Proceedings of the International Multiconference on
Computer Science and Information Technology pp. 67–72
ISBN 978-83-60810-27-9
ISSN 1896-7094

Reasoning in RDFgraphic formal system with quantifiers

Alena Lukasová
Ostravská Univerzita v Ostravě,
Přírodovědecká fakulta, katedra informatiky a počítačů. 30. dubna
22, Ostrava, Czech Republic.
Email: alena.lukasova@osu.cz

Marek Vajgl
Ostravská Univerzita v Ostravě,
Přírodovědecká fakulta, katedra informatiky a počítačů. 30. dubna
22, Ostrava, Czech Republic.
Email: marek.vajgl@osu.cz

Martin Žáček
Ostravská Univerzita v Ostravě,
Přírodovědecká fakulta, katedra informatiky a počítačů. 30. dubna
22, Ostrava, Czech Republic.
Email: martin.zacek@osu.cz

***Abstract*—Both associative networks and RDF model (here we consider especially its graph version) belong to formal systems of knowledge representation based on concept-oriented paradigm. To treat properties of both of them as common properties of the systems is therefore natural. The article shows a possibility to use universal and existential quantified statements introduced prior to associative networks also within RDF graphic system and to define a RDF formal system with extended syntax and semantic that can use inference rules of associative networks. As an example solution, a logical puzzle is presented.**

I. Introduction

ABSTRACT RDF metamodel forms a framework and approach how to express a piece of knowledge by means of a directed, labeled graph that links relevant resources. All entities described by RDF expressions have to be treated as resources that are identified by resource identifiers - URIs. According to the metamodel RDF graph notation RDF statement consists of two labeled nodes (subject and object) linked together by a labeled edge (predicate). All of the labels of subject or predicate are represented by names as URIs, in the case of object it is also possible to use literal value. RDF in its second part RDFS also has an ability to define vocabularies (new terms) for practical use in RDF statements to specify kinds or classes of resources with specific attributes.

The graph view is the easiest possible and easy-to-understand visual explanations of RDF statements.

A similar idea as in the case of RDF occurred several years ago at the beginning of knowledge representation approach by means of association (semantic) networks. The authors of the idea tried to express semantics of knowledge about a concrete topic by a certain context framework represented graphically in the form of a binary labeled network (see for example [1]).

In relation to the first order logic associative networks as well as RDF graphs have been built on atomic vectors - elementary network statements, represented in the first order predicate logic by binary predicates:
⟨ predicate_symbol ⟩ (⟨attribute_1⟩, ⟨attribute_2⟩), or graphically:

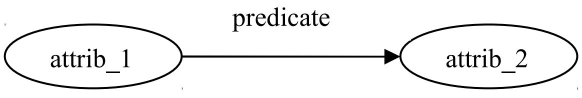

Fig.1 Associative network

In our paper [2], we have treated the RDF modeling as a special case of knowledge representation in associative networks (without functions). This fact brought a possibility to use notification and inference methods of associative networks also in the frame of RDF modelling. Moreover it has given a possibility to see RDF as a formal system.

The formal system of RDF(S) ought to be defined by its language, special axioms used to represent a knowledge base, and inference rules as a ground for creating special theories by means of formal proofs. The system ought to be presented as a formal system corresponding to predicate logic.

II. Extended Language Syntax of the RDF(S) Formal System

Our definition of RDF(S) language corresponds to the document W3C [3]. Authors of the document define names (node labels) as URIs or literals or *blank nodes* expressing the fact that there exists a URI reference making the statement of the triple true.

Graphs without blank nodes are ground RDF graphs.

According to [3], the RDF language does not contain a mechanism to represent universal or existential quantification. However, first order logic formulas can be transformed to special clausal form, where all individual variables (previously bounded to existential quantification) are eliminated by skolemization, and in this form (without quantification symbols) marked with special characters as existential terms. Universally quantified variables are (represented without quantification) conceived as variables of a universal character.

