and adverbs (e.g., 80% of the Serbian WordNet are nouns [9]). As building blocks of ontologies, concepts and their instances are described by nouns as well. Espinoza et al. [10] have empirically found that existing ontologies share the same lexical patterns. For example, approximately 60% of concept labels follow an adjective-noun pattern (e.g., *temporal region*), where as the others (about 30%) use the noun-noun pattern (e.g., *knowledge domain*) [10]. We limit our discussion to the above stated lexical patterns. Other lexical categories (e.g., verbs) are left for future research, since some of them can be represented by picturable entities as well.

Cognitive psychology studies have found that: i) there exists a correlation between visual and semantic similarity in the human visual system; ii) semantic categories are visually separable; iii) there exist visual prototypes for semantic categories [11]. More recently, Deselaers et al. [12] have experimentally confirmed that these conclusions hold in the field of computer vision. In addition, they have found that the visual variability within a category grows with its semantic domain.

There are plenty of images available on-line that can be used as visual representations of ontology concepts.

According to the aforementioned, visual representations of ontology concepts can be used and compared in order to find out the adequate mapping. Our idea is additionally supported by the fact that it is easy to cope with synonyms issues in visual domain since synonyms visual representations are similar or even the same (e.g., words *hound* and *dog* are synonyms and visually represent the same entity).

A proposed architecture for image-based multilingual ontology alignment is presented in the following section.

III. A SKETCH OF A POSSIBLE SYSTEM ARCHITECTURE

The proposed architecture is shown in Fig. 2. It is based on four main components: the Alignment Generator, the Visual Representations Provider, the Image Comparator, and the Alignment Repository.

The Alignment Generator receives two ontologies as input and generates alignments if possible. Firstly, for each concept pair of the matching ontologies, the component checks whether suitable alignment already exists in the Alignments Repository. If it does not, this component enquires the Visual Representations Provider to provide suitable visual representations (several images and their accompanied textual descriptions) of these concepts. If such representations can be found, they are compared using the Image Comparator component. The Image Comparator computes a degree in which these visual representations of concepts are related and chooses the best among these representations. Finally, the alignment is generated and stored in the Alignments Repository for sharing and reuse, along with the chosen visual representations.

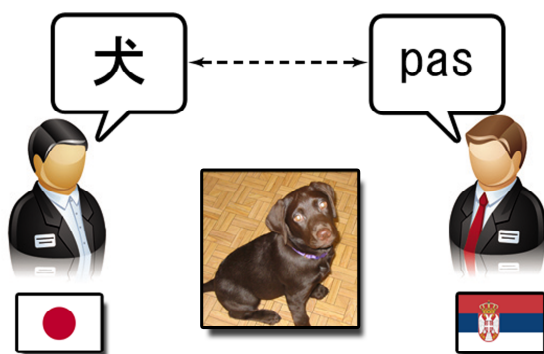


Figure 1. Natural way of learning terms of foreign language using its visual representation

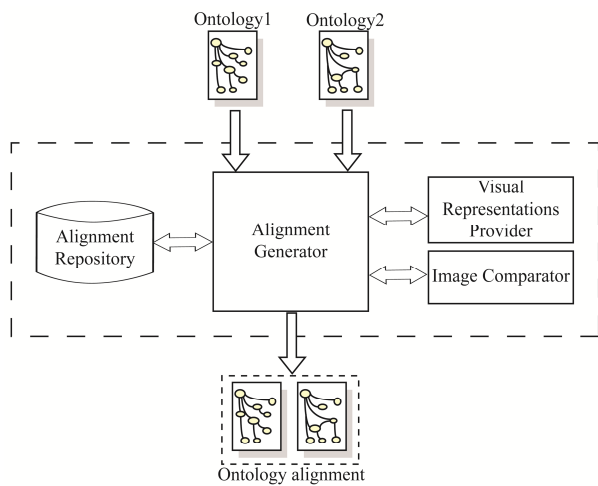


Figure 2. The proposed architecture for image-based multilingual ontology alignment

Later, the generated alignments can be employed for automatic query translation and distributed query answering.

The proposed system is in the early stages of development and provides guidelines for future work.

A. Acquisition of the visual representations of concepts

The process of multilingual ontology alignment discovery begins with a comparison of leaf concepts. The rationale for comparing leaf concepts first is that they often point towards specific entities, while concepts that are high in the hierarchy tend to represent more abstract and thus more ambiguous entities [1]. In addition, for those concepts the semantic domain is narrow, thus visual variability is small (see Section II).

