Development of a fuzzy knowledge base based on context analysis of problem area

N.G. Yarushkina[1], A.A. Filippov[1], V.S. Moshkin[1]

[1]Ulyanovsk State Technical University, Severny Venetz str. 32, Ulyanovsk, Russia, 432027

Abstract. The article describes the process of developing a fuzzy knowledge base (KB). The content of fuzzy KB is formed as a result of the analysis of the contexts of the problem area (PrA). In this case, the context is a certain "point of view" on the PrA and its features. A graph database (DB) is used as the basis for storing the contents of the KB in the form of an applied ontology. An attempt is made to implement the mechanism of inference by the contents of a graph database. The mechanism is used to dynamically generate the screen forms of the user interface to simplify the work with the KB.

Keywords: knowledge base, problem area, applied ontology, inference, context.

1. Introduction

Post-industrial society operates with huge volumes of information both in everyday and professional activities. A large amount of information causes difficulties in making decisions within the framework of rigid time constraints [1] [2].

A variety of software automation of human activities are used to solve this problem. However, it is necessary to adapt them to the specifics of a particular problem area (PrA) and its contexts for the effective operation of these tools [3] [4] [5] [6] [7] [8] [9].

Thus, "trained" automation software solves the tasks more efficiently, but they require considerable resources (human and temporary) for training.

In this paper, an attempt is made to construct a fuzzy knowledge base (KB). The content of the fuzzy KB is an applied ontology. The basic requirements for fuzzy KB are (fig.1):

- adaptation to the specifics of PrA based on contexts;
- reliability and speed of ontology storage;
- the presence of a mechanism of logical inference;
- availability of tools to simplify work with the KB for unprepared users;
- availability of mechanisms for importing data from external information resources.

As you can see from Figure 1, the KB consists of the following subsystems:

1. Ontology store:
    - Neo4j;
    - content management module;
    - ontology import/export module (RDF, OWL).
2. Inference subsystem:

- inference module.

3. A subsystem for interaction with users:
   - screen forms generation module.
4. A subsystem for importing data from external information resources:
   - a module for importing data from external wiki-resources;
   - a module for filling external wiki-resources.

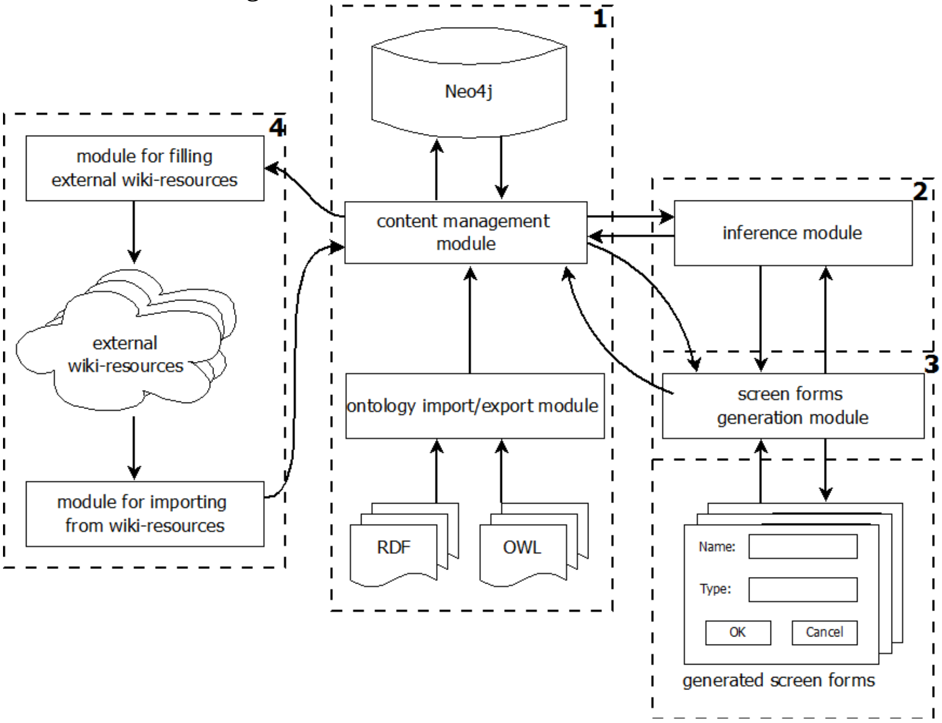

Figure 1. Fuzzy KB architecture.