For the sake of expressing inference rules in a manner corresponding to the clausal form of the first order logic we add a metalanguage extention to the original RDF graphic language:

1. We extend the set of node names of the graphic RDF language by:

   a. existential metasymbols (strings with the @ at the beginning) that give a possibility within a statement to express an existence of a name in subject or object label that satisfies the corresponding triple. (The original definition of the RDF language recommendation [3] do not suppose the necessity to distinguish existential symbols in different statements or different nodes and solves the problem universally by means of blank nodes only);
   b. universal variable metasymbols (strings with a capital at the beginning) that give a possibility within a graph or its special part to express its universal properties.
2. We introduce a new kinds of RDF graphs that can express conditions "if -then". So we use within knowledge bases similarly as in association networks graphs the following kinds of networks:
   a. *Unconditional networks:*
      i. *ground* representing facts, containing purely constant node names, without variable or existential metasymbols
      ii. *universal/existential* with an occurrence of a universal/existential metasymbol
   b. *Conditions* (*rules*) - represent rules: q if p_1, p_2 ,...p_n with antecedent p_1, p_2 ,...p_n and consequent q. Edges in an antecedent part of the rule are represented via dash-line, consequent is represented via solid line, as usual. Conditions can have more than one atom in antecedent, but exactly one atom in consequent. Again, they can be divided into:
      i. ground
      ii. universal/existential

Network with a node label containing universal/existential metasymbol is a universal/existential network; otherwise it is taken as a ground network.

III. 3 SEMANTICS OF THE EXTENDED RDF(S) LANGUAGE

Name in RDF(S) is a URI-reference or a literal (possibly ordered to a URI-reference). A set of used names of a language defines its vocabulary.

As the RDF graphs are described by RDF syntax triples, we have to consider an interpretation of the RDF language relative to a corresponding vocabulary, it means relative to a set of names (identifiers) - a set of node's and arcs' labels.

Resources corresponding to the URI-reference set of a vocabulary form a *universe of discourse* IR of an interpretation I.

To be able to use interpretation rules of a language, we need to have a *structure* of the interpretation including the universe of discourse IR that orders in the model theoretic semantics of a language:

- an object of the universe of discourse to each of the name in the model,
- a truth value to each of the statement of the model.

To define interpretation of our extended language RDF(S), we follow a two-part process corresponding to that one of syntax definition of the RDF(S) language:

- rdf-interpretation (rdfs-interpretation) concerning interpretation of the rdf vocabulary (rdfs vocabulary) that orders an object of the universe of discourse to each of the names in the model;
- rules of interpretation as methods how to derive truth values of basic statements (event. obtained from universal/existential statements) and consequently a truth value of the whole RDF graph.

Interpretation rules according to [6] (slightly modified and extended):

A simple interpretation ***I*** of a vocabulary **V** is defined by the structure of interpretation ***I*** as follows:

1. A non-empty set IR of resources, called the domain or universe of ***I***.
2. A set IP $\subseteq$ IR - the set of resources of IR corresponding to properties.
3. A mapping IEXT from IP into the powerset of IR x IR i.e. the set of sets of pairs <x,y> with x and y in IR.
4. A mapping IS from URI references in V into IR.
5. A mapping IL from typed literals in V into IR.
6. A distinguished subset LV of IR, called the set of literal values, which contains all the plain literals in V.

Points 4. - 6. concern interpretation of names. In this cases interpretation values are elements of the universe of discourse. The interpretation assigns to each URI reference its corresponding resource, to literals it assigns them.

The mapping of the point 3 concerns interpretation of a property (predicate) of a ground triple E:

If E is a ground triple (⟨subject⟩, ⟨predicate⟩, ⟨object⟩), then ***I***(E) = true if ⟨subject⟩, ⟨predicate⟩ and ⟨object⟩ are in V, ***I***(⟨predicate⟩) is in IP and the pair (***I***(⟨subject⟩), ***I***(⟨object⟩)) is in IEXT(***I***(⟨predicate⟩)), otherwise ***I***(E)= false.

It means in this case an interpretation I denotes a truth-value to a ground triple E: If E is a ground graph then ***I***(E) = false iff ***I***(E′) = false for some triple E′ in E, otherwise ***I***(E) = true.

