Resolution reasoning by RDF Clausal Form Logic

Alena Lukasova, Martin Zacek, Marek Vajgl, Martin Kotyrba

Department of Informatics and Computers, University of Ostrava
Ostrava, Czech Republic

Abstract
The article presents a simple way how knowledge represented via RDF triples provides an easier method of inference, managing and finding interrelations between knowledge objects than that an approach based on OWL language. The authors of the article come out of the T. Richard's inference system of the Clausal Form Logic (CFL) based on the formal manipulation with conditional „if – then" statements, and an idea of RDF extended model (with quantifiers) of knowledge representation to propose an inference mechanism RDF RR working over knowledge bases of RDF triples.
Keywords: *RDF extended model, CFL, Clausal Form Logic, graph.*

1. Introduction

There are a lot of approaches in AI, presenting solutions how knowledge can be managed and used for deductions of new knowledge. However, a lot of those approaches are built on requests of semantically rich language for knowledge capturing (like OWL) and powerful inference mechanism typically built on some kind of description logic. There are also many resources collect in simple relation information captured by simple knowledge representing tools, typically as pairs of objects bounded with some relations in a RDF triple.

For RDF model, which is typically aimed towards the conceptual level of knowledge representation, there are not very many approaches resolving an inference over such knowledge bases. Here we show one of them – with the usage of Clausal Form Logic (CFL) we present a resolution reasoning (RR) mechanism how the knowledge base using a format of RDF triples can be handled.

At the beginning we had a starting state composed of two distinct qualities:

1. an inference system of a Clausal Form Logic of T. Richards [1] based on the formal manipulation with conditional „if – then" statements; it belongs to the versions of the first order logic (FOL) mostly used within computer science as a method of formal reasoning,
2. an idea of RDF extended model (with quantifiers) of knowledge representation [2].

In our paper we would like to show that CFL based on extended RDF-triple knowledge representation could lead to an easy-to-use method of resolution reasoning within the space of RDF triple knowledge representation, Linked Data inclusive.

To realize our proposed RDF resolution reasoning (RDF RR) method, it is necessary to fulfil a further precondition as a third starting component. It consists in a clausal form style of knowledge representation. It means:

If the RDF RR method ought to be realized, it is necessary

3. to create a special "RDF-clausal form space" of the stored knowledge.

2. Richards´ CFL and a corresponding extent of RDF model

In the CFL language a general clause is of a form

<antecedent> → <consequent>,

one of them possibly could be empty,

or

$$P1\&\ldots\&\ Pm \rightarrow Q1\vee\ldots\vee Qn$$

or

$$P1,\cdots, Pm \rightarrow Q1,\cdots, Qn$$

("→" is a meta-symbol "implies", the antecedent is a conjunction of some set of positive atoms of first order predicate logic {P1,⋯, Pm} and the consequent is a disjunction of another set of positive first order predicate logic atoms {Q1,⋯, Qn}.

To be usable in RDF triple representation format the first order predicate logic atoms in clauses have to be rewritten as follows:

In relation to the first order logic a RDF triple

(〈subject〉, 〈predicate〉, 〈object〉)

or graphically:

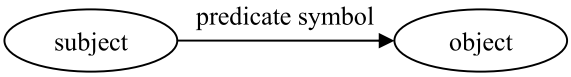

Fig. 1

has been built on the concept of an elementary statement - predicate atom represented in the first order predicate logic by binary predicate:

⟨ predicate_symbol ⟩ (⟨ subject ⟩, ⟨ object ⟩).

RDF model according to the document W3C [**Chyba! Nenalezen zdroj odkazů.**] does not contain a mechanism to represent universal or existential quantification in the subject or object parts of RDF triples. Authors specify within the graph based format only *blank nodes* expressing the fact that there exists a URI reference making the statement of the triple true. Graphs without blank nodes are then ground RDF graphs.

However, generally we are able to rewrite first order logic formulas into special clausal forms, where all the universally quantified individual variables are conceived as variables of universal character (now represented without quantifiers) and all the existentially (previously bounded to existential quantifier) are eliminated by skolemization, and in this form (without quantification symbols) marked with special characters as existential terms. This is the approach that uses also the Clausal Form Logic (CFL) [**Chyba! Nenalezen zdroj odkazů.**].

In order to a more expressive RDF model (to be able to express general or existential statements) we have undertaken the CFL convention to extend the set of RDF subject/object labels (previous only URIs or blanks). To express universal and existential quantified variables uses our RDF CFL format as well as the language of the CFL the following convention:

- Each of the variables within clauses be universal and labeled by a character chain with a capital letter at the first position.
- Skolem symbols as results of the former Skolem process within the transformation into clausal form are labeled by character chains with the @ symbol as a prefix.

Our RDF CFL format is now able to express statements like "Everybody likes anybody".

