Rough Set Based Approximations of Classes in the OWL Ontology of Places in Poland

Extended Abstract

Piotr Grochowalski[1], Krzysztof Pancerz[1] and Tomasz Szul[2]

[1] Chair of Computer Science, Faculty of Mathematics and Natural Sciences
University of Rzeszów
Prof. S. Pigonia Str. 1, 35-310 Rzeszów, Poland
{piotrg,kpancerz}@ur.edu.pl
[2] Department of Power Engineering and Automation of Agricultural Processes
Faculty of Production and Power Engineering
University of Agriculture in Krakow, Poland
Balicka Str. 116B, 30-149 Kraków
tomasz.szul@ur.krakow.pl

Abstract. The main goal of our research is to build the ontology of places in Poland covering a variety of aspects of places, mainly administrative and socio-economic. The ontology is being implemented using the OWL 2 Web Ontology Language. In the created OWL ontology, we can distinguish two kinds of classes, primary classes exactly defined in the ontology as well as secondary classes defined over the ontology on the basis of primary classes and properties of individuals considered in the ontology. We show how to use rough sets to approximate secondary classes by means of primary classes in the created ontology. Rough set approximations enable us to extract some useful knowledge about places.

Keywords: rough sets, approximation, ontology, OWL 2.

1 Ontologies and Semantic Relations

Ontologies, as formal representations of knowledge, have recently gained a significant popularity. They are currently used in knowledge engineering and data mining to capture knowledge about some domain of interest. Two reasons seem to be the main source of this popularity. Firstly, there exist well-defined standards of languages for the ontology representation. Secondly, ontologies cover various semantic aspects of information which are useful in data mining processes.

One of the key decisions to take in the ontology development process is to select the language in which the ontology will be implemented. Our ontology of places is built in accordance with the OWL 2 Web Ontology Language (shortly OWL 2). OWL 2 is the most recent development in standard languages defined by the World Wide Web Consortium (W3C) [2]. An OWL ontology consists of three components: classes, individuals, and properties. Classes are representations of concepts in a given domain of interest. Classes are interpreted as sets

that contain individuals. Individuals (also known as instances) represent objects in the domain of interest. Individuals can be referred to as being instances of classes. Properties (also known as roles or attributes) are binary relations on individuals. Properties link two individuals together. There are two main types of properties in OWL 2: object properties linking an individual to an individual and data properties linking an individual to a data value.

Semantic relations are very important components of ontologies as they describe the relationships that can be established between concepts. In the presented approach to rough set based approximations of classes in the OWL ontology, we are interested in the INSTANCE-OF relation as well as specific semantic relations describing relationships covering economic and social aspects of places. Such relations are represented in the OWL ontology by object and data properties. If i INSTANCE-OF c holds, it means that the individual i is an instance (example) of a given class c. It is worth recalling that i is also an instance of all superclasses of c.

2 The OWL Ontology of Places in Poland

In [9], we showed selected parts of the ontology of places in Poland (defined classes and class hierarchies, identified individuals, identified properties linking individuals). The ontology is being implemented using the OWL 2 Web Ontology Language. In general, the main goal of our research is to build the ontology of places in Poland covering a variety of aspects of places, mainly administrative and socio-economic. In the next section, we show how to use rough sets to approximate secondary classes by means of primary classes in the created ontology of places. Rough set approximations enable us to extract some useful knowledge about places. The ontology built by us can be used in various socio-economic research as the knowledge base. Moreover, it may constitute the basis for search engines and other computer tools used in the real-estate market.