Interpretation also has to specify the truth-value of a graph containing nodes with existential symbol labels.

If ***I*** is interpretation and ***A*** is a mapping from the set of nodes labeled by existential symbols to the universe IR of ***I*** then the interpretation [***I***+***A***](E) orders to every triple E containing an existential symbol label a ground triple via the mapping ***A*** and orders to the triple or to the whole graph a truth value as stated above.

Triples having universal metasymbols as node names give us a possibility to treat them as universal variables, to make valuations of the variables and then to order a truth value to the resulting ground triple (graph) as stated above.

Conditions representing rules q if p_1, p_2 ,...p_n hold universally.

Example 1 - representation of condition (with quantifiers)

“Cheetah belongs to order the Carnivore if it eats meat.”

We can rewrite the sentence as a universal conditional network. “Every animal belongs to the order of Carnivore if it eats meat.”

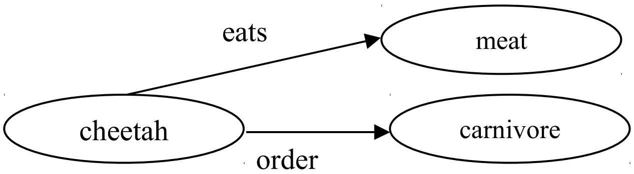

Fig.2 Unconditional network

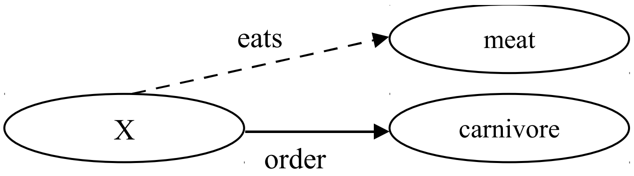

Fig.3 Conditional network

Using existential symbols we can express the sentence “Animal belongs to the order of Carnivore if there exists a meal that is meat and the animal eats it.”

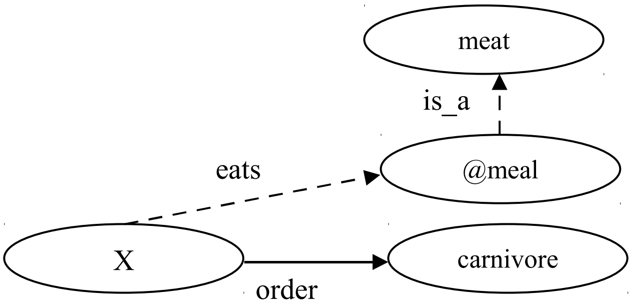

Fig.4 Extended network

IV. Inference in RDF

Deduction in RDF similarly to that in an associative network is a process of filling triples or terms in labels of the nodes from a *particular network* into a *main network* using one of the following rules:

Uniform substitution rule

a) If all the edges of a particular network also appear in the main network containing universal symbols as labels of their connected nodes, then the labels of nodes of the particular network can be substituted uniformly to the corresponding labels of nodes of the main network labeled by universal symbols.

b) If all the edges of a particular network also appear in the main network containing existential symbols as labels of their connected nodes and a maping *A* from the set of nodes labeled by existential symbols to the universe IR of ***I*** has been found then substitute uniformly corresponding elements of IR for existential symbols via the maping ***A***.

Transfer rule

If all dashed edges of the antecedent of a particular network appear (as solid) in the main network, and moreover if the labels of the corresponding nodes are the same then the solid vector of consequent can be added into the main network.

Negation of a triple as a special condition

To negate the triple E of RDF a special statement is used. It stands for the false consequent in the implication with a true antecedent. Due to no existence of a definition for a contradiction triple, a new special symbol of network has been created, called falsum, (notation ⊗), which is false in all interpretations.

Conditions, which consequent is falsum, are called negative facts.

