The ontology analysis based on relations on arcs of the formal context

Bato Merdygeev[1], Sesegma Dambaeva[1]

[1]East Siberia State University of Technology and Management, Ulan-Ude, Russia
mainisjusticeone@gmail.com
damseg@gmail.com

Abstract. The paper presents an approach to the analysis of the domain ontology and the criterion of the method of analysis based on relations on arcs of the formal context. The approach allows to evaluate the completeness of ontology relations. We used a lattice of concepts to analyze the relations of ontology.

Keywords: ontology, domain, ontology analysis, concept lattice, relation, evaluating, completeness of the ontology relations

1 Intoduction

Currently, many intelligent systems use ontology as a knowledge base. The effectiveness of this system depends on the effectiveness of knowledge represented in the ontology. Regardless of the type of ontology its creation is a laborious and expensive task. At the same time, there is a possibility to get inefficient or incorrect knowledge of ontology. To avoid it is necessary to evaluate the quality of ontology at every stage of its production. In the existing ontology analysis methods are based on the expert evaluation. Experts in this case often act domain experts or knowledge engineers. The main problem here is the amount of time required for checking the quality of the ontology. Modern methods provide a variety of tools for ontology analysis, but most of them are only effective in ontologies with a certain structure. Therefore, a search for new approaches to the analysis of the quality of ontology of various structures is needed.

One such approach could be the approach to ontology evaluation, based on an analysis of the relations between the terms of concept lattice [3]. This approach allows one to analyze the ontology structure based on relations of concept lattice. In this paper, we analyze the relations on arcs of the formal context.

Domain ontology contains a structured open data, which makes it possible to assess the application of certain properties of FCA-based methods. In contrast to [3], we do not use a specific ontology in this paper. Instead we consider basic relations related to a concept lattice.

2 Relations on arcs of Formal Context

With the help of [4] we define the concepts of relations on arcs of a formal context.

The formal context K is a triple <G, M, I>, where G, M are sets and $I \subseteq G \times M$ is the binary relation between G and M. The elements of G are objects, the elements of M are attributes, and I is the incidence of the context <G, M, I>.

$$A' := \{m \in M \mid (g, m) \in I, \forall g \in A\}, \text{where } A \subseteq G$$

$$B' := \{g \in G \mid (g, m) \in I \ m \in B\}, \text{where } B \subseteq M \quad (1)$$

A pair (A, B) is a formal concept of <G, M, I> if and only if

$$A \subseteq G, B \subseteq M, A' = B \text{ and } A = B' \quad (2)$$

A and B are called the extent and intent of the formal concept (A, B), respectively.

The arrow relations [4] of the formal context <G, M, I> are defined as follows: for g, $h \in G$ and m, $n \in M$, let's say:

$$g \swarrow m :\Leftrightarrow \begin{cases} (g, m) \notin I \text{ and} \\ \text{if } g' \subseteq h' \text{ and } g' \neq h', \text{then } hIm, \end{cases} \quad (3)$$

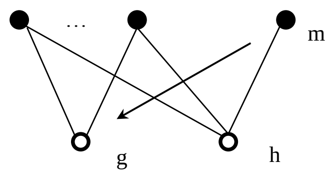

Fig. 1. Relation $g \swarrow m$

$$g \nearrow m :\Leftrightarrow \begin{cases} (g, m) \notin I \text{ and} \\ \text{if } m' \subseteq n' \text{ and } m' \neq n', \text{then } gIn, \end{cases} \quad (4)$$

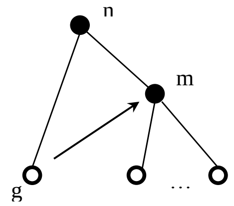

Fig. 2. Relation $g \nearrow m$

$$g \swarrow\!\!\!\nearrow m :\Leftrightarrow g \swarrow m \text{ and } g \nearrow m \quad (5)$$

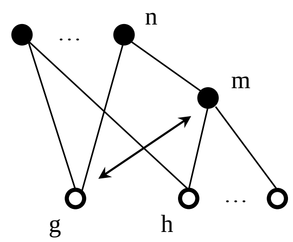

Fig. 3. Relation $g\swarrow\!\!\nearrow m$

For a given $g \in G$, there is an attribute $m \in M$, marked by $g\swarrow m$ if and only if γg is $\vee$-irreducible (minimal). Dual $g\nearrow m$ for the same $g \in G$ and only if μm is $\wedge$-irreducible.