Fig. 2

3. Knowledge representation by RDF CFL clauses

T. Richards [1] proposed the version CFL of the First Order Logic (FOL) that corresponds to the mostly used conditional „if – then" statements. Generally, a conditional statement (clause) proposed by T. Richards says that some consequent statement composed as a disjunction of some predicate atoms follows from another antecedent statement composed as a conjunction of some predicate atoms.

The case of RDF format of representation we consider here under a name RDF CFL is a special case of the original Richard's CFL that shares all the properties of semantically sound and complete formal system with an extended RDF model (see above) by universal and existential terms. RDF CFL format uses the following syntax of *clauses* based on RDF-triple form of knowledge representation atoms:

- *Unconditional clauses (ground/universal/existential)* without antecedent/consequent representing positive/negative facts,
- *Conditional clauses* (*rules*) *(ground/universal/existential)*

Conditions represent rules (see (1))

$$Q \text{ if } P_1 \& P_2 \& ... \& P_n$$

with antecedent

$$P_1 \& P_2 \& ... \& P_n$$

and consequent

$$Q = Q_1 \vee Q_2 \vee ... \vee Q_m.$$

Atomic statements $P_1,..,P_m$, $Q_1,\ldots,Q_n$ form generally the following structure of the general CFL clause

$$P_1 \& \ldots \& P_m \rightarrow Q_1 \vee \ldots \vee Q_n$$

or simply

<antecedent> → <consequent>.

To represent RDF-triples RDF CFL we use here a pseudo-format (without URIs) of the form like

anne **knows** john, john **isa** student,…

On the base of RDF-triple representation of positive atoms, we have schemes (C1), (C2), (C3) of clauses (written in a special pseudo-format without URIs).

(C1):

```
clause
    antecedent
        < a set of atoms of the antecedent in the form of RDF-triples>
    /antecedent
    consequent
        < a set of atoms of the consequent in the form of RDF-triples>
    /consequent
/clause
```

Antecedent or consequent of the conditional clause could become also an empty set of atoms.

The clause (C2):

```
clause
    consequent
        anne knows john
    /consequent
/clause
```

(C3)

has a meaning of a positive fact “Anne knows John.”

(C3):

```
clause
    antecedent
        anne knows john
    /antecedent
/clause
```

represents a negative fact “It is not true that Anne knows John.”

The structure of the clause allows only constructions of clauses with connections & in the antecedent and connection ∨ in the consequent. If necessary, CFL solves as well as our RDF RR format the problem by a following decomposition of the clause into n (m) separate clauses

$$P_1 \rightarrow Q_1 \vee \ldots \vee Q_n \qquad\qquad P_1 \& \ldots \& P_m \rightarrow Q_1$$

$$\ldots\ldots \qquad\qquad \ldots\ldots$$

$$P_m \rightarrow Q_1 \vee \ldots \vee Q_n \qquad\qquad P_1 \& \ldots \& P_m \rightarrow Q_n$$

To prepare a knowledge base for our proposed RDF RR method of resolution reasoning the following three steps have to be made:

1) For the sake of expression of the clause in RDF-triples, all the atoms of first order logic have to be transformed into a corresponding binary version. To order of a unary predicate (for example **wise**(john), **wise**(X) – “John is wise”, “everybody is wise”)) into the RDF-triple format, the use of a concept of the binary predicate **isa**(<term1>, <term2>) with the meaning “is a” (**isa**(john, wise), **isa**(X, wise)) is possible. Now the corresponding RDF triple is of the form <term1> **isa** <term2> (john **isa** wise, X **isa** wise).
2) To use RDF extended model with quantifiers.
3) To express "if - then" condition, we introduce RDF CFL clauses (C1) having an antecedent and a consequent part. Clauses with terms containing universal/existential symbols are universal/existential clauses; otherwise they are taken as ground clauses.

For the ground clause holds the following *“rule of the transfer of negative ground atoms”*:

- If a negative ground atom ought to be ordered into the antecedent set of atoms, transfer it as a positive atom into the consequent set of atoms.
- If a negative ground atom ought to be ordered into the consequent set of atoms, transfer it as a positive atom into the antecedent set of atoms.
- Similarly to CFL, our RDF CFL format also applies a special *transfer rule* in the case of universal/existential clause. For example the clause “It is not true that Jane knows everybody, so somebody is a stranger of Jane.” changes after the transfer into a clause “Jane knows everybody or somebody is a stranger of Jane.”

4) Knowledge base of the RDF RR is a set of the RDF clauses in the form of (C1), (C2) or (C3).