2. The organization of the ontology store of fuzzy KB

Ontology is a model of the representation of the PrA in the form of a semantic graph [10] [11]. Graph-oriented database management system (Graph DBMS) Neo4j is the basis of the ontology store for fuzzy KB. Neo4j is currently one of the most popular graph databases and has the following advantages:

1. Having a free community version.
2. Native format for data storage.
3. One copy of Neo4j can work with graphs containing billions of nodes and relationships.
4. The presence of a graph-oriented query language Cypher.
5. Availability of transaction support.

Neo4j was chosen to store the description of the PrA in the form of an applied ontology, since the ontology is actually a graph. In this case, it is only necessary to limit the set of nodes and graph relations into which ontologies on RDF and OWL will be translated.

The context of an ontology is some state of ontology, obtained during versioning or building an ontology using different “points of view”.

Figure 2 shows an example of the translation of the owl representation of ontology of family relations into the ontology of the fuzzy KB.

Formally, the ontology of the fuzzy KB can be represented by the following equation:

$$O = \left\langle T, C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i} \right\rangle, i = \overline{1,t},$$

where t is a number of the ontology contexts, $T = \{T_1, T_2, \ldots, T_n\}$ is a set of ontology contexts, $C^{T_i} = \left\{C_1^{T_i}, C_2^{T_i}, \ldots, C_n^{T_i}\right\}$ is a set of ontology classes within the i-th context, $I^{T_i} = \left\{I_1^{T_i}, I_2^{T_i}, \ldots, I_n^{T_i}\right\}$ is a set of ontology objects within the i-th context, $P^{T_i} = \{P_1^{T_i}, P_2^{T_i}, \ldots, P_n^{T_i}\}$ is a set of ontology classes properties within the i-th context, $S^{T_i} = \{S_1^{T_i}, S_2^{T_i}, \ldots, S_n^{T_i}\}$ is a set of ontology objects states within the i-th context, $F^{T_i} = \{F_1^{T_i}, F_2^{T_i}, \ldots, F_n^{T_i}\}$ is a set of the logical rules fixed in the ontology within the i-th context, R^{T_i} – is a set of ontology relations within the i-th context defined as:

$$R^{T_i} = \{R_C^{T_i}, R_I^{T_i}, R_P^{T_i}, R_S^{T_i}, R_F^{T_i}\},$$

where

$R_C^{T_i}$ is a set of relations defining hierarchy of ontology classes within the i-th context,

$R_I^{T_i}$ is a set of relations defining the 'class-object' ontology tie within the i-th context,

$R_P^{T_i}$ is a set of relations defining the 'class-class property' ontology tie within the i-th context,

$R_S^{T_i}$ is a set of relations defining the 'object-object state' ontology tie within the i-th context,

$R_F^{T_i}$ is a set of relations generated on the basis of logical ontology rules in the context of i-th context.