For each concept pair, the Visual Representations Provider component tries to provide suitable visual representations of these concepts, as well the accompanying text (tags, annotations, etc.). The success of the entire process of image-based alignment discovery is highly dependent upon this step. The image comparison process is more reliable and precise if the acquired images are true semantic visual representatives of concepts.

Thus, the component should attempt to acquire images from semantically rich sources, if possible. The component attempts to find the source that supports queries stated in the natural languages of both ontologies first. If no such source can be found or yields no results, the component opts for two monolingual sources: one for each natural language. If those sources cannot be found or yield no results, the component opts for sources with less support for semantic search. Four types of sources have been identified according to the semantics they incorporate: ontology-based image retrieval systems, hierarchy-based image databases (usually WordNet [13] hierarchy-based), content-based image retrieval systems, and syntax-based image search engines. The sources are listed in descending order of the semantics they incorporate.

In situations where none of aforementioned steps obtain visual representations, the system marks that pair as incomparable due to lack of data and steps to the next pair.

When querying for images, the context of the ontology concept is used to disambiguate the lexical meaning of a

concept label, as in [10]. For example, let us consider an ambiguous concept label *crane*. The term *crane* can have two senses in English: *a bird* and *a type of construction equipment*. Thus, an image search with the term *crane* results in images both of birds and of construction equipment. By adding parent concept label/s (e.g., *bird*) to the query, the obtained images are more appropriate than those using the concept label alone.

B. Semantic-based image comparison

After obtaining visual representations, the Image Comparator component performs semantic-based comparison of two visual representation sets and selects the best visual representations. This component is the core of the system and represents the most complex part of the system. For the time being, this component is in early stages of development. We plan to develop it as a multimodal probabilistic framework, inspired by [14, 15].

C. Generation of alignments

When computing, the confidence value reliability of the source must be taken into account. Source reliability is a weighting factor ranging from 0 to 1 which is used to define the influence of a particular retrieval option on the final result. Generally speaking, ontology-based retrieval is assigned high values, and syntax-based low values due to the greater reliability of the former.

If the confidence value is below a predefined threshold t , the concepts are considered unrelated. Otherwise, an alignment with a calculated confidence value is generated. In addition, to support alignment reuse, the algorithm stores alignments and respective visual representations in a shared alignment repository, similar to [16].

IV. USAGE SCENARIO

One possible usage scenario would be to use the generated alignments in the Alignments Repository to support automatic query translation into several natural languages.

For example, a Serbian teacher gives an assignment to her/his pupils, still in elementary school, to write an essay about the culture of modern Japan for a sociology class. Since pupils are not fluent in English nor do they know Japanese, it is very hard for them to acquire materials (including images) using common image search engines. First, they must face the problem of query translation in English and/or in Japanese. They do either this manually or by using some (machine) translation tools. It is highly unlikely that this approach would lead to acceptable result set. Secondary, they have to manually issue queries in both languages and compare them manually.

When relying on our approach, pupils can issue queries in their native language without the need to know any ‘major’ language. The query is parsed and concepts and context are extracted. The Alignment Repository is queried for those concepts. If alignments can be found, the concepts are translated in their respective equivalents in different languages. Since the alignments store the image data, which are visual representations of those concepts, these images can be used either as results and/or to support query-by-semantic-example [17] queries. The image search engines execute translated queries. The results are aggregated and presented to the user.

V. RELATED WORK

Various multilingual ontology alignment strategies have been proposed: manual processing, the corpus-based approach, the linguistic enrichment, and the two-step generic approach [4].

Laing et al. [18] used manual mapping to map agricultural thesaurus in English to the Chinese equivalent. Whilst the manual mapping is costly and error-prone task, this approach is feasible only for relatively small and simple ontologies. Thus, fully/semi-automated multilingual ontology mapping strategies have emerged.

Corpus-based approaches use bilingual corpora for alignment discovery. In [19], by using this approach, Dutch thesaurus is aligned with the English thesaurus WordNet. This approach is applicable in situations where corpora of similar granularity and quality exist. Alas, for many domain-specific ontologies there are no adequate corpora to be used. In addition, this approach does not consider structural aspect and thus cannot provide precise mappings for ontologies with complex structure [4].

In instance-based approach, analysis of instance similarity is used for obtaining matching correspondences. This approach is based on machine learning methods and thus, is applicable for ontologies with sufficiently large number of instances. In [20], Wang et al. used annotations of instances to compute a measure of similarity between instances. Later this similarity was used to determine similarity between concepts.

