StreamJess: Enabling Jess for Stream Data Reasoning and the Water Domain Case

Edmond Jajaga[1], Lule Ahmedi[2], Figene Ahmedi[3]

[1] South East European University, Department of Computer Science
Ilindenska n. 335, 1200 Tetovë, Macedonia
e.jajaga@seeu.edu.mk
[2] University of Prishtina, Department of Computer Engineering
Kodra e diellit pn, 10000 Prishtinë, Kosova
lule.ahmedi@uni-pr.edu
[3] University of Prishtina, Department of Hydro-Technic
Kodra e diellit pn, 10000 Prishtinë, Kosova
figene.ahmedi@uni-pr.edu

Abstract. This paper introduces *StreamJess*, a Stream Reasoning system that layers on top of a state of the art query processing system such as C-SPARQL to enable closed-world, non-monotonic and time-aware reasoning with Jess rules. The system is validated in the water quality monitoring domain by demonstrating water bodies' classification and pollution sources investigation.

Keywords: stream data, expert system, Semantic Web, rules, water quality monitoring, ontologies

1 Introduction

Even though the Semantic Web technologies have been extensively used for modelling stream data domains, e.g. SSN ontology[1], and for processing through SPARQL-like extensions, e.g. C-SPARQL [1], EP-SPARQL [2], etc., the recommended rules standards, SWRL and RIF, still remain not applicable in the domain of stream data applications. As a result, the stream data knowledge bases have been merely coupled with production rules, answer set programming or event processing systems [3]. In the vision of building a unique Semantic Web platform for reasoning over stream data, we have developed a production rules system, *StreamJess*, which layers Jess rules to reason over our water quality monitoring (WQM) ontology named InWater-Sense [4]. Jess supports closed-world and non-monotonic reasoning. However, extending Jess with stream data reasoning features is very expensive. Code could not be optimized even for simple temporal operations over event-streams [5]. Thus, we propose a much simpler approach by coupling stream data processing features, supported by state of the art Stream Reasoning (SR) query systems such as C-SPARQL, with

[1] Semantic Sensor Network Ontology, http://purl.oclc.org/NET/ssnx/ssn

adfa, p. 1, 2016.

Jess's reasoning abilities. C-SPARQL supports time-aware reasoning on stream data. However, as a query language, it is not intended to have any effect on the underlying ontology. In *StreamJess* we use Jess rules for populating the knowledge base. Moreover, they enable data modifications i.e. non-monotonic reasoning and the tools for archiving data. The system is validated with simulated data in the WQM domain, but it is developed for use within the InWaterSense project[2] with real data. The simulator randomly generates observation data for an arbitrary number of 70 measurement sites and 11 water quality parameters. A single sensor observation is arbitrarily set to be produced every second and includes 6 RDF streams representing time, location, device and quality of observation information. For example, in a 20 seconds window size 120 tuples will be produced. Moreover, the system supports registering multiple streamers to run concurrently.

2 Conceptual design and implementation

As depicted in Figure 1, *StreamJess* acts as a pipeline. Incoming RDF data streams, e.g. sensor observed values, are firstly filtered out and eventually aggregated by C-SPARQL queries. The query results are asserted into the knowledge base through JessTab functions. The ABox changes will eventually cause to fire Jess rules, which have been registered on application startup. The Jess engine inferences will be again published onto the ontology. The processing and reasoning over incoming streams is iterative for each window.

StreamJess is implemented as a Java console application. The application uses an instance of `jess.Rete` which is created at system start up. It provides the central access point of the application as it loads the ontology, builds the working memory, holds the list of rules and offers the methods for doing CRUD operations over facts [6]. Namely, Protégé functions of Jess in JessTab were used to manage with the knowledge base. All ontology modules are imported and loaded at application start up. Moreover, class instances are also mapped into the Jess's working memory. Different stream data ontologies can be loaded into the system and appropriate C-SPARQL queries and Jess rules can be defined to run over incoming data streams.

3 Validation

As a proof of concept, we have implemented *StreamJess* in a typical WQM scenario based on WSN. Sensors in InWaterSense WQM system are deployed in different measurement sites at different times. *StreamJess* will (1) classify the water body into the appropriate status (good, high or moderate) according to WFD regulations [7, 8] and (2) identify the potential sources of pollution in case of critical status detection. In general, each water quality is monitored and investigated with a *monitoring* rule (1) and an *investigation* one (2).

