

Towards a Practical Implementation of Contextual Reasoning on the Semantic Web

Sahar Aljalbout

Centre Universitaire d'informatique, University of Geneva, Switzerland
sahar.aljalbout@unige.ch

Abstract. Contextual knowledge representation and reasoning is an old issue in the semantic web. Despite the fact that context representation has for a long time been treated locally by many semantic web practitioners, a recognized and widely accepted consensus regarding the precise ways of encoding and even more reasoning on contextual knowledge has not yet been reached by far. In this dissertation, we introduce an approach to represent and reason over contextual knowledge in RDF, while committing to a formally defined semantics of a contextual description logic. Our key contribution is the definition of a formally solid contextual model (not only for contextual knowledge representation but also for contextual reasoning) which is practically applicable using existing semantic web languages and tools.

Keywords: contextual reasoning, contextual OWL, contexts...

1 Problem statement

The problem of representing and reasoning on contextual knowledge is a recognized, and open issue in the semantic web. Many data providers and semantic web practitioners have attempted local approaches for treating contexts representation; however, there is, so far, no consensus regarding the precise ways of encoding, and much less reasoning, on contextual knowledge. Nevertheless, the representation of contexts in the semantic web has been considered separately as, first, a data problem, giving rise to several proposals to encode contexts into RDF [18][10][20][1]. And second, as a theoretical problem where several attempts to include contexts in description logics have emerged [3][14]. However, contrary to the field of context representation where research is abundant, reasoning with contextual knowledge has been notably less explored.

This PhD proposal aims to strengthen the links between the theoretical and practical communities working on diverse aspects of representing and reasoning with contextual knowledge; which could hopefully lead one day to some forms of standardization or at least good practice guidelines accepted by the community. We propose an approach to represent and reason over contextual knowledge in RDF while committing to a formally defined semantics of a contextual description logic. The key idea behind this approach is the definition of a formally

solid contextual model but also practically applicable to data while using existing semantic web languages and tools. Throughout this work, we have adopted McCarthy's theory of contexts [17], primarily because this theory offers an instrumental view on contexts, where contexts are considered as formal objects, describable in first-order languages.

The preliminary results of this dissertation are the following:

- A survey of logical and practical models to encode contexts on the semantic web.
- A contextual extension of the web ontology language, that we called OWL^C .
- A contextual profile of the web ontology language inspired from $OWL-RL$, that we called $OWL-RL^C$, with contextual entailment rules for reasoning, not only on contextual statements, but also on contexts.
- A study of the practical implementation of the contextual reasoning.

2 Relevancy

The goal of the semantic web (GSW) is to promote machine-understandability of the information published on the web. For instance, if a dataset is published on the web by an arbitrary agent, it should ideally lend itself correctly interpretable by any other agent accessing it independently. However, in many situations, information cannot be fully understood without making explicit assumptions about the context in which it is stated. Therefore, if the web of data is not accompanied by a clear *standard* for the representation of contextual information, the goal of machine understandability can never be fully achieved.

Succeeding in positioning our proposed approach will benefit many stakeholders in the semantic web community. We are introducing a method to bridge the gap between the theoretical and practical communities. If proven successful, we will contribute to advance the field towards achieving the GSW as described above.

3 Related works

In 1969, McCarthy [17] proposed a theory of contexts which consists of three major postulates: 1) Contexts as formal objects. 2) Contexts having properties 3) Contexts organized in relational structures. Then, in 1993, F.Guinchiglia [11] discussed the concept of contextual reasoning considering reasoning always local to a subset of the known facts. More recent research can be divided in two groups: theoretical and practical.