According to proponents of linguistically enrichment strategy current ontologies suffer from unreadability due to badly chosen labels, lexical ambiguity etc., and thus, impeding the interoperability. They enrich the ontology's linguistic expressivity, through the exploitation of existing linguistic resources. A linguistically motivated mapping method has been proposed in [21]. Although linguistically enrichment of ontologies is beneficial, it is difficult to apply due to lack of linguistic resource standards.

In the generic two-step method, which was proposed in [7], the source ontology labels are translated into target language first and then monolingual matching techniques are applied. Since the translation does not take into account the semantics of involved ontologies, it can introduce inadequate translations that hamper the matching process. In these systems, the translation phase is crucial to success of ontology alignment. Therefore, obtaining the most suitable label translation is the key to generation of high quality alignments [4]. Fu et al. [4] addressed these issues with appropriate translation selection component. Namely, this component chooses the most appropriate translation amongst candidates with regard to target ontology semantics, the mapping intent, the operating domain, the time and resource constraints and user feedback.

Still, none of these approaches presents a comprehensive solution to the multilingual ontology alignment problem [22]. Thus, the multistrategy approaches have emerged. Various papers report that combination of strategies is highly dependent of the ontologies used. In [22], Li et al. presented a dynamic multistrategy ontology framework. They have used various similarity factors to select dynamically the most appropriate strategy for each individual alignment task. On the other hand, Songyun et al. propose an iterative supervised-learning weighted multistrategy alignment approach [23]. For each

alignment task, system computes weights for every strategy available and uses those weights to combine correctly the strategies.

We propose a conceptual idea to use images for alignment discovery between two multilingual ontologies. Unlike previous approaches images are used (visual representations of ontology concepts) to perform alignment discovery. Our idea is based on the fact that majority of ontology concepts are picturable entities, which can be found on the Web as images. Our approach complements the aforementioned approaches and adds a new dimension to the research field of multilingual ontology alignment.

VI. CONCLUSION AND FUTURE WORK

This paper is a concept paper, which introduced the idea of indirect alignment between multilingual ontologies by discovering alignments based on semantic-similarity of visual representations of ontology's concepts. Thus, the problem of finding adequate alignment between two concepts is reduced to the problem of matching their visual representations.

To the best of our knowledge, this is the first time the images as visual representations of ontology concepts, are exploited for multilingual ontology alignment.

Our idea is appealing, but has following limitations:

- The approach is suitable in situations where visual representations of concept exist and are available. The more the concept is visually discriminating, the easier is to obtain alignment using image similarity, and vice versa. For some broad and abstract concepts, the approach is not feasible because of their visual diversity (e.g., *animal* concept has very broad visual diversity). For some others concepts no appropriate image/s can be found on the Web.
- Comparing two images is complex, computationally expensive and context-dependant task itself.

As future work, we want to: i) conduct experiments for evaluation of proposed idea and level of applicability; ii) implement a prototype that is build upon the presented idea; iii) investigate how this idea could be combined with existing strategies into synergy-based dynamic multistrategy alignment framework to enhance alignments accuracy and precision.

ACKNOWLEDGMENT

This research is financial supported by Ministry of Science and Technological Development, Republic of Serbia; under the project number III47003 "Infrastructure for Technology Enhanced Learning in Serbia", 2011-2014.

REFERENCES

- [1] A. Popescu, "Image retrieval using a multilingual ontology". In Large Scale Semantic Access to Content (Text, Image, Video, and Sound) (RIAO '07). Le Centre De Hautes Etudes Internationales D'informatique Documentaire, Paris, France, pp. 461-474 (2007)
- [2] S. Wang, F. Jing, J. He, Q. Du, and L. Zhang, "IGroup: presenting web image search results in semantic clusters". In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '07). ACM, New York, NY, USA, pp. 587-596. doi:10.1145/1240624.1240718 (2007)
- [3] B.J. Jansen, A. Spink, and J. Pedersen, "An Analysis of Multimedia Searching on Altavista". In Proc. of the 5th ACM SIGMM MIR Workshop, Berkeley, CA, USA, (2004).

