LAAOS: Logical Aspects of Adaptable Ontological Schemas (Project Overview)

Michal Čertický[1,3], Martin Homola[1], Ján Kľuka[1], Simone Serra[2], Vojtěch Svátek[2], Alexander Šimko[1], Miroslav Vacura[2], and Ondřej Zamazal[2]

[1] Comenius University in Bratislava, Mlynská dolina, 842 48 Bratislava, Slovakia
[2] University of Economics, Prague, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic
[3] Agent Technology Center, Czech Technical University, Charles Sq. 13, 121 35 Prague 2, Czech Republic

Abstract. The Czech–Slovak cooperation project aimed at new results in the field of Semantic Web ontologies, via synergy between the formal logics background of the Bratislava group and conceptual modelling and Linked Data technology background of the Prague group. The most important achievements of the project are the PURO ontological background modelling language, an associated higher-order description logic, a metamodelling ontology allowing for conceptual coherency checking of OWL ontologies, and an implemented Protégé plugin for annotation of ontologies with PURO distinctions.

1 Project and Its Objectives

The advent of the Semantic Web (SW), and especially the more recent Linked Data (LD) initiative, resulted in an unprecedented amount of ontologies being published on the Web, and reused in diverse applications. This mass popularity of ontologies has brought forward a number of issues, among them the needs to combine, transform, and merge ontologies, and also different modelling of similar entities in several ontologies. This is often the case when the ontologies are crafted by Linked Data practitioners, majority of whom are not trained ontology experts, but it also results from limited expressiveness of the Semantic Web ontology language OWL.[4]

The LAAOS project addressed some of these issues. (a) Different ontologies may capture the same or similar concepts differently and hence recognizing and representing a uniform *background model* that helps to align the ontologies is often useful. (b) Background modelling may require constructs of higher orders, which are sometimes implicitly present in ontologies, hence there is a need for *logical languages* that can capture such constructs and model with them. (c) Different SW languages and their fragments support different expressive constructs, and there is a need to *transform* ontologies from one fragment to another while keeping the most of the logical meaning.

The Czech–Slovak cooperation project (reg. as 7AMB12SK020 with AIP ČR, and SK-CZ-0208-11 with APVV) was funded in 2012–2013 by 90,000 CZK and 4,000 EUR. The results, as summarized in this paper, have been published with additional support (on either one or other side) from the EU ICT FP7 project no. 257943, CSF project no. P202/10/1825, and VEGA project no. 1/1333/12.

[4] http://www.w3.org/TR/owl2-overview/

2 Results

2.1 PURO Ontology Background Modelling Language

SW ontologies, and especially LD vocabularies, are often crafted by non-experts in ontology engineering. What is more, the SW ontology languages have limited expressiveness. Consequently, imprecise or approximative modelling is often used. In particular, entities intuitively corresponding to general concepts may be modelled by OWL individuals (syntactical instances), for example, the individual Album[5] denotes the respective type of musical product. Similarly, relations are often reified (modelled as OWL classes), e.g., the relationship between a musician and a work s/he composed is expressed by an instance of class Composition rather than by an instance of an OWL property (i.e. binary relation). Also, different ontologies may model the same or similar concepts differently.

We propose that the "ontological background" that is implicit in the actual SW ontologies (also called foreground models) should be made explicit, and represented in backgrounds models. We proposed the *PURO background modelling language*, which allows to capture the background nature of the entities in foreground models by a set of annotation labels. PURO concentrates especially on the distinction between particulars (i.e., objects) and universals (i.e., classes) – labels $\mathcal{B}$-object vs. $\mathcal{B}$-type (the $\mathcal{B}$-prefix indicates background modelling) – and also on the distinction between (reified) relationships and true objects – labels $\mathcal{B}$-relationship vs. $\mathcal{B}$-object, and so on.

It is desired for the background model to be *coherent*, in the sense of satisfying constraints such as that, e.g., $\mathcal{B}$-objects should not have instances, that classes, property domains and property ranges should always contain a homogeneous set of entities, and so on. The set of constraints was codified in the *PURO ontology*. With the help of annotation labels and this ontology, the background coherence can be verified by a description logic (DL) reasoner. The results appeared at WIKT 2012 [9], OWLED 2013 [6], and in more detail in a technical report [7].

2.2 Typed Higher-Order Description Logics

While annotation labels and the PURO ontology allow to make the ontological background model explicit, and the background coherence can be checked, we would like to also represent the model in a suitable ontology language with logical semantics. However, background entities of higher orders may be implicit in a number of ontologies, such as classes that contain other classes (represented as individuals in the foreground model), etc. Current SW ontology languages do not allow higher-order classes. We proposed a typed higher-order extension of the $\mathcal{SROIQ}$ DL (the base of OWL 2), which allows for such treatment. In this language background models can be directly encoded and reasoned with. Strict type separation ensures background coherence of any such model. The language is also suitable for some metamodelling applications. The results appeared at DL 2013 [2] and 2014 [3].

[5] This as well as next example are from the popular Music Ontology, `http://musicontology.com/`.

