Semantic Web Trends on Reasoning Over Sensor Data

Edmond Jajaga[1], Lule Ahmedi[2] and Lejla Abazi Bexheti[1]

[1] South East European University, Faculty of Contemporary Sciences and Technologies, Ilindenska n. 335, 1200 Tetova, Macedonia
{e.jajaga, l.abazi}@seeu.edu.mk
[2] University of Prishtina, Department of Computer Engineering, Nene Tereze nn, 10000 Prishtine, Kosova
lule.ahmedi@uni-pr.edu

Abstract. One of the most challenging features of Semantic Web applications is the reasoning module. Based on the Semantic Web paradigm, the tendency of its community is to build applications with well known standards or recommendations. OWL and SWRL are the first to be considered while expressing data semantics. Their support of monotonic reasoning and open world assumption is not always fruitful on the domain of sensor data. Sensor data, which are the case of this study, are specific in terms of their dynamic nature. Addressing the reasoning issues over the gigantic flow of sensor data the community has considered different approaches mostly resulting by building hybrid systems. This paper outlines preliminary work on reasoning issues over sensor data while describing the approaches that have been taken so far focusing on the discipline of water quality management.

Keywords: sensor data, Semantic Web, reasoning, rules, OWL, SWRL, Semantic Sensor Web

1 Introduction

Semantic Web applications are growing day to day. Meanwhile Semantic Web standards are also maturing. Sensor rapid development and deployment in different disciplines including weather forecasting, water quality management, civic planning for traffic management etc. requires efficient machine communication. Many organizations and institutions have taken initiatives to take advantage from the synthesis of both "worlds" to provide semantics on different application domains. In 2008, Kno.e.sis[1] initiated a project for building Semantic Sensor Web assembling sensor metadata from all over the world. The initiative is aligned-well with standardization efforts of W3C and Open Geospatial Consortium (OGC), in particular

[1] The Ohio Center of Excellence in Knowledge-enabled Computing at the Wright State University, http://knoesis.wright.edu/

with Semantic Web[2] and Semantic Web Enablement[3] (SWE) activities, respectively. In fact, Semantic Sensor Web represents a synergy of both initiatives by semantic annotating of simple sensor data i.e. time, spatial and thematic data. In line with Semantic Sensor Web the W3C Semantic Sensor Network Incubator group (the SSN-XG) recently produced an OWL 2 [11] ontology named SSN[4] [4], which enhances OGC SWE simple spatial and temporal concepts with semantic annotation for analyzing and Linked Data publishing. The SSN ontology models sensor data in four main perspectives: sensor, observation, system and feature and property perspectives.

Sensor data are an example of stream data which are rapidly changing data. These huge amounts of data need to be quickly consumed and reasoned over. For example, if a particular water quality parameter drops from its allowed threshold then this information needs to be captured quickly and an appropriate decision should follow. Sensors continually produce water quality parameter values. Historical and real-time data produced by sensors require a flexible knowledge management system. An area which deals with continues execution of queries over stream data is Data Stream Management Systems (DSMS). As indicated in [22] it lacks the ability to reason about complex tasks and lacks a protocol for wide publication. The Semantic Web fulfills these gaps but caching all the knowledge for rapidly changing information is inappropriate. Similar to DSMS is Complex Event Processing (CEP) which provides on-the-fly analysis of event streams, but cannot perform reasoning tasks [15]. Following the pros and cons of DSMS and CEP a new research area has been investigated by the community, namely Stream Reasoning [22]. Stream Reasoning integrates data streams, the Semantic Web and reasoning techniques into a unique platform. Unlike in a traditional reasoning environment, where all the information is taken into account, in stream reasoning there are two concepts which indicate the distinguished approach. The *window* concept restricts the reasoning to a certain subset of statements recently observed on the stream while previous information is ignored, furthermore *continuous processing* means continuous evaluation of streams against the knowledge base which is constantly changing.

In general, querying RDF triples of stream data has been leveraged with different SPARQL extensions like: Streaming SPARQL [24], Continuous SPARQL (C-SPARQL) [23] and Time-Annotated SPARQL [25].