Figure no. 8 represents negative fact „It is not true that dragonfly lives in Iceland. “

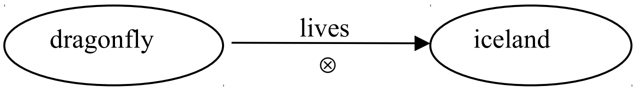

Fig.5 Negated associative network

Example 2 (logical puzzle)

We have 4 animals that live somewhere else and eat something else. Moreover:

1. The cheetah lives in Africa.
2. An animal that eats the flying insect lives in Europe.
3. The falcon doesn’t live in the Atlantic.
4. The dragonfly doesn’t hunt small fish.
5. The cheetah eats the mostly mammals.
6. The swordfish doesn’t live in Europe.
7. The falcon doesn’t eat the flying insects.
8. The dragonfly doesn’t live in the Iceland.
9. An animal that lives in the Atlantic doesn’t eat little birds.

Where does the falcon live and what does it eat?

Universe of discourse IR consists of all the resources used within solving the puzzle. Form of a list of pairs (graph label, URI of corresponding resource) is shown in tab 1.

Tab.1
URIs

mammal	http://dbpedia.org/resource/Mammal
bird	http://dbpedia.org/resource/Bird
fish	http://dbpedia.org/resource/Fish
insect	http://dbpedia.org/resource/Insect
cheetah	http://dbpedia.org/resource/Cheetah
falcon	http://dbpedia.org/resource/Peregrine_Falcon
swordfish	http://dbpedia.org/resource/Swordfish
dragonfly	http://dbpedia.org/resource/Dragonfly
africa	http://dbpedia.org/resource/Africa
iceland	http://dbpedia.org/resource/Iceland
atlantic	http://dbpedia.org/resource/Atlantic_Ocean
europe	http://dbpedia.org/resource/Europe
eat	http://en.wikipedia.org/wiki/Eat
live	http://en.wikipedia.org/wiki/Living

At the beginning we can imagine the main network graph consisting of four connected subgraphs with mutualy different existential node labels (fig. 5).

Next, we can define particular networks representing sentences 1 - 9 (as all the subgraphs are recognizable, the existential labels can be omitted) (fig 6).

Searching a common solution (model) of the main network and all of the nine participating particular networks consists in sequential binding of the label names occurring in particular networks (names of objects of the universe of discourse) to a set of existential nodes symbols (here blank nodes) in the main network until no contradiction is obtained and the proper mapping has been found.

Partially, rules 1. and 5. (similarly the case 4. and 8. and the case 3. and 7.) with the common cheetah label give a possibility to find a mapping within two subgraphs and make a corresponding substitution of labels (fig. 7, 8 and 9).

We can rewrite rules with negation (here for example 11 and 12) by RDF:bag. As an example, for a subgraph "dragonfly" holds a universal (within the puzzle) condition (fig. 10 and 11).

By applying these steps, we come to the conclusion, who lives where and what eats (fig. 12).

