Expert Systems with Applications 41 (2014) 5201–5211

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Evaluating techniques for learning non-taxonomic relationships of ontologies from text

Ivo Serra [a], Rosario Girardi [a,*], Paulo Novais [b]

[a] Computer Science Department, Federal University of Maranhão (UFMA), São Luís, MA, Brazil
[b] Department of Informatics, University of Minho, Braga, Portugal

ARTICLE INFO

Keywords:
Learning non-taxonomic relationships
Ontology
Ontology learning
Natural language processing
Machine learning

ABSTRACT

Learning non-taxonomic relationships is a sub-field of Ontology Learning that aims at automating the extraction of these relationships from text. Several techniques have been proposed based on Natural Language Processing and Machine Learning. However just like for other techniques for Ontology Learning, evaluating techniques for learning non-taxonomic relationships is an open problem. Three general proposals suggest that the learned ontologies can be evaluated in an executable application or by domain experts or even by a comparison with a predefined reference ontology. This article proposes two procedures to evaluate techniques for learning non-taxonomic relationships based on the comparison of the relationships obtained with those of a reference ontology. Also, these procedures are used in the evaluation of two state of the art techniques performing the extraction of relationships from two corpora in the domains of biology and Family Law.

1. Introduction

Manual construction of ontologies by domain experts and knowledge engineers is a costly task, thus automatic and/or semi-automatic approaches for their development are needed. Ontology Learning (OL) (Buitelaar, Cimiano, & Magnini, 2006; Cimiano, Volker, & Studer, 2006; Girardi, 2010) aims at identifying the constituent elements of an ontology, such as non-taxonomic relationships (Serra, Girardi, & Novais, 2012), from textual information sources.

Several techniques for learning non-taxonomic relationships have been proposed. Some of them use linguistic patterns (Girju, Badulescu, & Moldovan, 2003), while others use statistical solutions (Sanchez & Moreno, 2008; Serra, Girardi, & Novais, 2013) or even machine learning (ML) (Fader, Soderland, & Etzioni, 2011; Maedche & Staab, 2000; Mohamed, Junior, & Mitchell, 2011; Villaverde, Persson, Godoy, & Amandi, 2009). All of them compare their results with a reference ontology. However, there are few studies on the comparison of results from one technique to another and moreover, there is a lack of formalization of evaluation procedures.

According to Dellschaft and Staab (2006) there are three ways to evaluate a learnt ontology: the resulting ontology can be evaluated in an executable application; by domain experts or even by comparing it with a predefined reference ontology (gold standard).

The use of an ontology in an executable application aims at measuring the effectiveness of a system that uses the ontologies being evaluated. A disadvantage of this approach is that other factors may impact the output of the system and sometimes the ontology is, in fact, a small part of the system with little interference in its results. The manual evaluation approach has its advantages, since it is expected that experts know the concepts and relationships of their domains of expertise, and therefore they are supposedly able to tell whether a given domain ontology is good or not. Disadvantages of these two proposals are their subjectivity and delay. Moreover, these methods are not feasible for large-scale evaluations. Thus, the comparison with a reference ontology is a plausible alternative since it permits the automation of the evaluation process. Proposals based on the comparison with reference ontologies are shown in Maedche and Staab (2000) and Dellschaft and Staab (2006). The main disadvantage of this approach is that a reference ontology is a handmade artifact and if it presents modeling problems, the evaluation method rewards ontologies with similar problems and penalizes ontologies with concepts or relationships that do not appear in the reference ontology.

This paper formally defines two procedures for evaluating techniques for Learning Non-Taxonomic Relationships of Ontologies (LNTRO) with respect to a reference ontology and uses them to

* Corresponding author.
E-mail addresses: ivocserra@gmail.com (I. Serra), rosariogirardi@gmail.com (R. Girardi), pjon@di.uminho.pt (P. Novais).

http://dx.doi.org/10.1016/j.eswa.2014.02.042
0957-4174/

comparatively evaluate two state of the art LNTRO techniques: Technique for Learning Non-taxonomic Relationships (TLN) (Serra et al., 2013) and Learning relationships based on the Extr action of Association Rules (LEAR) (Villaverde et al., 2009), to extract relationships from the corpus Genia (Rinaldi et al., 2004) and Family Law doctrine (FindLaw, 2013).

The paper is organized as follows: Section 2 introduces a general process for LNTRO (Serra et al., 2012). Section 3 presents a discussion about related work. In Section 4, two procedures for the evaluation of LNTRO techniques according to the generic process for LNTRO are presented. In Section 5, two LNTRO techniques that are used to illustrate the application of the evaluation procedures are briefly described. In Section 6, the results of the application of the two evaluation procedures to perform benchmarking of these LNTRO techniques are presented and discussed. Section 7 presents the conclusions and points out future lines of research for this work.