This study is focused on the Semantic Web rule layer. State-of-the-art rule-based systems for dealing with sensor data reasoning are mainly:

- *Hybrid systems* e.g. CEP with Semantic Web in [14] and [15], production rules with Semantic Web in [13] and [27].
- *Pure Semantic Web rule systems* as given in [5], [6] and [21], but which do not deal with the streaming nature of sensor data.

This paper is organized as follows. Section 2 presents the current trends on ontology and query processing for sensor networks. Section 3 describes the current state-of-the-art of rule-based implementations with focus on sensor application areas.

---

[2] See http://www.w3.org/2001/sw/
[3] See http://www.opengeospatial.org/domain/swe
[4] Available at http://purl.oclc.org/NET/ssnx/ssn

The main discussion takes place on Section 4. Finally, the paper is concluded with a discussions section.

2 Ontologies and Queries

Ontologies are defined as formal specifications of a shared conceptualization [19]. Because of its knowledge reuse and sharing, the ontological knowledge model has been widely leveraged for representing wireless sensor networks (WSNs). One of the first WSNs which has seen benefits from including the ontological model into its knowledge base is OntoWEDDS [9], a decision-support system for wastewater management, which extends its previous version's case-based and rule-based reasoning with the WaWO [10] ontology. The evaluation results have yielded an improvement of 70-100% successful diagnosis and no impasse situations including WaWO reasoning, against 60-70% and 10 out of 57 impasse situations without using it.

Interoperability between sensors and sensing systems was enabled with the development of the SSN ontology. Its foundation is based on the DOLCE-UltraLight[5] (DUL) ontology. To model a knowledge base of sensor networks one would include SSN interested features extending it with units, location, feature and time ontologies [4]. Additional classes and properties can be defined and added to model domain specific knowledge.

There are also initiatives dealing with sensor streaming data on query level. In [26], Shahriar et al. propose a smart query system considering both streaming data and historical data from marine sensor networks. ES3N [18] and C-SPARQL [23] are also dealing with sensor stream data. C-SPARQL is an extension of SPARQL for supporting stream data querying. Query processing is an important issue on the Semantic Sensor Web [26], but it is out of the scope of this paper. Instead, we focus on rule layer reasoning.

3 Rule-Based Reasoning

As claimed in the previous section almost every sensor network knowledge base is modeled through OWL ontologies. The Semantic Sensor Web foundation has enabled semantic enrichment of simple sensor data through these ontologies. However, inferring new and implicit knowledge from known facts represented in ontological terms is enabled through a powerful mechanism known as rule-based reasoning. In general, the limited expressivity of the Sematic Web Rule Layer (SWRL) [26], which currently has the status of W3C submission, has forced the community to consider hybrid systems while keeping the knowledge base modeled in the form of ontologies. Specifically, for the domain of sensor data an obstacle appears from the continuous flow of data. These data need to be consumed quickly by the reasoning engine which

[5] http://www.loa-cnr.it/ontologies/DUL.owl

in turn will efficiently infer new knowledge by combining these data with background knowledge. Because of this nature when trying to infer logical consequences from sensor data, different rule systems are considered by the community. In general, rule systems fall into three categories: first-order, logic programming (LP), and action rules [29]. In the rest of this section we will briefly describe the main rule systems developed for modeling water quality management systems.

3.1 Association Rule Mining

Association rule mining is about finding frequent paths and correlations between items in the database. In [8], Ding et al. have proposed a framework for association rule mining and scoping in spatial datasets [8]. For example, they have used an association rule to infer dangerous arsenic levels with 100% confidence.

As envisioned by Bhatnagar and Kochhar [7], association rule mining performing on stream data are increasingly in need. They are employed in the estimation of missing data streams of data generated by sensors and frequency estimation of internet packet streams [7].

Association rule mining is more concerned with predicting what may happen in the future, while our aim is to deal with the current state of water quality.