- [4] B. Fu, R. Brennan, and D. O'Sullivan, "Cross-Lingual Ontology Mapping and Its Use on the Multilingual Semantic Web". In 1st Workshop on the Multilingual Semantic Web, at the 19th International World Wide Web Conference (WWW 2010), Raleigh, USA, April 27th, 2010. CEUR Vol.571, 13 – 20 (2010)
- [5] E. Beisswanger, "Exploiting relation extraction for ontology alignment". In: Patel-Schneider, P. F., Pan, Y., Hitzler, P., Mika, P., Zhang, L. (eds.) ISWC'10, LCNS, Vol. 6497/2010 Part II, pp. 289-296, Springer, Heidelberg (2010)
- [6] J. Euzenat, and P. Shvaiko, *Ontology matching*. Springer, Heidelberg, Germany (2007)
- [7] C. Trojahn, P. Quaresma, and R. Vieira, "A framework for multilingual ontology mapping". In: Proceedings of the International Conference on Language Resources and Evaluation, LREC. pp. 1034- 1037 (2008)
- [8] Rosetta Stone, a language-learning software, [Online] Available at: <http://www.rosstattstone.co.uk>, 11/28/2011
- [9] S. Antonic, and C. Krstev, "Serbian Wordnet for Biomedical Sciences". In INFORUM 2008: 14th Conference on Professional Information Resources, Prague, May 28-30. (2008)
- [10] M. Espinoza, A. Gomez-Perez, and E. Mena. "Enriching an ontology with multilingual information". In Proceedings of the 5th European semantic web conference on The semantic web: research and applications (ESWC'08), S. Bechhofer, M. Hauswirth, J. Hoffmann, and M. Koubarakis (Eds.). Springer-Verlag, Berlin, Heidelberg, 333-347. (2008)
- [11] R. J. Sternberg. *Cognitive Psychology*. Wadsworth, 5th edition, 2008.
- [12] T. Deselaers and V. Ferrari, "Visual and semantic similarity in ImageNet", in CVPR, IEEE, , pp. 1777-1784 (2011)
- [13] WordNet, a lexical database for the English language, [Online] Available at: <http://wordnet.princeton.edu>, 11/28/2011
- [14] R. Paredes and F.M. Segarra, "Interactive Image Retrieval". *Multimodal Interactive Pattern Recognition and Applications*. A. H. Toselli, E. Vidal and F. Casacuberta (Eds.), Springer (2011)
- [15] J. Milgram, R. Sabourin and M. Cherié, "Combining Model-based and Discriminative Approaches in a Modular Two-Stage Classification System: Application to Isolated Handwritten Digit Recognition", *Electronic Letters on Computer Vision and Image Analysis*, Volume 5, Issue 1, pp. 14-29 (2005)
- [16] Jung, J.J., Hakansson, A., Hartung, R.: Indirect Alignment between Multilingual Ontologies: A Case Study of Korean and Swedish Ontologies. In: Hakansson, A., Nguyen, N.T., Hartung, R.L, Howlett, R.J, Jain, L.C. (eds.) KES-AMSTA '09, LCNS, Vol. 5559/2009, pp. 233-241, Springer, Heidelberg (2009)
- [17] E. G. K. Chang, G. Sychay and W. Gang, "CBSA: content-based soft annotation for multimodal image retrieval using bayes point machines". *IEEE Transactions on Circuits and Systems for Video Technology* 13(1) pp. 26–38 (2003)
- [18] A. C. Liang, and M. Sini, "Mapping AGROVOC and the Chinese agricultural thesaurus: definitions, tools, procedures". *New Review of Hypermedia and Multimedia*, vol. 12:1, pp. 51-62 (2006)
- [19] V. Malaise, A. Isaac, L. Gazendam, H. Brugman, "Anchoring Dutch Cultural Heritage Thesauri to WordNet: Two Case Studies". In: Proceedings of the Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2007), pp. 57 – 64, Prague, Czech Republic (2007)
- [20] S. Wang, G. Englebienne, S. Schlobach, "Learning concept mappings from instance similarity". In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T.W. Thirunarayan, K. (eds.) *The Semantic Web - ISWC 2008, 7th International Semantic Web Conference, ISWC 2008, Karlsruhe, Germany, October 26-30, 2008, LCNS, vol. 5318, pp. 339-355 (2008)*
- [21] T. M. Paziienza, A. Stellato, "Linguistically Motivated Ontology Mapping for the Semantic Web". In: Proceedings of SWAP 2005, the 2nd Italian Semantic Web Workshop, Trento, Italy, December 14-16, 2005, CEUR Workshop Proceedings, vol. 166, (2005)
- [22] J. Li, J. Tang, Y. Li, Q. Luo, "RiMOM: A Dynamic Multistrategy Ontology Alignment Framework Journal", *IEEE Transactions on Knowledge and Data Engineering - TKDE* , vol. 21, no. 8, pp. 1218-1232 (2009)
- [23] D. Songyun, F. Achille, S. Kavitha, "One Size Does Not Fit All: Customizing Ontology Alignment Using User Feedback", In: Patel-Schneider, P. F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J., Horrocks, I., Glimm, B. (eds.) ISWC'10, LCNS, vol. 6496. pp. 177-192. (2010)