In determining the arrow relations, the infimum and supremum of the lattice are not taken into account.

3 The approach to ontology analysis

3.1 Description of the approach

The approach to analysis is based on the approach used in [3]. In this paper, only the lattice internal structure is analyzed.

The purpose of the analysis of this approach is **the completeness of the ontology relations**. This property concerns the knowledge about the relations between domain terms displayed in the ontology.

To evaluate this property it is necessary to determine whether the ontology relations are complete and consistent.

The basis of the analysis is to search for possible missing relations on the lattice using arrow relations. The approach consists of the following steps (Figure 4):

1. Select the type of term relation that you want to analyze. Every relation type has its semantics, so the result of the analysis is interpreted according to the selected type.
2. Construct concept lattices for contexts where terms are taken as objects and attributes. Depending on the relation type it is possible to use different methods of constructing a formal context to maximize the effectiveness of the analysis. [11-17]
3. Search for possible missing relations on the lattice. In addition to the sets of relations, the result of this step is the set of values of the metrics of completeness (criterion) of ontology relations: the number of missing relations for each analyzed type of relations.
4. Result analysis. This analysis is performed by an expert, however, the evaluation on the criterion of completeness of ontology relations can automatically be made based on metric values.

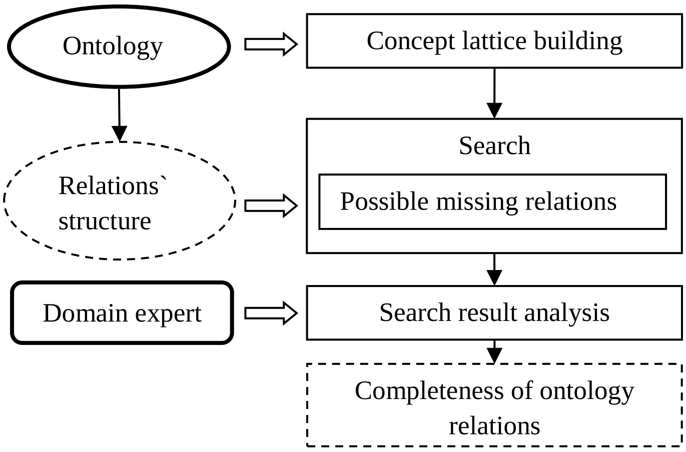

Fig. 4. Sequence of analysis steps

Arrow relations are divided into three types. On the basis of these types of relations a conclusion is made about possible missing of some necessary relations.

3.2 Search for possible missing relations

The criterion for determining the possible missing relations between an object and an attribute of the concept lattice is the number of arrow relations.

To determine the set of possible missing elements of the relations, it is necessary to search for all arrow relations between unrelated objects and attributes of the context.

Let us consider an example of the qualitative relation "Class-Kind" between the terms of an ontology. In this example, we use a simple method of the construction of the formal context: all terms that do not take the role of "Class" in any relations are formal objects, and other terms are formal attributes.

Table 1 provides a formal context for the relation.

Table 1. Example of formal context

G\M	Dog	Turtle	Chord	Predatory	Mammals
Dan	X		X	X	X
Bobby			X	X	X
Ellis		X			X
Katy			X		

Figure 5 shows a concept lattice on the formal context.