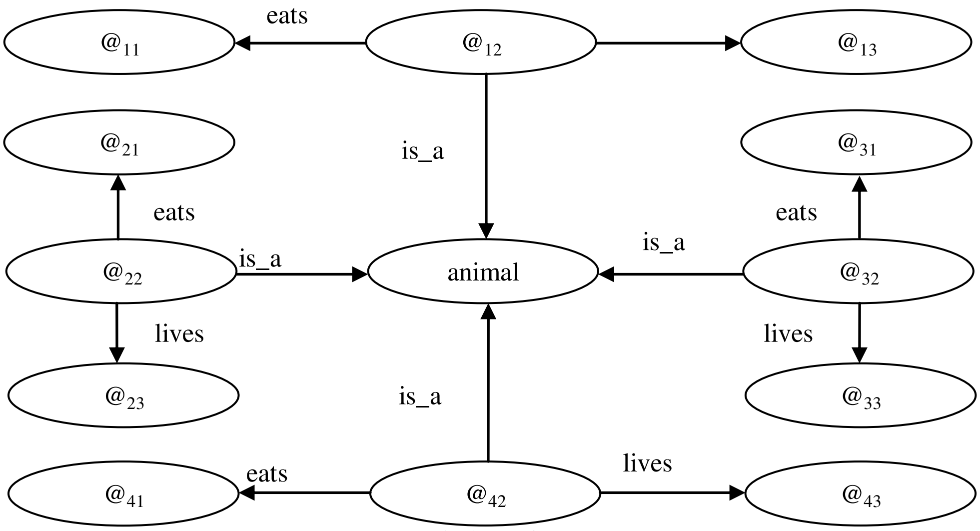

Fig.5 Base fact

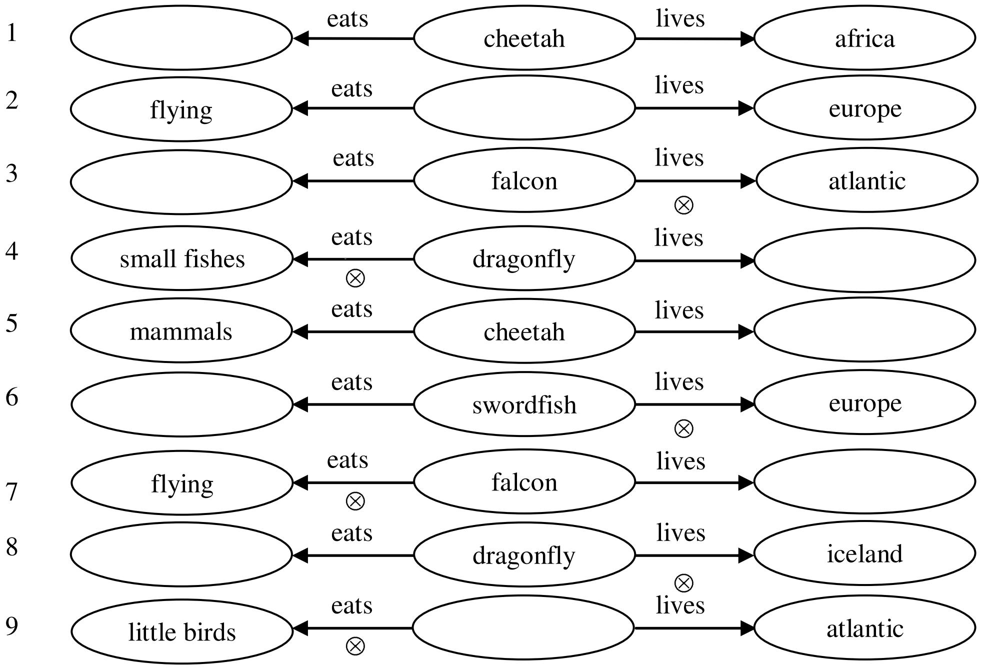

Fig.6 Extending facts

Fig.7 Rules over networks 1 and 5.

Fig.8 Rules over network 4 and 8.

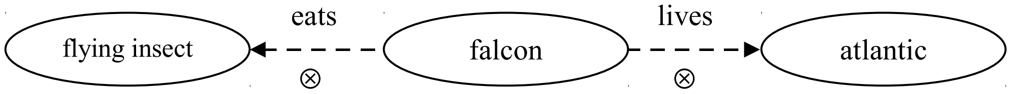

Fig.9 Rules over network 3 and 7.