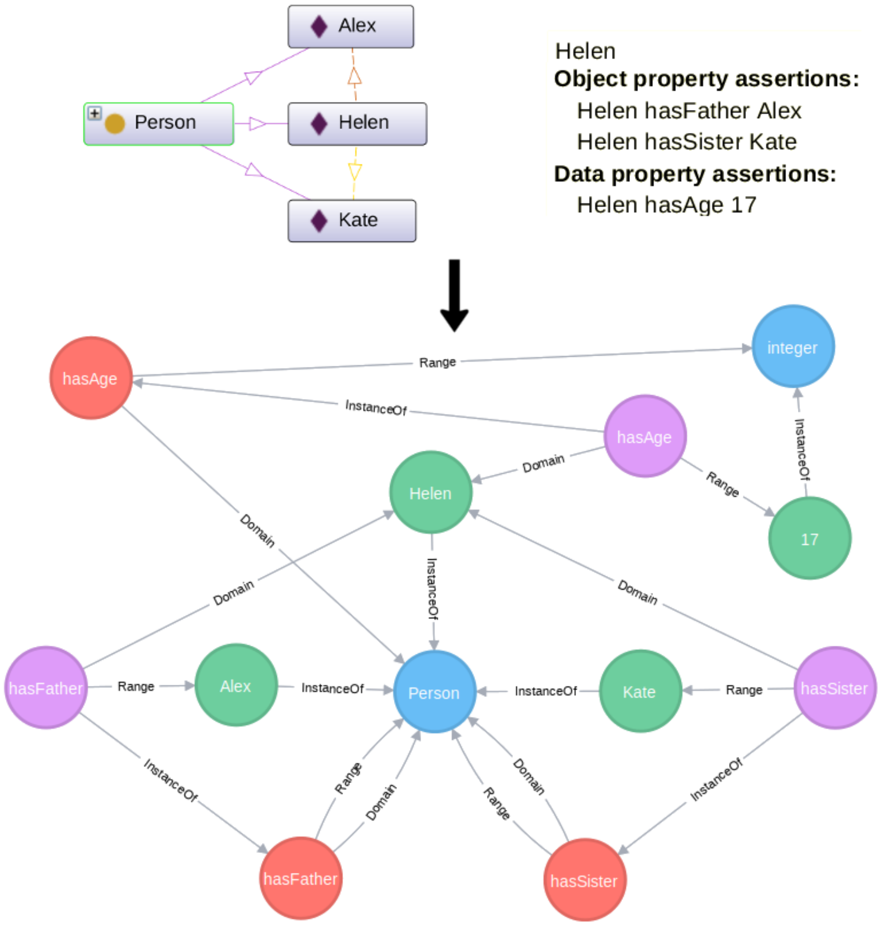

Figure 2. Example of the translation of the owl representation of ontology of family relations into the ontology of the fuzzy KB.

Principles similar to the paradigm of object-oriented programming are at the basis of the ontology of the fuzzy knowledge base:

- ontology classes are concepts of the PrA;
- classes can have properties, the child-class inherits properties of the parent class;
- objects of ontology describe instances of the concepts of the PrO;
- specific values for the properties of objects inherited from the parent class are determined by the states;

- logical rules are used to implement the functions of inference by the content of fuzzy KB.

3. The inference on the contents of fuzzy KB

The inference is the process of reasoning from the premises to the conclusion. Reasoners are used to implement the function of inference. Reasoners form logical consequences on the basis of many statements, facts and axioms [12][10]. The most popular at the moment reasoners are:

- Pellet;
- FaCT ++;
- Hermit;
- Racer, etc.

These reasoners are actively used in the development of intelligent software. However, Neo4j does not assume the possibility of using similar default reasoners. Thus, there is a need to develop a mechanism for inference based on the content of a fuzzy KB.

Currently the Semantic Web Rule Language (SWRL) is used to record logical rules.
These SWRL rules describe the conditions under which object *a* has «nephew-uncle» relation with object *c*. Formally the logical rule of the ontology of the fuzzy knowledge base is:

$$F^{T_i} = \left\langle A^{Tree}, A^{SWRL}, A^{Cypher} \right\rangle,$$

where T_i – i -th context of the ontology of the fuzzy KB; A^{Tree} – a tree-like representation of a logical rule F^{T_i}; A^{SWRL} – SWRL representation of the logical rule F^{T_i}; A^{Cypher} – Cypher representation of the logical rule F^{T_i}.

The tree-view A^{Tree} of a logical rule F^{T_i} is:

$$A^{Tree} = \left\langle Ant, Cons \right\rangle,$$

where $Ant = Ant_1 \Theta Ant_2 \Theta \ldots Ant_n$ – is the antecedent (condition) of the logical rule F^{T_i}; $\Theta \in \{AND, OR\}$ – is a set of permissible logical operations between antecedent atoms; $Cons$ – consequent (consequence) of a logical rule F^{T_i}.

Figure 3 shows an example of a tree-like representation of two logical rules for the ontology of family relations. That rules describes the father-child relationships.

The tree-like logical rule is translated into the following SWRL:

```
hasFather(?a,?b) => hasChild(?b,?a)
hasSister(?c,?a) & hasFather(?c,?b) => hasChild(?b,?a)
```

and the following Cypher view:

```
MATCH (s1:Statement{name: "hasChild", lr: true})
MATCH (r1a)<-[:Domain]-(:Statement{name:"hasFather"})-[:Range]->(r1b)
MERGE (r1b)-[:Domain]->(s1)
MERGE (r1a)-[:Range]->(s1)
MATCH (s1:Statement{name: "hasChild", lr: true})
MATCH (r2c)<-[:Domain]-(:Statement{name:"hasSister"})-[:Range]->(r2a)
MATCH (r2c)<-[:Domain]-(:Statement{name:"hasFather"})-[:Range]->(r2b)
MERGE (r2b)-[:Domain]->(s1)
MERGE (r2a)-[:Range]->(s1)
```

Thus, the rules are translated into their tree-view when imported into the KB of logical rules in the SWRL language.