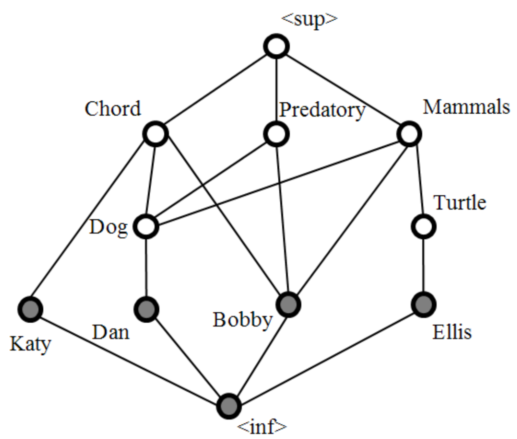

Fig. 5. Example of concept lattice

Consider the object "Bobby" and the attribute "Dog". They do not have an obvious relation. However, there are several $\swarrow\!\!\!\nearrow$-relations elements (Bobby$\swarrow\!\!\!\nearrow$Dog).

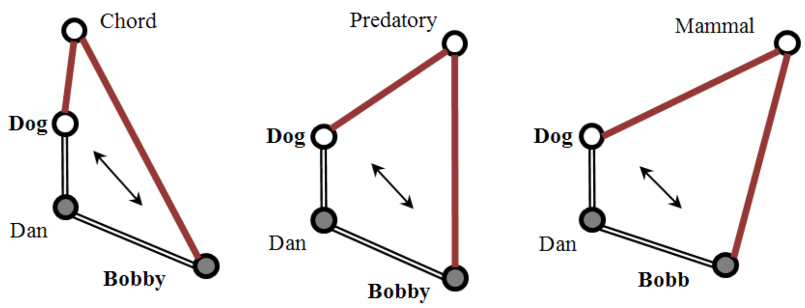

Fig. 6. Relation Bobby$\swarrow\!\!\!\nearrow$Dog

On the basis of formula 3, the set $S_{\swarrow}(g, m)$, consisting of terms by which the relation $g \swarrow m$ is defined.

$$S_{\swarrow}(g, m) = \{h : (g, m) \notin I \text{ and if } g' \subseteq h' \text{ and } g' \neq h', \text{then } hIm\} \tag{6}$$

On the basis of formula 4, the set $S_{\nearrow}(g, m)$, consisting of terms by which the relation $g \nearrow m$ is defined.

$$S_{\nearrow}(g, m) = \{n : (g, m) \notin I \text{ and if } m' \subseteq n' \text{ and } m' \neq n', \text{then } gIn\} \tag{7}$$

On the basis of formula 5, the pair (g,m) satisfies

$$S_{\swarrow\!\!\!\nearrow}(g, m) = (S_{\swarrow}(g, m), S_{\nearrow}(g, m)) \tag{8}$$

and represents all the terms that support the possibility that between g and m there must be the relation.

The found sets allow assuming that in the construction of ontology some relation elements were omitted. Also it can allow deducing new knowledge from concept lattice on certain type of relations. These data are provided by domain expert.

In this example:

$S_{\swarrow}$(Bobby, Dog) = {Dan},

$S_{\nearrow}$(Bobby, Dog) = {Chord, Predatory, Mammal},

$S_{\swarrow\nearrow}$(Bobby, Dog) = ({Dan}, {Chord, Predatory, Mammal}),

…

$S_{\swarrow\nearrow}$(Katy, Dog) = ({Dan}, {Chord}),

$S_{\swarrow\nearrow}$(Dan, Turtle) = ({}, {Mammals}),

$S_{\swarrow\nearrow}$(Bobby, Turtle) = ({}, {Mammals}),

$S_{\swarrow\nearrow}$(Ellis, Dog) = ({}, {Mammals}).

Supports of relation elements (Dan, Turtle), (Bobby, Turtle) and (Ellis, Dog) are very low, hence these pairs will not be considered.

A set of all found terms can be used to evaluate the completeness of ontology relations.

$$S_K = \{S_{\swarrow\nearrow}(g,m) : \forall g \in G, \forall m \in M, \text{где } S_{\swarrow}(g,m) \neq \varnothing \text{ или } S_{\nearrow}(g,m) \neq \varnothing\} \quad (9)$$

S_K is the set of pairs that support possible missing relations. For each type of relations, a separate set of S_K is generated.