[2] http://inwatersense.uni-pr.edu/

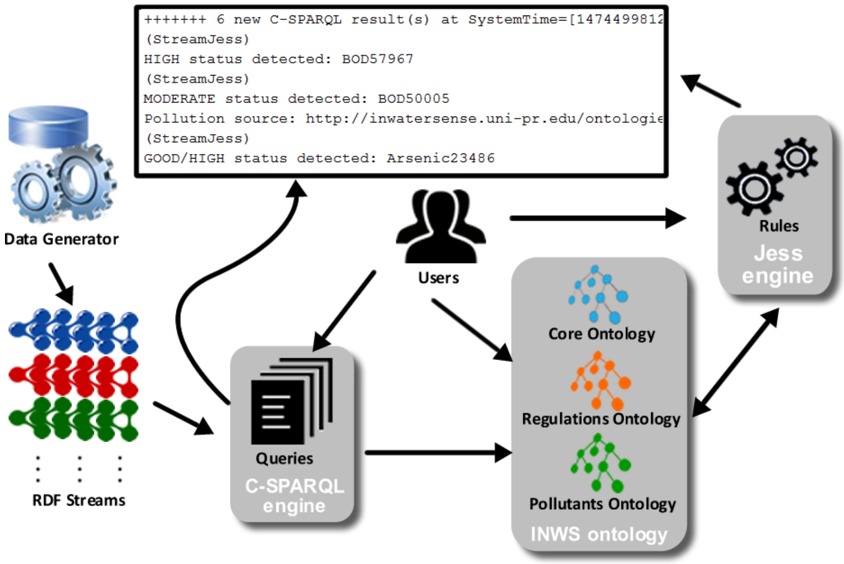

Fig. 1. *StreamJess* architecture

For brevity, we will demonstrate the case of Biochemical Oxygen Demand (BOD_5) and pH observations. Like most of water quality parameter observations, BOD_5 observations are classified based on the average value of measurements within a time interval while pH ones are considered one by one [8]. Two C-SPARQL queries are deployed into the system to match each of the types of observations. Moreover, four rules are deployed, one for monitoring and one for investigation of BOD_5 observations and another couple for pH observations. After loading all start up components, the user is asked to specify the window type of the queries. Namely, if he specifies time-based windows then he is presented with another question for setting the window size in seconds. Otherwise, he may specify to use tuple-based windows by further providing the number of tuples to be processed within a window. Each query eventually outputs triples of values: the water quality name, the location of measurements and the observed value i.e. the average value of BOD_5 measurements or individual pH measurements. Every output triple is mapped into a temporary observation class. Furthermore, for each new incoming triple a new call to the Rete method `run()` is invoked for doing rule-based reasoning. As illustrated in Figure 1, the Jess engine runs the rules against the temporary observation facts, produced by C-SPARQL, and it eventually activates the rule's RHS actions. The inferred knowledge forms another set of RDF data which is stored back into the ontology for further reasoning. Namely, monitoring rules do the water quality classifications based on the WFD regulation rules which general form looks like follows: `{observation details} => {classify and archive the observation}`. In case a critical status is detected, investigation rules act to identify the pollution source which general form is: `{moderate status observation} => {get and display the sources of pollution present on the measurement site}`.

```
++++ 3 new C-SPARQL result(s) at SystemTime=[1462808878402] ++++
#1 (C-SPARQL) Observed WQ: BiochemicalOxygenDemand Value: 1.6622
(StreamJess)(42693) BOD status is MODERATE
On: Mon May 09 17:47:58 CEST 2016
In: http://inwatersense.uni-pr.edu/ontologies/inws-core.owl#ms11
BOD5 pollution source: Urban stormwater discharges
BOD5 pollution source: Fish farming
#2 (C-SPARQL) Observed WQ: BiochemicalOxygenDemand Value: 1.9215
(StreamJess)(42181) BOD status is MODERATE
On: Mon May 09 17:47:58 CEST 2016
In: http://inwatersense.uni-pr.edu/ontologies/inws-core.owl#ms10
BOD5 pollution source: Urban stormwater discharges
#3 (C-SPARQL) Observed WQ: BiochemicalOxygenDemand Value: 0.6389
(StreamJess)(50374) BOD status is HIGH
On: Mon May 09 17:47:58 CEST 2016
In: http://inwatersense.uni-pr.edu/ontologies/inws-core.owl#ms12
```

Fig. 2. An excerpt of the output of the running example

An output of the running example is illustrated in Figure 2, where C-SPARQL processing of RDF streams has resulted with 3 new observations on 3 measurement sites: `ms10`, `ms11` and `ms12`. Two observations have been classified as of “moderate” status (line #1 and #2) and one of “high” status (line #3). Potential sources of pollution include urban stromwater discharges and fish farming on site `ms11` while urban stormwater is potential source of BOD_5 discharges on site `ms10`. An online demo of *StreamJess* can be found on the following link http://inwatersense.uni-pr.edu/streamjess/demo.html.

4 Conclusion

Until recently most of the SR research has been dedicated on ontology and stream processing developments. Our work goes beyond the query processing achievements and thus focusing on rule level implications of stream data reasoning. SWRL lacks the required expressivity level to reason over stream data. Thus, we built *StreamJess*, a production rule system supporting time-aware and non-monotonic reasoning.

5 References

1. Barbieri, D. F., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: C-SPARQL: a continuous query language for RDF data streams, J. of Semantic Computing, 4(1) 3–25 (2010)
2. 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)
3. Jajaga, E., Ahmedi, L., Abazi-Bexheti, L.: Semantic Web trends on reasoning over sensor data. In: 8th South East European Doctoral Student Conf., pp. 284-293. Greece (2013)
4. Ahmedi, L., Jajaga, E., Ahmedi, F.: An Ontology Framework for Water Quality Management. In Corcho, Ó., Henson, C. A., Barnaghi, P. M. (eds.) SSN@ISWC. pp. 35-50. Sydney (2013)
5. Ermert, L.: Comparing Jess and Esper for Event Stream Processing. Bachelor Thesis, Faculty IV - department of computer science, Fachhochschule Hannover, Germany (2009)
6. Hill, E. F. Jess in action: Java rule-based systems, Manning Publications Co., CT (2003)
7. Environment Agency. (2011) Method statement for the classification of surface water bodies, v2.0 (external release) [online], Monitoring Strategy v2.0, July 2011.
8. Statutory Instruments. European Communities Environmental Objectives (Surface Waters) Regulations 2015, S.I. No. 386 of 2015. http://www.irishstatutebook.ie/eli/2015/si/386/made/en/pdf (2009)