3.2 Production Rules

Production rules are IF-THEN rules which fire actions (the THEN part) based on the precondition (the IF part) matching the current "state of the world". In [27], the authors model a hybrid Environmental Decision Support System (EDSS) for Waste-Water Treatment Plants (WWTP). As an example of production rules they infer invalid NO3 measurement values. They argue that the WWTP domain should be modeled through ontologies, for modeling sensor data, paired with decision-making rules, for processing incoming sensor data and recommending actions to be taken.

Another system implemented in terms of production rules has been designed by Chau [13] in the domain of water quality modeling. Namely, the system simulates human expertise during the problem solving of coastal hydraulic and transport processes. Both forward-chaining and backward-chaining are used collectively during the inference process [13].

The W3C in 2005 has created the Rule Interchange Format (RIF)[6], a standard for rule exchange, with one of its dialects RIF-PRD[7] implementing production rules.

As proved in [27] production rules can be efficiently employed on a hybrid model for dealing with sensor networks. Production rules can modify and retract knowledge base facts. RIF-PRD enables interoperability of different rule systems with Semantic Web standards [29]. However, the sensor networks community is committed to Semantic Web approach, which standards (e.g. RDF, OWL) embrace monotonic inference only.

[6] http://www.w3.org/TR/rif-overview/
[7] http://www.w3.org/TR/2013/REC-rif-prd-20130205/

3.3 Event Processing

Another hybrid approach which deals with sensor data reasoning is using OWL ontologies and CEP, which is an area similar to Stream Reasoning. CEP provides on-the-fly real-time processing of streams of atomic events (e.g. sensor readings) [14], [15]. Taylor and Leidinger [16] translate the whole OWL ontology, which models the event definition and optimization and extends an early version of the SSN ontology, into CEP statements for processing in an event processing engine. Unlike this approach, Anicic et al. [15] have taken the advantage of both "worlds" synthesizing the ability of CEP systems to process real-time complex events within multiple streams of atomic occurrences and the Semantic Web i.e. ontological ability to effectively handle background knowledge and perform reasoning. The later approach has resulted in a new rule-based language ETALIS [14] and EP-SPARQL [15], a query language extending the SPARQL language with event processing and stream reasoning capabilities. Both are implemented in Prolog, which has its foundations in LP.

All in all, CEP systems are very useful while dealing with real-time detection of sequential events e.g. if X water quality parameter drops from its allowed threshold, then another Y water quality parameter should be measured and if it is lower (or greater) than some constant *a,* than the river will be classified as 'polluted' and a water expert should be alerted. However, CEP cannot be used as a standalone system for reasoning in sensor networks. In fact, they lack the reasoning module, which is complemented by Semantic Web as it is the case with ETALIS [14].

3.4 Semantic Web

Dealing with sensor data, a pure Semantic Web approach has been implemented by Keßler et al. [5]. They have leveraged the SWRL's ability to express free variables and the use of its built-ins for modeling mathematical functions which has fulfilled the OWL's lack of mathematical processing capabilities. The approach is tested for geographical information retrieval (GIR) task for recommending personalized surf spots based on user location and preferences. A similar approach is taken by Wei and Barnaghi [6] who demonstrate how rule-based reasoning can be performed over sensor observation and measurement data within the terms of Semantic Sensor Web. They emphasize the ability of rules not just to infer accurate but also approximate knowledge.

Approaches [5] and [6] have proved that expressing SWRL rules in the domain of sensor data is feasible. However, they lack implementation and do not deal with reasoning obstacles i.e. monotonicity and open world assumption and their implications. These will be addressed in more detail in the next section.

Henson et al. [21] have used the Jena Semantic Web Framework [12] as an engine for reasoning with rules implemented for Semantic Sensor Web on weather domain. Using Jena rules they infer new knowledge about sensor observation data and link the newly generated relations with original observation time and location data. This approach suffers from the monotonicity issue, since Jena supports monotonic inference by default.