The presence of a tree-like representation of a logical rule allows to form both a SWRL-representation of a logical rule and a Cypher-representation based on it.

Relations of a special type are formed by using Cypher to represent the logical rule between entities of the ontology of the fuzzy KB. Figure 4 shows the content of fuzzy KB after executing the Cypher queries that were built for the logical rule shown in Figure 3. These relations correspond to the antecedent atoms of the logical rule. Formed relationships provide the inference from the contents of the fuzzy KB.

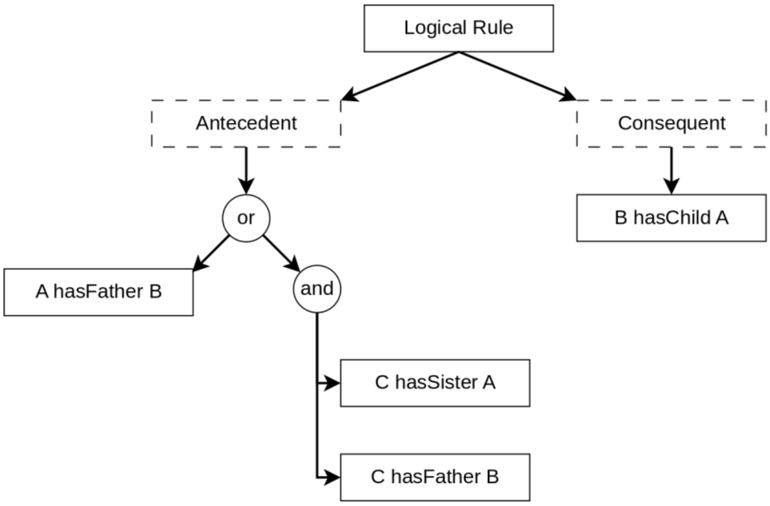

Figure 3. Example of a tree-like representation of a logical rule.

4. Building a Graphical User Interface (GUI) based on the contents of a fuzzy KB

The dynamic graphical user interface (GUI) mechanism is used to simplify the work with KB of untrained users and control of user input [14].

You need to map the fuzzy KB ontology entities to the GUI elements to build a GUI based on the contents of the fuzzy KB. Formally, the GUI model can be represented as follows:

$$UI = \langle L, C, I, P, S \rangle, \tag{1}$$

where $L = \{L_1, L_2, \dots, L_n\}$ is a set of graphical GUI components (for example, ListBox, TextBox, ComboBox, etc.); $C = \{C_1, C_2, \dots, C_n\}$ – is a set of ontology classes; $I = \{I_1, I_2, \dots, I_n\}$ – is a set of ontology objects; $P = \{P_1, P_2, \dots, P_n\}$ – is a set of properties of ontology classes; $S = \{S_1, S_2, \dots, S_n\}$ – is a set of states of ontology objects of fuzzy KB.

The following function is used to build a GUI based on fuzzy KB:

$$F(O) = \{C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i}\} \to \{L, C, I, P, S\},$$

where $\{C^{T_i}, I^{T_i}, P^{T_i}, S^{T_i}, F^{T_i}, R^{T_i}\}$ is a set of ontology entities of fuzzy KB represented by expression 1 within the *i*-th context; $\{L, C, I, P, S\}$ is a set of GUI entities of fuzzy KB represented by the expression 2.

Thus, the contents of the fuzzy KB are mapped to many GUI components. This makes it easier to work with KB for a user who does not have skills in ontological analysis and knowledge engineering. It also allows you to monitor the logical integrity of the user input, which leads to a reduction in the number of potential input errors.

5. Conclusion

Thus, the use of fuzzy KB stored in the Graph DBMS in the decision support process presupposes the existence of a certain set of mechanisms:

- organization of inference on the content of fuzzy KB by translating SWRL-rules into Cypher-structures;
- building a graphical user interface based on the contents of fuzzy KB;
- automated import of knowledge from internal and external wiki-resources.

These mechanisms allow to automate the learning process of KB and simplify the work of specialists with KB.