4. Resolution reasoning on RDF RR knowledge bases

RDF RR works with the help of two rules - *the substitution rule* and *the cut rule*, both are well known in the first order logic. Together the two rules form as well as in the CFL, the *resolution rule.*

- The *substitution rule*:
  From a clause with variables we can obtain a new clause by a uniform substitution of a term for some of the variables.
- The *cut rule*:
  If the knowledge base contains two clauses sharing the same atom, once in the antecedent and once in the consequent, then we can obtain a new clause by cutting out the same atoms at both sides and create a new clause with an antecedent (consequent) that contains all the atoms of the original clauses antecedents (consequents).

In our paper [4] we have treated a special case of the RDF CFL modeling as the knowledge representation tool in the graph-based RDF notation.

The RDF-tiples are captured in the similar xml-element oriented way like the content of the „clause“ is. Due to lack of space the following examples presents rdf triples only in the way of sentences containing three words. In full formula, Theky are also decomposed into free element describing subject, predicate and object.

Example 1

Immigration rules of an EU country for citizens of other countries forms in the language of RDF RR a knowledge base that consists in the following set of clauses {1., 2., 3., 4.}.

1. <clause>
<antecedent>
X **citizen** Z, Z **isa** EUcountry, Y **isa** EUcountry
</antecedent>
<consequent>
X **may_enter** Y
</consequent>
</clause>
2. <clause>
<consequent>
anne **citizen** aus
</consequent>
</clause>
3. <clause>
<consequent>
aus **isa** EUcountry
</consequent>
</clause>
4. <clause>
<consequent>
@EUcountry **isa** EUcountry
</consequent>
</clause>

Clauses 1.- 4. – Prerequisites of the proof.
Proof of the statement „Anne may enter.“ on the knowledge base {1., 2., 3., 4.} in the language of RDF RR:

5. <clause>
<antecedent>
X **citizen** Z, Z **isa** EUcountry, @EUcountry **isa** EUcountry
</antecedent>
<consequent>
X **may_enter** @EUcountry
</consequent>
</clause> *1. after the substitution [@EUcountry/Z]*
6. <clause>
<antecedent>
X **citizen** Z, Z **isa** EUcountry
</antecedent>
<consequent>
X **may_enter** @EUcountry
</consequent>
</clause> *cut 4.,5.*
7. <clause>
<antecedent>
anne **citizen** aus, aus **isa** EUcountry
</antecedent>
<consequent>
anne **may_enter** @EUcountry
</consequent>
</clause> *6. after the substitution [anne/X, aus/Z]*
8. <clause>
<antecedent>
anne **citizen** aus
</antecedent>
<consequent>
anne **may_enter** @EUcountry
</consequent>
</clause> *cut 3.,7.*
9. <clause>
<consequent>
anne **may_enter** @EUcountry
positive statement as a result of inference
</consequent>
</clause> *cut 2.,8.*

Example 2

Indirect proof of the conclusion “If Chen has visa, he may enter country EU” if a knowledge base (prerequisites) consists in the following <clause>s 1. – 6..

1. <clause>
<antecedent>
X **citizen** Z, Z **isa** EUcountry, Y **isa** EUcountry, Z **needs visa** Y,
X **has visa** Y
</antecedent>
<consequent>
Z **isa** EUcountry, X **may_enter** Y
</consequent>
</clause>
2. <clause>
<consequent>
chen **citizen** china
</consequent>
</clause>
3. <clause>
<antecedent>
china **isa** EUcountry
</antecedent>
</clause>
4. <clause>
<consequent>
china **needs_visa** EUcountry
</consequent>
</clause>
5. <clause>
<consequent>
@EUcountry **isa** EUcountry
</consequent>
</clause>
6. <clause>
<antecedent>
chen **has_visa** @EUcountry
</antecedent>
</clause>
7. <clause>
<antecedent>
chen **may_enter** @EUcountry
</antecedent>
<consequent>
chen **has_visa** @EUcountry
</consequent>
</clause> *(clause to be proved - negated)*
8. <clause>
<antecedent>
chen **citizen** china, @EUcountry **isa** EUcountry,
china **needs_visa** @EUcountry,
chen **has_visa** @EUcountry
</antecedent>
<consequent>
china **isa** EUcountry, chen **may_enter** @EUcountry
</consequent>
</clause>
(1. after the substitution [chen/X, china/Z, @EUcountry/Y] and transfer china isaEUcountry)

9. <clause>
<antecedent>
chen **citizen** china, @EUcountry **isa** EUcountry, china **needs_visa** @EUcountry
</antecedent>
</clause> *(cut 8.,3., cut 8.,7)*

10. <clause>
*(inconsistent empty clause – **clash**)*
</clause> *(cut 2., 9., cut 5., 9., cut 4., 9.)*

5. Semantics of RDF CFL clauses

The formal system RDF RR has been defined above by its language RDF CFL, syntax of clauses ((C1), (C2), (C3)) used to represent knowledge bases, and resolution inference rules as a ground for creating special theories by means of formal proofs. The system ought to be presented as well as the CFL as a formal system corresponding to the first order predicate logic.