4 Discussion

Semantic Sensor Web enables semantic sharing and reasoning over sensor data spread all over the world. Inferring new knowledge from these data represents a huge improvement. However, this achievement cannot be easily obtained, because of the streaming nature of sensor data. During our previous works in [1] and [2], attempting to implement a database normalization tool through Semantic Web technologies we have encountered different issues on the ontological and rule level. OWL and SWRL's open world assumption and monotonic reasoning implied difficulties while dealing with negation as failure, classical negation, knowledge base modifications and disjunction. We think we will encounter the same reasoning challenges as sensor networks community is committed to Semantic Web, and Description Logic (DL), on which are based Semantic Web formal foundations, is the one to blame for this. Because of this inadequacy of DL, the preferred implementation has often been to use LP instead. As an example, ETALIS [14] is a rule-based system which reasons over sensor data and it is implemented in Prolog, which is a LP language.

4.1 Streaming Nature of Sensor Data

Wei and Barnaghi [6] and Keßler et al. [5] represent how SWRL reasoning can be effectively leveraged in Semantic Sensor Web. Henson et al. [21] use Jena rules [12] to infer new knowledge from sensor observation data. All these approaches do not consider the streaming nature of sensor data and thus the continual modification of sensor outputs. As envisioned by Barbieri et al. [28], because Semantic Web is still focusing on static data the continuous processing of data streams together with rich background knowledge, requires specialized reasoners. As a result, they have developed C-SPARQL to enable querying over stream data in combination with static background knowledge. We envision a need for a rule system which will deal with rapid flow of sensor data in combination with static knowledge. Perhaps a similar approach to Barbieri et al. [23] should be considered by SWRL developers to enable continues execution of rules over sensor data.

4.2 Open World Assumption and Negation As Failure

The open world assumption deals with the completeness of the knowledge about a particular domain. In contrast to closed world assumption where the absence of a fact will return false, in open world semantics what is not known is undefined. For example, one cannot be sure that a particular measurement site is not polluted just because the water quality parameters are within their allowed thresholds. If a property stating that a measurement site is not polluted is absent then it does not mean that this property does not exist. This uncertainty makes it difficult to consider pure Semantic Web approaches. As a consequence of open world assumption we cannot assume the negation as failure. Negation as failure means that we can infer not P if we fail to prove every possible proof of P. For example, water temperature should not reach the rates greater than 1.5°C plus ambient temperature for general conditions. Expressing this rule is not possible in DL, because of not supporting the negation as failure.

LP and production rules approaches have the advantages of supporting closed world assumption with negation as failure. In this manner their knowledge is defined. This is why the sensor networks community has implemented a hybrid Semantic Web approach, layering closed world reasoning on top of open world assumption i.e. OWL. In general, CEP is attractive, but we do not have any basis for confidence in how it works within a system making mixed assumptions about open and closed worlds. A possible approach is using query languages adopting closed world assumption, which are out of the scope of this discussion. Instead, we are focused on rule-based reasoning.

4.3 Monotonicity

OWL and SWRL support monotonic inference only i.e. earlier conclusions cannot be invalidated. Dealing with sensor data, we agree with Keßler et al. [5], who state that the same reasoning steps may lead to different results, thus the problem of monotonicity need to be addressed.

Let's consider the following example:

A sensor node, which is a collection of sensing devices (e.g. sensor probes) placed in an observation point of the river, will report values about water quality parameters which are within the allowed threshold. So, the sensor node will be classified as "clean". Suddenly, a parameter drops from its allowed threshold thus the sensor node should be classified as "polluted".

Because of the monotonic inference, the sensor node firstly asserted as "clean" cannot be later modified as "polluted". Additionally, as indicated in [22] the windows opened over streams can determine changes in the static information sources. This inadequacy has forced the sensor networks community to enable nonmonotonic reasoning through hybrid approaches described in the previous section. They have layered LP e.g. Prolog system, or production rule systems reasoning with updated working memory, to enable nonmonotonic inference.