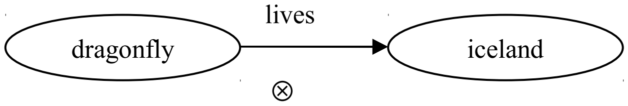

Fig.10 Negative facts

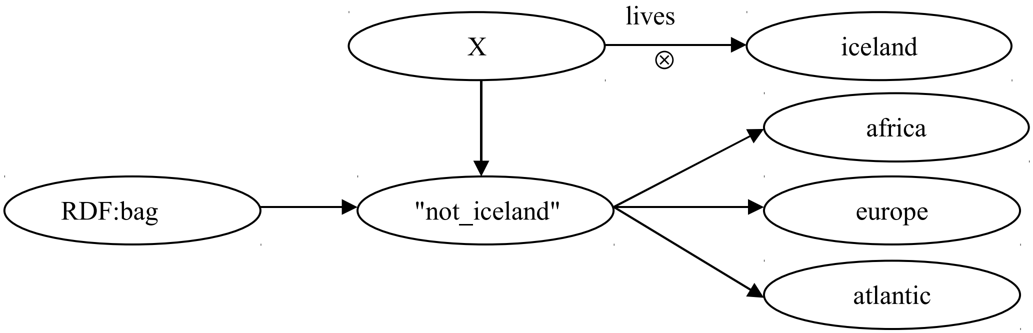

Fig.11 Negative facts

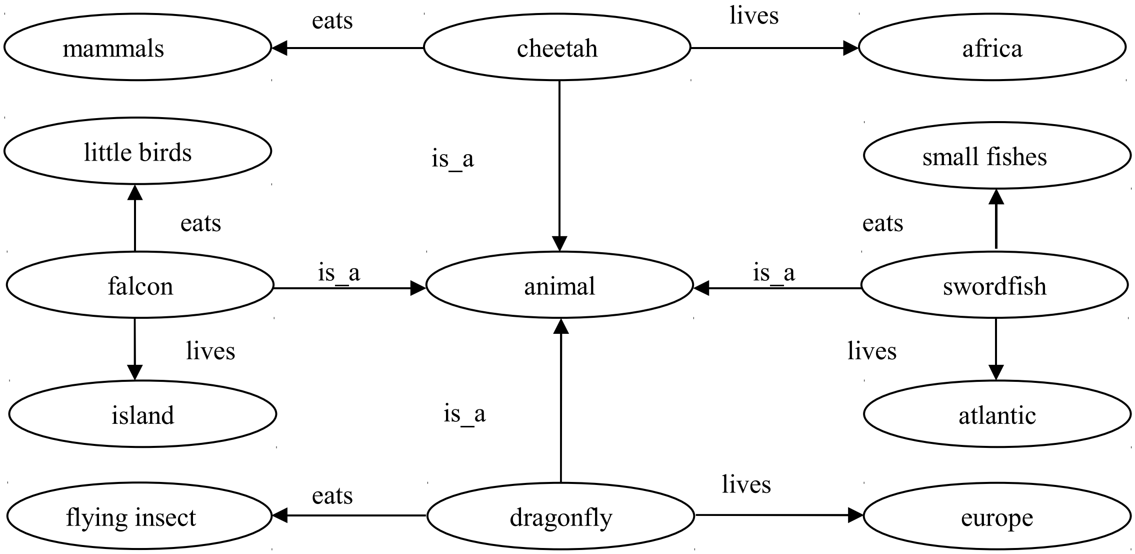

Fig.12 Extended network

V. Conclusion

We can treat the RDF modeling as a special case of knowledge representation in associative networks (without functions). This fact offers us a possibility to use the notification and inference rules of associative networks also in the frame of RDF modeling. Reasoning, possibly supported by the graphical version of representation, then becomes more understandable and easier to use than that of rewriting RDF models by description logic tools (OWL language) because of using their inference mechanisms. However, some issues like how negative and positive vectors are matched together are not fully formalized. Although article presents a mechanism (and an example) how can be semantic network used for inference over RDF models via graphical way affording better understandability, especially for human.

REFERENCES

[1] Lukasová, A.: Reprezentace znalostí v asociativních sítích.Proc. Znalosti 2001 (in Czech).
[2] Vajgl, M., Lukasová, A.: RDF model as Associative Network. Datakon 2009, ISBN 978-80-245-1568-7.
[3] W3C: Resource description Framework (RDF). RDF Primer http://www.w3.org/TR/2004/REC-rdf-primer-20040210/
[4] W3C: RDF Vocabulary Description Language 1.0: RDF Schema, 2004 http://www.w3.org/TR/rdf-schema/
[5] W3C: Status for Resource Description Framework (RDF) Model and Syntax Specification, Recommendation 1999, http://w3.org/TR/REC-rdf-syntax.
[6] W3C: RDF Semantics. Recommendation 10 February 2004. http://www.w3.org/TR/2004/REC-rdf-mt-20040210/