In the theoretical group, in 2001, [9] introduced the idea of locality and compatibility where reasoning is considered mainly local and uses only part of what is potentially available. Compatibility is argued to be used among the reasoning performed in different contexts. In 2003, [5] introduced the concept of distributed description logics. The authors consider that there are binary relations that describe the correspondences. An advantage of DDLs is its support for multiple

ontologies. However, the coordination between a pair of ontologies can only happen with the use of bridge rules. In 2004, a new concept called E-connections [15] emerged: ontologies are interconnected by defining new links between individuals belonging to distinct ontologies. One major disadvantage is that it does not allow concepts to be subsumed by concepts of another ontology, which limits the expressiveness of the language. Then, in 2006, [3] attempted to extend description logics with new constructs with relative success. In 2011, a proposition was argued to use a two dimensional- description logics with a context language supporting context descriptions and an object language equipped with context operators for representing object knowledge relative to contexts. Results showed that this approach does not necessarily increase the computational complexity of reasoning. In 2012, [6] argues that treating contexts in the semantic web needs more advanced means, such that contexts should be explicitly presented and logically treated...

In the practical group, many attempts to find a solution to the syntactic restriction of RDF binary relations emerged, because an RDF property holding for a specific context is a relation involving three resources (a subject, an object, and a context). Three types of works were proposed:

- (a) Extending the data model: the triple data structure is extended by adding a fourth element to each triple, which is intended to express the context [7] of a set of triples.
- (b) Extending the semantic of RDF: In 2014, RDF* [12] was proposed. The idea is to extend the RDF data model with a notion of nested triples. Another approach is Singleton property [18] which recommends the creation of a special instance for every triple predicate for which we want to provide the context. A drawback of the singleton properties proposal is that it introduces a large number of unique predicates.
- (c) Using design patterns: Pat Hayes [13] presents ways to attach temporal indexing to sentences of the form $R(a,b)$. It could be categorized along three axis: 3D, 3D+1, 4D. We extend this categorization to many dimensions of contexts and use it to classify contextual patterns.
 - 3D representation: the contextual index co is attached to the sentence $R(a,b)$ and thus $R(a,b)$ holds for co such as RDF reification [4]. The major drawback of this method is that it is supported in DL reasoning.
 - 3D+1 representation: the contextual index co is attached to the relation $R(a,b,co)$. An example of this representation is situation pattern [8]. One advantage is being able to talk about assertions as (reifying) individuals, but the disadvantage is being unable to use them as properties. A second example is FluentRelations [1] [2]. Two advantages of this method are 1) considers contexts as objects; 2) don't cause objects proliferation.
 - 4D representation: the contextual index co is attached to the object terms $R(a@co, b@co)$ where co is the contextual-slice of the thing named. The first example is context slices [19]. This pattern introduces new entities: the contextual projections and assignments of the individuals as well as a context index that takes the indexes. Another example is Ndfloents

which is an extension of the 4dfuents model [20] for the temporal dimension to a generic contextual model. The drawback of this method is that it introduces many contextualized individuals which causes objects proliferation.

4 Research questions and hypotheses

We have formulated the following research questions:

Q1 *What are the practical requirements of contextual reasoning that are not fulfilled by the existing approaches and are essential for accomplishing a community consensus?*

Q2 *Can we significantly extend the web ontology language with contexts without increasing the computational complexity of the language? Is it possible to implement contextual reasoning using the existing reasoners?*

Q3 *To what extent linking the contexts by means of semantic relations¹ can enhance the expressiveness of the contextual language and push forward the discovery of hidden knowledge? And what are the drawbacks in terms of computational complexity?*

Q4 *What is the cost of the transformation of existing knowledge graphs to adhere to the proposed model? And how can we identify hidden contextual knowledge?*

Our hypotheses derive directly from the research questions:

H1 There are many ways to encode contexts on the graph level, yet we believe this is not enough to provide the semantic web with a contextualized state provided that reasoning on contexts is still an open problem. We hypothesize that adopting a two language approach with an object language and a context language will reduce the reasoning cost on the semantic web.

H2 The use of design patterns to encode the notion of contexts is more realistic than extending the RDF data model in the semantic web community. Although there is no best design pattern, each one is suitable for a specific target and dimension of contexts.