2. A general process for learning non-taxonomic relationships of ontologies

Based on the analysis of some techniques of the state of art (Fader et al., 2011; Girju et al., 2003; Maedche & Staab, 2000; Mohamed et al., 2011; Sanchez & Moreno, 2008; Villaverde et al., 2009) we have developed a generic process for LNTRO (Fig. 1) (Serra et al., 2012). The objectives were to have a guideline to suggest new LNTRO techniques and to facilitate comparative evaluations between techniques regarding the solutions they adopt for each one of its phases.

The corpus construction task selects documents of the domain from which relationships can be extracted. This is usually a costly task and the outcome of any LNTRO technique depends on the quality of the used corpus.

The extraction of candidate relationships task identifies a set of possible relationships. It has the corpus built in the previous phase as input and candidate relationships as its product. It is composed of two sub-activities: corpus annotation and extraction of relationships. The corpus annotation task tags the corpus using Natural Language Processing (NLP) techniques that are necessary for the next steps of LNTRO. In the extraction of relationships activity, the annotated corpus is searched for evidence suggesting the existence of relationships. For example, Maedche and Staab (2000) consider the existence of two instances of ontology concepts in a sentence as evidence that they are non-taxonomically related. For Villaverde et al. (2009), a relationship is identified by the presence of two ontology concepts in the same sentence with a verb between them.

The relationships obtained from the previous task should not be recommended to the specialist since there is usually a substantial amount of them that do not correspond to good suggestions. For this reason, in the refinement phase, machine learning (ML) techniques could be used to deliver the best suggestion to the specialist.

In the evaluation by the specialist task, he/she selects and possibly edits the relationships to be added to the ontology from those outputted from the previous phase. Finally, in the ontology update activity, the ontology is updated with the relationships that were chosen by the specialist.

One aspect of particular interest regarding LNTRO techniques is the type of representation adopted for the learned relationships. In the following we present some of the most common. The first is the one used by techniques that receive ontology concepts as input. There are two subtypes for this representation, depending if labels (typically verb phrases) are recommended. For the first subtype, the representation is $\langle c_1, vp, c_2 \rangle$ where c_1 and c_2 are ontology concepts and vp is a verb phrase. For example, considering the sentence "The court decree protects the property rights of the parties and provides support for the children" and "decree" and "property" as two ontology concepts, the relationship $\langle decree, protect, property \rangle$ would be extracted. Examples of techniques that use this representation are LEAR (Villaverde et al., 2009) and TLN (Serra et al., 2012). For the second subtype, the representation is $\langle c_1, c_2 \rangle$, where c_1 and c_2 are two concepts. For example, considering "court" and "decree" as ontology concepts and the sentence "The court decree protects the property rights of the parties and provides support for the children", the relationship $\langle court, decree \rangle$ would be extracted. An example of a technique that use this representation is the LNTRO based on the extraction of generalized association rules (Maedche & Staab, 2000).

The second type of representation is used when ontology concepts are not given as input to the LNTRO technique. In this case, noun phrases extracted from the corpus are used as ontology concepts. Here again there are two subtypes depending if labels are recommended. For the first subtype, the representation is $\langle np_1, vp, np_2 \rangle$ where: np_1 and np_2 are noun phrases and vp is a verb phrase. For example, from the sentence "The judge granted the custody of the child to his grandmother." the relationship $\langle \textit{the judge}, \textit{granted}, \textit{the custody} \rangle$ would be extracted. Examples of techniques that use this representation are: LNTRO based on Web queries (Sanchez & Moreno, 2008) and LNTRO based on logistic regression (Fader et al., 2011). The second subtype is $\langle np_1, np_2 \rangle$.

The procedures and evaluation measures (recall, precision and F-measure) used in the case studies presented in Section 6 are suitable for use with LNTRO techniques that adopt relationships of the

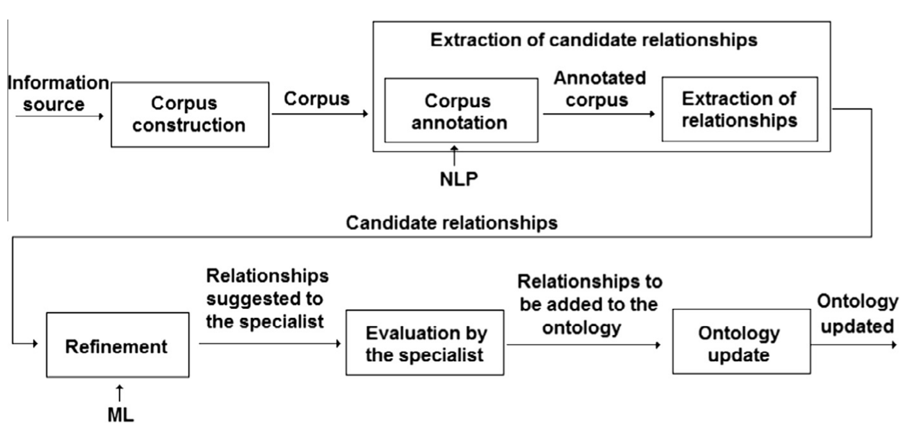

Fig. 1. A generic process for LNTRO.