5 Conclusion

In this paper we described current trends on representing sensor data with semantic technologies and their pros and cons. Because of the streaming nature of sensor data the community has mainly taken two approaches: building a hybrid system or a pure Semantic Web system. Both approaches prove that the Semantic Web rule layer lacks the capabilities needed to efficiently reason over sensor data. This is due to the issues of open world semantics and monotonicity. With some examples from the domain of water quality management we have demonstrated the inability of a pure Semantic Web approach. Because of this shortage the sensor networks community has integrated ontologies with closed world and nonmonotonic reasoning.

Acknowledgements. The work described in this paper was supported by "InWaterSense: Intelligent Wireless Sensor Networks for Monitoring Surface Water Quality", an EU funded project managed by European Union Office in Kosovo, implemented by University of Prishtina.

References

1. Ahmedi, L., Jajaga, E.: Normalization of relations and ontologies, In: The 10th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Databases, pp. 419-425, Cambrige, UK (2011)
2. Ahmedi, L., Jajaga, E.: A database normalization tool using Semantic Web technologies, In: International Journal of Systems Applications, Engineering and Development, vol. 5 (2011)
3. Sheth, A., Henson, C., Sahoo, S.: Semantic Sensor Web, In: IEEE Internet Computing, pp. 78-83 (2008)
4. Compton, M., Barnaghi, P., Bermudez, L., Garcıa-Castro, R., Corcho, O., Cox, S., Graybeal, J., M. Hauswirth, C. Henson, A. Herzog, V. Huang, K. Janowicz, W. D. Kelsey, D. Le Phuoc, Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheth, A., Taylor, K.: The SSN Ontology of the W3C Semantic Sensor Network, Incubator Group, Journal of Web Semantics (2012)
5. Keßler, C., Raubal, M., Wosniok, C.: Semantic Rules for Context-Aware Geographical Information Retrieval, In: Barnaghi, P. (eds.) European Conference on Smart Sensing and Context, (EuroSSC 2009), LNCS, vol. 5741, pp. 77–92, Springer (2009)
6. Wei, W., Barnaghi, P.: Semantic Annotation and Reasoning for Sensor Data, In: Smart Sensing and Context, pp.66-76 (2009)
7. Bhatnagar, V., Kochhar, S.: Association Rule Mining, Encyclopedia of Artificial Intelligence, pp. 172-178, IGI Global (2009)
8. Ding, W., Eick, C. F., Yuan, X., Wang, J., Nicot, J.: A Framework for Regional Association Rule Mining and Scoping in Spatial Datasets, Geoinformatica, vol. 15, issue 1, pp. 1-28 (2011)
9. Ceccaroni, L., Cortes, U., Sanchez-Marre, M.: OntoWEDSS: an ontology-underpinned decision-support system for wastewater management (2001)
10. Ceccaroni, L., Cortes, U., Sanchez-Marre, M.: WaWO-An ontology embedded into an environmental decision-support system for wastewater treatment plant management, In: Proceedings of ECAI2000 - Wo9: Applications of ontologies and problem-solving methods, pp. 2.1-2.9, Berlin, Germany (2000)
11. Hitzler, P., Krotzsch, M., Parsia, B., Patel-Schneider, P., Rudolph, S., OWL 2 Web Ontology Language Primer, http://www.w3.org/TR/owl2-primer/ (2009)
12. Jena Semantic Web Framework, http://jena.sourceforge.net/, last accessed July, 2nd 2013
13. Chau, K. W.: An ontology-based knowledge management system for flow and water quality modeling, In: Advances in Engineering Software, vol. 38, no. 3, pp. 172-181 (2007)
14. Anicic, D., Fodor, P., Rudolph, S., Stuhmer, R., Stojanovic, N., Studer, R.: A Rule-Based Language for Complex Event Processing Reasoning, In: Proceedings of the Fourth