¹ Temporal contexts can be linked using Allens interval algebra (https://en.wikipedia.org/wiki/Allen%27s_interval_algebra), spatial contexts with RCC8 (https://en.wikipedia.org/wiki/Region_connection_calculus) etc.

5 Approach

In order to achieve a formally solid contextual model, but also practically applicable to linked data, we proceed as follows:

- To begin, we provided a contextual extension of the web ontology language and we called it OWL^C . This extension is based on a two-dimensional description logic [14] with one language for the representation of contexts-dependent concepts, roles, or axioms; and, a second language for the representation of contexts and their relations. The reasons behind this choice are: first, there is no additional cost in the complexity of reasoning², and second, the approach was designed to be applied to several practical scenarios.
- Then, we plan to define a generic upper vocabulary for describing contextual metadata and meaningful relations between contexts. If this vocabulary is adopted by the community, it can facilitate data interoperability.
- Additionally, we adapted the OWL-RL profile to OWL^C and we called it $OWL-RL^C$. The latter contains new contexts-dependent rules and novel rules for handling the new constructs.
- At this point, we must choose an adequate reasoning approach to validate our contextual model. We have preliminary results utilizing SPARQL inferencing notation (SPIN); however, we intend to apply other choices.
- We also propose to apply the entailments rules on different graphs implementing a variety of design patterns; by doing that, we aim to identify possible effects that design patterns have on the reasoning.
- Finally, we propose to test the complete model on different types of knowledge graphs among them Wikidata where qualifiers and references are attached to every statement.

6 Preliminary results

A contextual web ontology language OWL^C . It is based on a two-dimensional description logic [14] that includes a core and a context vocabulary. A contextual interpretation is a pair of interpretations: a core interpretation and a context interpretation. The core vocabulary defines contexts-dependent description of the concepts, roles and axioms of the ALCO fragment³.

- $[2000]$ *Student* is a contextualized concept which refers to students in the year “2000” where “2000” is the temporal context.
- $[wikipedia]$ *birthPlace* is a contextualized role which refers to the *birthPlace* property in the context of “wikipedia” where “wikipedia” is the provenance context.

² Because as mentioned by Klarman the cost is already hidden in the shift from one dimensional to two dimensional semantics

³ which is proven to be sound [14]

- $[before1970]$ $(CanVote \sqsubseteq Aged21orMore)$ is a contextualized axiom that illustrates the fact that, “before 1970”, voting was restricted to people who were at least 21. The same applies for the other axioms included in ALCO.
- $[2000]$ $Student(John)$ is a contextualized concept assertion which means that John was a student in the year “2000”.

We additionally use the rigid designator hypothesis [16] for individuals, which means that the interpretation of an individual is the same in all contexts.

The context language introduces two contextual constructors:

- $(\langle Country \rangle Citizen)$ illustrates the existential contextual operator. The example refers to the concept citizen in some context of type country.
- $([Country] Citizen)$ illustrates the universal contextual operator. The example refers to the concept citizen in all context of type country.

Contextual Reasoning with OWL-RL^C. We defined a profile for the contextual web ontology language that we presented previously, by adapting⁴ the idea of OWL 2 RL to OWL^C. We call this new profile OWL-RL^C. Due to space limitations, we introduce only one rule of each language in table 1. We use a quaternary predicate $Q(s; p; o; co)$ ⁵ where s is the subject, p is the predicate, o is the object and co is the context for which the predicate holds. Variables in the implications are preceded with a question mark. For instance, the contextual entailment rule of the universal constructor of the core language $(\forall role.Concept)$ takes three forms where: 1) both the corresponding class and role are contextual, 2) only the class is contextual, 3) only the role is contextual. Table 1 illustrates the first case. On the other hand, defining the rules of the context language is crucial because it imposes the declaration of new predicates such as: $owl-rl^c:onClass$ (i.e. declares the class on which the constructs apply), $owl-rl^c:inAllContextOf$ (i.e. similar to $owl:allvaluesFrom$ but for the contexts only) among others.