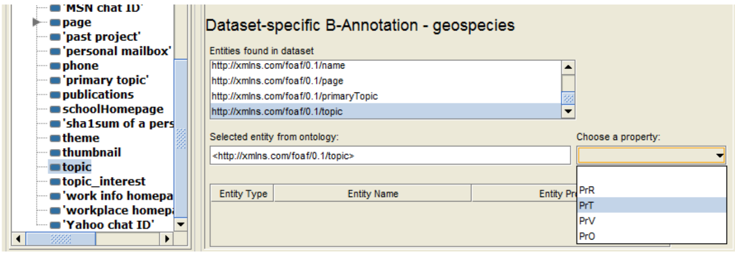

Fig. 1. B-Annot tool interface: annotating a property

2.3 Analyzing Ontology Patterns via Background Models

Besides concrete ontologies, the apparatus of PURO background models can also be applied to (empirical or explicit best-practice) patterns recurring in multiple ontologies. The prominent W3C-endorsed best practice pattern family, “Classes as Property Values”,[6] has been analysed in the PURO style and decomposed into three variants corresponding to different states of affairs [5]. A new pattern family for a related problem, “Multi-Instance Fact”, has also recently been formulated [4].

2.4 Annotation Experiments

To verify that the PURO language is at the same time non-trivial and potentially intuitive to users, two annotation experiments have been carried out. In the first one, two experienced ontological engineers annotated all classes from 3 popular LD vocabularies: FOAF, Music ontology and GoodRelations.[7] Although initially some discrepancies between the annotators appeared, perfect agreement was established through a short mutual clarification phase: 92 out of 94 classes were jointly assigned a dominant label, which indicates that with proper guidelines, PURO annotation can be intuitive in most cases. Furthermore, 35 of the classes (i.e., 37 %) received other than the ‘obvious’ label (which would be ‘level-1 $\mathcal{B}$-type’), which indicates that PURO models are not trivial counterparts to the foreground OWL models. Later on, 13 undergraduate students, after a 1-hour explanation of PURO, achieved, on average, a 57% match to expert annotation on a representative sample of 20 classes from the above-mentioned vocabularies.

2.5 *B-Annot* Protégé Plug-In

To support the annotation of ontologies with PURO labels we developed a Protégé plug-in called *B-Annot*.[8] Generic (w.r.t. ontology specification) but also dataset-specific annotation is enabled [8], the latter addressing the issue that in some dataset, certain

[6] `http://www.w3.org/TR/swbp-classes-as-values/`

[7] `http://xmlns.com/foaf/spec/`, `http://purl.org/ontology/mo/`, `http://purl.org/goodrelations/v1#`

[8] Available from `http://patomat.vse.cz/cz.vse.bannotation.plugin.view.jar`.

entities may be used differently than originally intended by the author of the ontology. The screenshot in Fig. 1 shows an assignment of a label ("PrT": property whose range is a type) to the topic property of FOAF. B-Annot is general enough to also capture other background modelling languages apart from PURO; OntoClean [1] is also supported.

2.6 Language Profiling

Within the project we work on the language profiling scenario, where an user needs (from diverse reasons, e.g., complexity downgrading) to replace language constructs while keeping the most of the logical meaning. Our approach aims at an automatic transformation of forbidden language constructs in the resulting/target ontology. The automation is enabled by an automated search for transformation patterns based on an evolutionary algorithm. Patterns that keep the most of the logical meaning are directly forwarded to the next generation. Others undergo mutation and crossover operations before they enter the next generation. This process involves a transformation patterns generator generating all three transformation pattern parts, i.e., a source and a target ontology pattern and a pattern transformation. Currently, we prepare further experiments where the generation of transformation patterns is restricted to axioms already present in the transformed ontology.

3 Future Work

Further results concerning evaluation of existing linked data ontologies based on PURO background models, annotation experiments, and language profiling are being prepared. The project sparked our interest in follow-up topics, such as metamodelling and investigation of the relationship between materialization and higher-order ontology entities.

References

1. Guarino, N., Welty, C.: An overview of ontoclean. In: Staab, S., Studer, R. (eds.) The Handbook on Ontologies, pp. 151–172. IHIS, Springer (2004)
2. Homola, M., Kľuka, J., Svátek, V., Vacura, M.: Towards typed higher-order description logics. In: DL (2013)
3. Homola, M., Kľuka, J., Svátek, V., Vacura, M.: Typed higher-order variant of $\mathcal{SROIQ}$ – why not? In: DL (2014)
4. Svátek, V., Homola, M., Kľuka, J., Vacura, M.: Referring to multiple unspecified objects of a type: multi-instance fact pattern family. Submitted to Semantic Web Journal
5. Svátek, V., Homola, M., Kľuka, J., Vacura, M.: Mapping structural design patterns in OWL to ontological background models. In: K-CAP (2013)
6. Svátek, V., Homola, M., Kľuka, J., Vacura, M.: Metamodeling-based coherence checking of OWL vocabulary background models. In: OWLED (2013)
7. Svátek, V., Homola, M., Kľuka, J., Vacura, M.: Ontological distinctions for linked data vocabularies. Tech. Rep. TR-2013-0039, Comenius University (2013), available online: `http://kedrigern.dcs.fmph.uniba.sk/reports/download.php?id=54`
8. Svátek, V., Serra, S., Vacura, M., Homola, M., Kľuka, J.: B-Annot: Supplying background model annotations for ontology coherence testing. In: WoDOOM (2013)
9. Svátek, V., Vacura, M., Homola, M., Kľuka, J.: Towards conceptual structure verification of linked data vocabularies. In: WIKT (2013)