International Conference on Web reasoning and rule systems, pp. 42-57, Springer-Verlag Berlin, Heidelberg (2010)
15. Anicic, D., Fodor, P., Rudolph, S., Stojanovic, N.: EP-SPARQL: A Unified Language for Event Processing and Stream Reasoning, In: WWW 2011, pp. 635-644 (2011)
16. Taylor, K., Leidinger, L.: Ontology-Driven Complex Event Processing, In: ESWC, LNCS, pp. 285-299, Springer, Greece (2011)
17. Amato, G., Caruso, A., Chessa, S., Massi, V., Urpi, A.: State of the Art and Future Directions in Wireless Sensor Network's Data Management, In: Istituto di Scienza e Tecnologie dell'Informazione del CNR, (2004)
18. Lewis, M., Cameron, D., Xie, S., Arpinar, I. B.: ES3N: A Semantic Approach to Data Management in Sensor Networks, In: A workshop of the 5th International Semantic Web Conference ISWC (2006)
19. Gruber, T. R.: A translation approach to portable ontologies, In: Knowledge Acquisition, vol. 5(2), pp. 199-220, (1993)
20. Barbieri, D., Braga, D., Ceri, S., Della Valle, E., Huang, Y., Tresp, V., Rettinger, A., Wermser, H.: Deductive and Inductive Stream Reasoning for Semantic Social Media Analytics. In: Proceedings of the Seventh Extended Semantic Web Conference (ESWC'10), L. Aroyo, G. Antoniou, E. Hyvonen, A. Teije, H. Stuckenschmidt, L. Cabral, and T. Tudorache, Eds. LNCS, vol. 6088, pp. 1–15. Springer-Verlag (2010)
21. Henson, C. A., Pschorr, J. K., Sheth, A. P., Thirunarayan, K.: SemSOS: Semantic Sensor Observation Service, In: Proceedings of the 2009 International Symposium on Collaborative Technologies and Systems (CTS 2009), Baltimore, MD (2009)
22. Della Valle, E., Ceri, S., van Harmelen, F., Fensel, D.: It's a Streaming World! Reasoning upon Rapidly Changing Information, In: IEEE Intelligent Systems, vol. 24(6), pp. 83–89 (2009)
23. Barbieri, D. F., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: C-SPARQL: SPARQL for Continuous Querying. In: Proceedings of the Eighteenth International Conference on World Wide Web (WWW'09), Quemada, J., Leon, G., Maarek, Y., Nejdl, W. (eds.), pp. 1061–1062, ACM Press (2009)
24. Bolles, A., Grawunder, M., Jacobi, J.: Streaming SPARQL - extending SPARQL to process data streams. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 448–462. Springer, Heidelberg (2008)
25. Rodriguez, A., McGrath, R. E., Liu, Y., Myers, J.: Semantic Management of Streaming Data. In: Proceedings of International Workshop on Semantic Sensor Networks, SSN, pp.80-85 (2009)
26. Shahriar, M. S., De Souza, P., Timms, G.: Smart Query Answering for Marine Sensor Data, Sensors 2011, vol. 11(3), pp. 2885-2897 (2011)
26. Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M.: SWRL: A Semantic Web Rule Language Combining OWL and RuleM, W3C Member Submission, http://www.w3.org/Submission/SWRL/, last accessed April, 22nd 2013 (2004)
27. Sottara, D., Bragaglia, S., Mello, P., Pulcini, D., Luccarini, L., Giunchi, D.: Ontologies, Rules, Workflow and Predictive Models: Knowledge Assets for an EDSS, In: Seppelt, R.,. Voinov, A. A., Lange, S., Bankamp, D. (eds.), International Environmental Modelling and Software Society (iEMSs), 2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany (2012)
28. Barbieri, D., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: Stream Reasoning: Where We Got So Far. In: Proceedings of the 4th International Workshop on New Forms of Reasoning for the Semantic Web: Scalable and Dynamic (NeFoRS) (2010)

29. Kifer, M., Boley, H.: RIF Overview (Second Edition), W3C Working Group Note 5 February 2013, http://www.w3.org/TR/rif-overview/, last accessed July, 2nd 2013
30. Schaaf, M., Grivas, S. G., Ackermann, D., Diekmann, A., Koschel, A., Astrova, I.:Semantic Complex Event Processing, In: Proceedings of the 5th WSEAS congress on Applied Computing conference, and Proceedings of the 1st international conference on Biologically Inspired Computation, pp.38-43 (2012)