The application of a contextual approach to the storage of knowledge raises the effectiveness of the use of subject ontologies, allowing to adapt the KB to the characteristics of the PrA and to the requirements of specialists. This approach provides them with a tool that is convenient in a software dynamically changeable depending on the contents of the KB.

Figure 4. The result of executing Cypher queries for logical rule.

6. Acknowledgments

This work was financially supported by the Russian Foundation for Basic Research (Grant No. 16-47-732054).

7. References

[1] Bova, V.V. Problems of representation of knowledge in integrated systems of support of administrative decisions / V.V. Bova, V.V. Kureychik, Ye.V. Nuzhnov // Izvestiya YUFU. – 2010. – Vol. 108(7). (in Russian).

[2] Chernyakhovskaya, L.R. Intellectual decision support in the operational management of business processes of the enterprise / L.R. Chernyakhovskaya, N.I. Fedorova, R.I. Nizamutdinova // Bulletin of the USATU. Management, Computer Science and Informatics. – 2011. – Vol. 15. – N 42(2). (in Russian).

[3] Vagin, V.N., Mikhaylov I.S. Development of programs for the integration of information systems based on metamodelling and domain ontology / V.N. Vagin, I.S. Mikhaylov // Software products and systems. – 2008. – Vol. 1. (in Russian).

[4] Gavrilova, T.A. Ontological approach to knowledge management in the development of corporate information systems // News of Artificial Intelligence. – 2003. – Vol. 2. (in Russian).

[5] Zagorul'ko, YU.A. Building portals of scientific knowledge on the basis of ontology / YU.A. Zagorul'ko // Computing technologies. – 2007. – Vol. 12(S2). (in Russian).

[6] Karabach, A.Ye. Information integration systems based on semantic technologies / A.Ye. Karabach // Science, technology and education. – 2014. – Vol. 2(2). (in Russian).

[7] Smirnov, S.V. Ontological modeling in operational management / S.V. Smirnov // Ontology of design. – 2012. – Vol. 2. (in Russian).

[8] Tuzovskiy, A.F. Development of a knowledge management system based on a single ontological knowledge base / A.F. Tuzovskiy // Proceedings of Tomsk Polytechnic University. – 2007. – Vol. 310(2). (in Russian).

[9] Filippov, A.A. Single Ontological Data Mining Platform / A.A. Filippov, V.S. Moshkin, D.O. Shalayev, N.G. Yarushkina // Proceedings YOU international scientific and technical conference OSTIS-2016. – Minsk, 2016. (in Russian).

[10] Moshkin, V.S. Methods for constructing fuzzy ontologies of complex subject domains / V.S. Moshkin, N.G. Yarushkina // Proc. 5 international. scientific-techn. conf. OSTIS-2015. – Minsk, 2015. – P. 401-406. (in Russian).

[11] Tarasov, V.B. Granular, fuzzy and linguistic ontologies to provide mutual understanding between cognitive agents / V.B. Tarasov, A.P. Kalutskaya, M.N. Svyatkina // Proc. 2 international. scientific-techn. conf. «OSTIS-2012». – Minsk, 2012. – P. 267-278. (in Russian).

[12] Makhortov, S.D. On the Algebraic Model of a Distributed Productive System / S.D. Makhortov // Proc. 15 nat. Conf. on AI with intern. concert. – Smolensk. – 2016. – Vol. 1. – P. 64-72. (in Russian).

[13] Moshkin, V.S. Logical inference based on fuzzy ontology / V.S. Moshkin, N.G. Yarushkina // Integrated models and soft computations in artificial intelligence. Collection of scientific works of the VIII-i International Scientific and Practical Conference. – 2015. – Vol. 1. – M.: Fizmatlit. – P. 259-267. (in Russian).

[14] Gribova, V.V. Process and solution management / V.V. Gribova, A.S. Kleshchev. – 2006. – Vol. 2. (in Russian).

[15] Shestakov, V.K. Development and maintenance of information systems based on ontology and Wiki-technologies / V.K. Shestakov // Proceedings of the 13th All-Russian Scientific Conference "Digital Libraries: Advanced Methods and Technologies, Digital Collections" - RCDL'2011. – Voronezh, 2011. (in Russian).

[16] Suchanek, F.M. YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia / F.M. Suchanek, G. Kasneci, G. Weikum // Proceedings of the 16th International Conference on World Wide Web. – USA, 2007.