3 Approximations of Classes in OWL Ontologies

Rough sets proposed by Z. Pawlak [10] are an appropriate tool to deal with rough (ambiguous, imprecise) concepts in the universe of discourse. There are various approaches to applying rough sets for a representation of vague knowledge and reasoning over it in ontologies (e.g., [4], [6], [7]). For example, in [4], a rough set approach to vague concept approximations was presented. The concept approximations were constructed on the basis of data sets (decision tables with condition attributes representing, e.g., sensory measurements) and an additional domain knowledge (the so-called concept ontology) using approximate reasoning schemes. In the current paper, we consider a situation where the whole knowledge is included in a domain ontology (implemented using the OWL 2 Web Ontology Language). Following the approach presented in [7], we propose to apply rough sets to approximate secondary classes by means of primary classes in the created OWL ontology of places.

Let us consider, as an example, a part of the ontology of places in Poland devoted to administrative districts. In our ontology, we have distinguished three administrative types of communes:

- urban commune,
- rural commune
- urban-rural commune.

Moreover, according to [3], where functional structures of communes in Poland were considered, we have distinguished eight basic functional types of communes:

- urban commune,
- urbanized commune,
- multifunctional transitional commune,
- overwhelmingly agricultural commune,
- prevalently agricultural commune,
- tourism and recreational function commune,
- forestry function commune,
- mixed function commune.

In our ontology of places, all of the types of communes shown above are represented by primary classes, i.e., classes exactly defined in the ontology.

Some of the socio-economic aspects of places considered in our ontology are issues related to waste management. They are expressed especially by means of data properties of individuals, for example, the rate $MAHW$ of mass accumulation of household waste. Using this rate, we can define a secondary class representing the concept "administrative district with the rate of mass accumulation of household waste greater than or equal to 100 $\frac{kg}{person \cdot year}$". One can see that such a class is not exactly defined in the ontology, but it can be derived from the primary class and one of the properties of individuals.

The semantics of the OWL 2 Web Ontology Language is complex (see [1]). Therefore, we omit the formal description of the considered problem of rough set based approximations of classes in the OWL ontology and give only its brief review, rather informal. Let C be a set of classes and I be a set of individuals in a given OWL ontology $\mathcal{O}$. For a given class $c \in C$, we consider a set $INST(c)$ of all individuals from I that are instances of c.

Analogously to rough approximation of sets defined in rough set theory [10], we can define rough approximation of a given secondary class c^* by means of primary classes. The lower approximation $lower(c^*)$ of c^* is given by:

$$lower(c^*) = \{c \in C : \underset{i \in INST(c)}{\forall} i \in INST(c^*)\}.$$

The lower approximation of a secondary class c^* consists of each primary class c such that all individuals being instances of c are also instances of c^*.

The upper approximation $upper(c^*)$ of c^* is given by:

$$upper(c^*) = \{c \in C : \underset{i \in INST(c)}{\exists} i \in INST(c^*)\}.$$

The upper approximation of a secondary class c^* consists of each primary class c such that there exists at least one individual being an instance of c which is also an instance of c^*.

Conventionally, the boundary region $bound(c^*)$ is defined as:

$$bound(c^*) = upper(c^*) - lower(c^*).$$

Let us consider a part of individuals in our ontology which are communes in the Lubelskie Voivodship. In this voivodship, we have 213 communes distributed into administrative types of communes as follows: 20 urban communes, 169 rural communes, and 24 urban-rural communes. In case of functional types of communes, we have the following distribution: 20 urban communes, 5 urbanized communes, 7 multifunctional transitional communes, 57 overwhelmingly agricultural communes, 102 prevalently agricultural communes, 11 tourism and recreational function communes, 2 forestry function communes, and 9 mixed function communes.

On the basis of data included in Tables 1 and 2, we obtain the following approximations of the class representing the concept "administrative district with the rate of mass accumulation of household waste greater than or equal to 100 $\frac{kg}{person \cdot year}$":

- the lower approximation consists of a class representing the concept "urban commune" (as a functional type), only,
- the upper approximation consists of classes representing the concepts "urban commune" (as an administrative type), "rural commune", "urban-rural commune", "urban commune" (as a functional type), "urbanized commune", "multifunctional transitional commune", "prevalently agricultural commune", "tourism and recreational function commune", and "mixed function commune".