If all the variables of a RDF triple have been evaluated by constant terms, the triple becomes a *ground atom*. Truth value of the atom (triple) in an interpretation *I* then corresponds to the truth value of a corresponding predicate atom in the first order logic model-theoretic interpretation. It means:

A *ground atom* is *true* in an interpretation *I* iff there exists a pair of elements within the relation ordered in the *I* as a denotation to the triple's predicate that equals to a pair of constant terms of the ground atom. Otherwise it is *false*.

A *ground conditional clause* is *false* iff all of the antecedent are true and all the vectors of the consequent are false. Otherwise it is true.

In the case of a clause with an empty antecedent, the logical constant true stays for it and in the case of an empty consequent fulfils the role the logical constant false.

A conditional clause is *consistent* in an interpretation *I* given if there is a valuation of all the variables that makes the clause true.

6. Conclusions and furthers ideas

As each of the formal system ought to operate on knowledge bases because of obtaining new consequents and interrelations between them it is natural to find and develop a knowledge system that can manipulate with structured data in a straightforward way without a necessity of rewriting simple knowledge representing only relations between objects into a semantically more complex languages like OWL.

At present, realization of a Semantic Web idea (or Web3 idea to harness collective intelligence on the base of a special structure of linked pieces of knowledge) consists more or less in a concept of Web of Linked Data by means of RDF model knowledge representation. It means large scale integration of, and reasoning on, data on the Web or creating new interesting interrelations. Briefly, a knowledge system of realizing the Web3 idea of a Semantic Web has a goal:

- to create new and useful knowledge from multiple or large collections of data based on human contributions augmented by the RDF technology of structured data,
- to provide answers, solutions, discoveries or other results beyond the original data based on computation and inference.

To make the Web of Linked Data a reality, it becomes important to have the huge amount of "topic centred" data on the Web available in a common standard format RDF, reachable and manageable by Semantic Web tools.

A willingness to develop a knowledge system that would be able to manipulate with RDF structured data in a straightforward way without rewriting knowledge into a language like OWL is natural.

One of the possible approaches to reach the second goal mentioned above consists in a formal system of Resolution Reasoning based on RDF-clause based linked data.

A useful tool of resolution reasoning on linked knowledge would be based on the RDF RR formal system of reasoning that we have presented here.

Acknowledgments

The research described here has been financially supported by University of Ostrava grant SGS02/PřF/2012. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

References

[1] Richards, T.: Clausal Form Logic. An Introduction to the Logic of Computer Reasoning. Addison-Wesley, 1989.
[2] Lukasová, A., Vajgl, M., Žáček, M.: Reasoning in RDF graphic formal system with quantifiers.Proceedings of the International Multiconference on Computer Science and Information Technology. 2010. pp. 67-72.
[3] W3C: Resource description Framework (RDF). RDF Primer http://www.w3.org/TR/2004/REC-rdf-primer-20040210/
[4] Lukasová, A., Žáček, M. , Vajgl, M.: Reasoning in Graph-based Clausal Form Logic. IJCSI International Journal of Computer Science Issues. 2012, year 2012, volume 9, pp. 37-43.

Alena Lukasova has been working as a professor at the Department of Informatics and Computers at the University of Ostrava (Czech Republic). Her interests include theoretical principles of information and knowledge systems and tools of representation of semantically structured knowledge (database systems with knowledge bases components and their semantics),

formal deduction in concept oriented languages, formal ontology for information systems, principles of ontology driven (based) information systems, and sharable knowledge patterns. She is author more than 50 scientific publications.

Martin Zacek graduated at the University of Ostrava (Czech Republic). The focus of his dissertation is a formal deduction in graph knowledge representation systems. The PhD thesis includes four graph systems: semantic (associative) networks, conceptual graphs of Sowa, RDF model and Topic Maps. The topic of the PhD thesis corresponds to the content of this article. He is author more than 10 scientific publications. In 2010 he won award “Young Scientist” at the International Multiconference on Computer Science and Information Technology awarded by the International Fuzzy Systems Association Distinction for the presentation of article Reasoning in RDFgraphic formal system with quantifier.

Marek Vajgl graduated at the University of Ostrava (Czech Republic) and defended PhD. thesis with title “A proposal of tool for creating and updating knowledge bases for semantic web”. He has been working as a lecturer at the Department of Informatics and Computers, University of Ostrava. His interests include descriptive logic formalism which is frequently used as semantic web formalism. He is author more than 20 scientific publications.

Martin Kotyrba graduated at the University of Ostrava (Czech Republic). The focus of his dissertation is Artificial intelligence methods for automatic pattern recognition with fractal structure. The topic of his work at university corresponds to the content of this article. He is author more than 10 scientific publications.