To reduce the number of unimportant relations, we should use threshold values for $S_{\swarrow}(g, m)$ and $S_{\nearrow}(g, m)$ before $S_{\swarrow\nearrow}(g, m)$ is included in the set S_K.

4 The evaluation of completeness of relations on the basis of arrow relations

4.1 The importance of the attribute in the concept lattice

During the step of searching arrow relations, it should be taken in mind that the importance (weight) of the attributes can be different. The importance of attributes can be calculated by different methods. In this paper we consider the following method.

The importance of the attribute m is calculated based on the number of related attributes.

$$w(m) = |\{n : n \in M, m' \subset n'\}| + 1 \quad (10)$$

In (10) we must write “+1” for the accounting concept lattices top element.

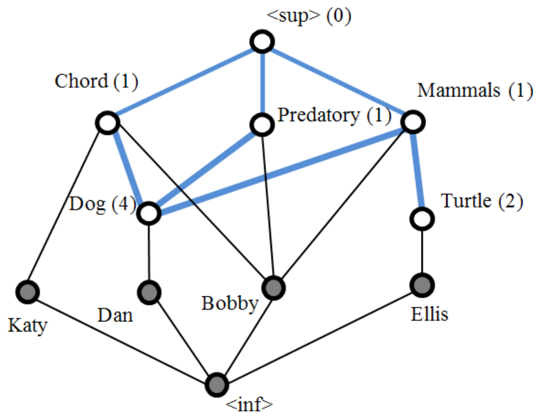

Fig. 7. The distribution of the importance of attributes

This indicator shows how many elements will actually be added to the formal context when adding a relation elements. The higher the value of this indicator, the higher the attribute's importance in the arrow relations analysis.

When evaluating the completeness of the ontology relations based on the obtained S_K, account should be taken of the importance of the attributes for each set $S_{\nearrow}(g, m)$ in $S_{\swarrow\nearrow}(g, m) \in S_K$. Thus, the indicator of importance $S_{\swarrow\nearrow}(g, m)$ is calculated as follows.

$$F(S_{\swarrow\nearrow}(g,m)) = |S_{\swarrow}(g,m)| + \sum_{i}^{|S_{\nearrow}(g,m)|} w(n_i), \tag{11}$$

where $n_i \in S\,(g, m)$, $w(n_i)$ is the importance of the attribute.

4.2 The importance of the attribute in the concept lattice

The measure of the criterion for evaluating the completeness of a relations is the number of possible missing relations.

Evaluation using this criterion can be carried out in several stages:

1) It is necessary to filter S_K separately for each type of relations, using the threshold for the importance indicator $F\,(S_{\swarrow\nearrow}(g, m))$ of possible missing relations.
   The threshold value can be calculated in different ways. For example, the threshold can be equal to the average value of the importance indicator of possible relations.

$$p = \frac{\max(F(S_{\swarrow\nearrow}(g,m))) + \min(F(S_{\swarrow\nearrow}(g,m)))}{2}. \tag{12}$$

   As a result, we get the updated set S_K'.
2) To get the value of the indicator of criterion for a certain type of relations it is necessary to calculate the sum of all the remaining indicators of the importance of S_K'.

$$F_s = \sum \mathrm{F}(S_{\swarrow\nearrow}(g,m)), \tag{13}$$

   where $S_{\swarrow\nearrow}(g,m) \in S_K'$.

3) In order to take into account the selection threshold p, one of the following formulas can be used.

$$F_s' = pF_s$$

$$F_s' = \frac{\min(F(S_{\nearrow}(g,m)))}{p} F_s$$

$$F_s' = \frac{p}{\max(F(S_{\nearrow}(g,m)))} F_s \tag{14}$$

4) To calculate the total value of the indicator for all types of relations, you can, for example, use one of the following formulas.

$$F_o = \sum F_{si}$$

$$F_o = \frac{\sum_i^N F_{si}}{N}$$

$$F_o = \left(\prod_i^N F_{si}\right)^N, \tag{15}$$

where i is the index of the relations type, N is the number of relations types.

Thus, the value of the indicator is obtained by the criterion of completeness of ontology relations. The smaller the value obtained, the more consistent and complete the structure of relations in ontology.