Table 1. Results of approximation for administrative types of communes

Administrative type	#Communes with $MAHW \geq 100$	#Communes with $MAHW < 100$
urban	19	1
rural	11	158
urban-rural communes	8	16

The obtained approximations enable us, for example, to make the following generalizations for communes in the Lubelskie Voivodship:

- an urban commune (as a functional type) is an administrative district with the rate of mass accumulation of household waste greater than or equal to 100 $\frac{kg}{person \cdot year}$ in the Lubelskie Voivodship (according to the lower approximation),

Table 2. Results of approximation for functional types of communes

Functional type	#Communes with $MAHW \geq 100$	#Communes with $MAHW < 100$
urban	20	0
urbanised	4	1
multifunctional transitional	2	5
overwhelmingly agricultural	0	57
prevalently agricultural	4	98
tourism and recreational function	7	4
forestry function	0	2
mixed function	1	8

– an urbanised commune may be an administrative district with the rate of mass accumulation of household waste greater than or equal to 100 $\frac{kg}{person \cdot year}$ in the Lubelskie Voivodship (according to the boundary region).

Such generalizations are useful knowledge derived from the ontology of places. One can see that the presented approach can be used in search engines for ontologies.

A valuable way of developing further research is to consider various approaches for determining approximations, for example, the Variable Precision Rough Set Model (VPRSM) [11] or those based on combined rough sets and fuzzy sets (cf. [5]).

References

1. OWL 2 Web Ontology Language direct semantics. W3C Recommendation. Tech. rep., W3C (2012)
2. OWL 2 Web Ontology Language structural specification and functional-style syntax. W3C Recommendation. Tech. rep., W3C (2012)
3. Bański, J.: Współczesne typologie obszarów wiejskich w Polsce - przegląd podejść metodologicznych (Contemporary typologies of rural areas in Poland - an overview of methodological approaches). Przegląd Geograficzny 86(4), 441–470 (2014)
4. Bazan, J., Skowron, A., Swiniarski, R.: Rough sets and vague concept approximation: From sample approximation to adaptive learning. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets V, Lecture Notes in Computer Science, vol. 4100, pp. 39–62. Springer-Verlag, Berlin Heidelberg (2006)
5. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems 17(2-3), 191–209 (1990)
6. Ishizu, S., Gehrmann, A., Nagai, Y., Inukai, Y.: Rough ontology: Extension of ontologies by rough sets. In: Smith, M.J., Salvendy, G. (eds.) Human Interface and the Management of Information. Methods, Techniques and Tools in Information Design, Lecture Notes in Computer Science, vol. 4557, pp. 456–462. Springer-Verlag, Berlin Heidelberg (2007)
7. Keet, C.M.: Ontology engineering with rough concepts and instances. In: Cimiano, P., Pinto, H.S. (eds.) Knowledge Engineering and Management by the Masses,

Lecture Notes in Artificial Intelligence, vol. 6317, pp. 503–513. Springer-Verlag, Berlin Heidelberg (2010)
8. Nguyen, L.A.: Paraconsistent and approximate semantics for the OWL 2 Web Ontology Language. In: Szczuka, M., Kryszkiewicz, M., Ramanna, S., Jensen, R., Hu, Q. (eds.) Rough Sets and Current Trends in Computing, Lecture Notes in Artificial Intelligence, vol. 6086, pp. 710–720. Springer-Verlag, Berlin Heidelberg (2010)
9. Pancerz, K., Grochowalski, P., Derkacz, A.: Towards the ontology of places in Poland: an example of the Mazowieckie voivodship. Barometr Regionalny. Analizy i Prognozy 14(3) (2016)
10. Pawlak, Z.: Rough Sets. Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
11. Ziarko, W.: Variable precision rough set model. Journal of Computer and System Sciences 46(1), 39–59 (1993)