Practical implementation:

- Encoding contexts in RDF: we showed in [1] that the fluent model is capable of supporting semantic relations between multiple time intervals. As a first attempt, we extended this model to any dimension of contexts and adapt it to support the representation of the context dependent concepts (roles, and axioms too) of the core language. We implicitly used the standard mapping⁶

⁴ Syntactically, we are considering only a subset (fragment) of OWL 2 RL whose constructors and axioms correspond to ALCO, i.e. (approximately) the intersection of OWL 2 RL and ALCO. Or, equivalently, ALCO with the sub class axiom restrictions of OWL 2 RL. Additionally, the semantics is contextual.

⁵ In the original version, they use a predicate T which is a generalization of RDF triples

⁶ <https://www.w3.org/TR/owl2-mapping-to-rdf/>

Table 1. Two rules of *OWL – RL^c*

	IF	THEN
$\forall p.Y$	$T(?x, owl:allValuesFrom, ?y)$ $T(?x, owl:onProperty, ?p)$ $Q(?u, rdf:type, ?x, ?co)$ $Q(?u, ?p, ?v, ?co)$	$Q(?v, rdf:type, ?y, ?co)$
$([C]D)$	$T(?e, owl^c:onClass, ?d)$ $T(?e, owl^c:inAllContextOf, ?c)$ $T(?x, rdf:type, ?e)$ $Q(?x, rdf:type, ?d, ?y)$	$T(?y, rdf:type, ?c)$

of OWL to RDF to represent the concepts, roles and axioms of the core vocabulary and extended it to handle the contextual constructs of the contexts vocabulary.

- Implementation of OWL-RL^C using SPIN rules: the majority of the rules for the core vocabulary generate a quadruple Q. That means, there is a generation of new objects, some of which, are the context instances that could be handled neither by OWL reasoners, nor by an addition of SWRL rules. Therefore, we decided to use SPARQL spin notation. SPIN can be used to encapsulate reusable SPARQL queries as templates. Then, they can be instantiated in any RDF or OWL ontology to add inference rules and constraint checks. Using Sparql rules, we managed to generate the new objects while committing to some predefined constraints, for instance, the non-generation of existing contextual statements which is incorporated directly as a filter in the sparql rule.

7 Evaluation plan

We have designed a multi-dimensional design space for the evaluation of the overall model :

- *Expressiveness of the model in terms of description logics varieties.* This will involve a theoretical investigation of the expressiveness of the logic as well as a user study to check to what extent the proposed logic addresses specific application scenarios. In particular, we tend to evaluate the model on an ongoing project in digital humanities [2], where linguists are interested in revealing temporal aspects of Ferdinand de Saussure’s manuscripts, using the associated contextualized knowledge graph.
- *Usability of the contextual language.* To measure the usability of the language, we will check if domain specialists (e.g. historians, linguists, etc.) can use it to express contextual facts and axioms.
- *Degree of object proliferation.* Using a model that introduces many properties and objects can lead to undesirable graph size increases, which oftentimes cause detrimental memory performance. The worst case scenario could lead to an explosion of the number of triples. Therefore, we plan to measure the

number of objects and predicates using our context representation approach and compare it to other approaches.

- *Time needed to check the consistency of the model.*
- *Ability to deal with polymorphism when adding new dimensions of contexts.* In other words, the model should be flexible enough to make the knowledge base grow linearly when adding new dimensions of contexts.
- *Generation of non-desired inferences.* A previous work [10] showed that the adoption of certain patterns can generate undesired inferences. To avoid that, we will study the behavior of every rule separately to check if there is a generation of non-desired inferences.