5 Conclusion

The proposed approach to the analysis of the completeness of ontology relations allows us to identify hidden dependencies between terms and provide them to the expert for further evaluation. Thus, the automation of the search for possible missing ontology relations is achieved, which speeds up the work of the expert, and also allows the identification of hidden knowledge that can be derived from the original structure of the ontology.

The criterion for evaluation of the completeness of the ontology relations allows one to obtain a numerical value of the general indicator of incompleteness of ontology relations. When you change or correct an ontology, this indicator can help to control the structure of the ontology relations.

In [3], the structure of ontology relations from the category of terms was considered and compared to the structure with respect to the concept lattice of this relation.

In this paper, the internal structure of the relation is considered only on the basis of concept lattice and an analysis is made on the basis of arrow relations.

These methods of evaluating the structure of an ontology are based on a uniform approach to evaluation, and it is possible to combine the methods for more efficient analysis.

References

1. Bezdushnii A.N., Gavrilova E.A., Sehebryakov V.A., Shkotin A.V. The place of ontologies in a single integrated system of the Russian Academy of Sciences. – url: http://www.olap.ru/home.asp?artId=2362.
2. Klesh’eva A.S., Shalfeeva E.A.: Classification of ontology properties. Ontologies and their classification. In: IACP, 2005, 20 p., Preprint, Vladivostok (2005)
3. Merdygeev B, Dambaeva S.: Evaluation of Ontology Quality based on Analysis of Relations in Concept Lattices. CLA 2016 (CDUD 2016). – http://ceur-ws.org/Vol-1625/proposal1.pdf
4. Ganter B., Wille R. Formal Concept Analysis: Mathematical Foundations, Springer (1999)
5. Ganter B., Wille R. Elliot H.: Applied Lattice Theory: Formal Concept Analysis. – http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.9907
6. Rudolf Wille: Formal Concept Analysis as Applied Lattice Theory. In: CLA 2006, pp.42-67 (2006)
7. Shalfeeva E.A., Gribova V.V. The internal properties of ontology. In: SICPRO'05. IV International conference. pp. 1109-1128, Moscow (2005)
8. Uta Priss: Knowledge Discovery in Databases Using Formal Concept Analysis. In: Bulletin of the American Society of Information Science 27, vol. 1, p. 18-20 (2000)
9. Dahlberg Ingetraud: Knowledge Organization: its scope and possibilities / Knowledge Organization: problems and trends // Proceedings of Conference reports, Moscow (1993)
10. Nikitina S.E.: Semantic analysis of the language of science. In linguistics material. Monograph. Book House "LIBROKOM", 146 p., Moscow (2010)
11. Stumme G., Maedche A.: FCA-merge: Bottom-Up Merging of Ontologies // IJCAI'01 Proceedings of the 17th international joint conference on Artificial intelligence, 2001 - https://www.researchgate.net/publication/2475502_FCA-Merge_Bottom-up_merging_of_ontologies
12. Stumme G., Cimiano P., Hotho A., Tane J.: Conceptual Knowledge Processing with Formal Concept Analysis and Ontologies. – http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.1946
13. Wille R.: Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts In: I. Rival (ed.): Ordered sets. Reidel, Dordrecht-Boston, p. 445-470 (1982)
14. Uta Priss: Lattice-based Information Retrieval Knowledge Organization. – http://www.upriss.org.uk/papers/ko00.pdf
15. Poelmans J., Kuznetsov S., Ignatov D., Dedene G.: Formal Concept Analysis in knowledge processing: A survey on models and techniques. In: Expert Systems with Applications, vol. 40, No. 16, pp. 6601-6623 (2013)
16. Tam T. Nguyen, Siu Cheung Hui, Kuiyu Chang: A lattice-based approach for mathematical search using Formal Concept Analysis. In: Expert Systems with Applications, vol. 39, No. 5, pp. 5820-5828 (2012)
17. Kuznetsov S., Poelmans J.: Knowledge representation and processing with formal concept analysis. In: Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol.3(3), pp. 200-215 (2013)