8 Reflections

To conclude, we identified from the state of the art approaches that the task of contextual reasoning on the web of data is still in early stages. In our research, we intend to push this forward with a practical implementation. We introduced a contextual extension of the web ontology language OWL^C based on a two-dimensional description logic. Additionally, we created a profile for contextual reasoning by adapting the idea of OWL-RL to OWL^C . $OWL-RL^C$ contains new context-dependent rules and novel rules for handling the new constructs. We did not consider the semantic relations that could exist between the contexts, but we plan to work on this in the next phase. We would also like to study the requirements to extend the model to a fragment larger than ALCO. On a lower level, we consider that the problem of encoding contextual knowledge in RDF datasets is a minor issue because it is already performed locally by a lot of data providers. We believe that what should be settled is an upper vocabulary to be commonly used for describing such metadata.

Acknowledgments: I would like to thanks my PhD advisors Prof. Gilles Falquet and Prof. Didier Buchs.

References

1. Aljalbout, S., Falquet, G.: Un modele pour la representation des connaissances temporelles dans les documents historiques. arXiv preprint arXiv:1707.08000 (2017)
2. Aljalbout, S., Falquet, G.: A semantic model for historical manuscripts. arXiv preprint arXiv:1802.00295 (2018)
3. Benslimane, D., Arara, A., Falquet, G., Maamar, Z., Thiran, P., Gargouri, F.: Contextual ontologies. *Advances in Information Systems* pp. 168–176 (2006)
4. Berners-Lee, T., Hendler, J., Lassila, O., et al.: The semantic web. *Scientific american* 284(5), 28–37 (2001)
5. Borgida, A., Serafini, L.: Distributed description logics: Assimilating information from peer sources. *J. Data Semantics* 1, 153–184 (2003)
6. Bozzato, L., Homola, M., Serafini, L.: Context on the semantic web: Why and how. *ARCOE-12* p. 11 (2012)

7. Dividino, R., Sizov, S., Staab, S., Schueler, B.: Querying for provenance, trust, uncertainty and other meta knowledge in rdf. *Web Semantics: Science, Services and Agents on the World Wide Web* 7(3), 204–219 (2009)
8. Gangemi, A., Mika, P.: Understanding the semantic web through descriptions and situations. In: *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*. pp. 689–706. Springer (2003)
9. Ghidini, C., Giunchiglia, F.: Local models semantics, or contextual reasoning= locality+ compatibility. *Artificial intelligence* 127(2), 221–259 (2001)
10. Giménez-García, J.M., Zimmermann, A., Maret, P.: Ndfuents: An ontology for annotated statements with inference preservation. In: *European Semantic Web Conference*. pp. 638–654. Springer (2017)
11. Giunchiglia, F.: Contextual reasoning. *Epistemologia, special issue on I Linguaggi e le Macchine* 16, 345–364 (1993)
12. Hartig, O., Thompson, B.: Foundations of an alternative approach to reification in rdf. *arXiv preprint arXiv:1406.3399* (2014)
13. Hayes, P.: Formal unifying standards for the representation of spatiotemporal knowledge a report on arlada task 02ta4-sp1-rt1" formalisms for spatio-temporal reasoning" advanced decision architectures alliance (2004)
14. Klarman, S., Gutiérrez-Basulto, V.: Two-dimensional description logics for context-based semantic interoperability. In: *AAAI* (2011)
15. Kutz, O., Lutz, C., Wolter, F., Zakharyashev, M.: E-connections of abstract description systems. *Artificial intelligence* 156(1), 1–73 (2004)
16. LaPorte, J.: Rigid designators (2006)
17. McCarthy, J.: Generality in artificial intelligence. *Communications of the ACM* 30(12), 1030–1035 (1987)
18. Nguyen, V., Bodenreider, O., Sheth, A.: Don't like rdf reification?: making statements about statements using singleton property. In: *Proceedings of the 23rd international conference on World wide web*. pp. 759–770. ACM (2014)
19. Welty, C.: Context slices: representing contexts in owl. In: *Proceedings of the 2nd International Conference on Ontology Patterns-Volume 671*. pp. 59–60. CEUR-WS.org (2010)
20. Welty, C., Fikes, R., Makarios, S.: A reusable ontology for fluents in owl. In: *FOIS*. vol. 150, pp. 226–236 (2006)