Object (B)logging: Semantic Self-Description for Cyber-Physical Systems

Giovanna Capurso, Michele Ruta, Floriano Scioscia, Eugenio Di Sciascio

Polytechnic University of Bari, Via E. Orabona, 4, I-70125 Bari, Italy
name.surname@poliba.it

Abstract. This paper proposes the *object (b)logging* framework, a novel general approach for the Semantic Web of Things. It allows associating semantic annotations to real-world objects and events as well as triggering complex objects interactions through advanced resource discovery. Machine learning algorithms are combined with non-standard reasoning services in order to produce a rich and meaningful semantic representation of events starting from a low-level statistical analysis of data gathered form sensing devices. The acquired knowledge is exposed to the outside world like in a blog and progressively enriched during the object's lifetime. The presented paradigm ideally applies to cyber-physical systems, where several mobile heterogeneous micro-devices cooperate to connote and modify appropriately the environment they are dipped in.

1 Introduction

The *Semantic Web of Things* (SWoT) vision [5] joins together the Semantic Web and the Internet of Things (IoT). Its goal is to embed intelligence in everyday objects and environments via semantic-enhanced pervasive computing. Heterogeneous and voluminous data streams are retrieved continuously and processed locally by mobile ad-hoc networks of smart objects dipped in a given environment, in order to detect events of interest in observed areas and achieve objects cooperation. This gives rise to distributed cyber-physical systems, able to make decisions and interventions on the environment according to the detected context and events in a fully automated fashion.

A *smart object* [1] is an intelligent agent running on a device equipped with embedded sensors, actuators, communication ports as well as (usually constrained) computation, storage and energy resources. For effective interaction, each smart object should describe itself and the context where it operates toward a variety of external devices and IoT applications. The existing approaches for knowledge extraction in smart objects use heuristic data mining methods which appear inadequate in complex scenarios. Moreover, state-of-the-art machine learning techniques, though constantly improving, require processing and storage capabilities that are ill-suited to resource-constrained pervasive scenarios and to the need of real-time responses. In order to improve flexibility and interoperability, Semantic Web standard technologies can be adopted for rich, unambiguous information exchange and manipulation.

The proposed *object (b)logging* approach aims to leverage the integration of standard supervised machine learning techniques with non-standard semantic-based inference services [6] to enable smart objects to produce and share a high-level description about themselves and the environment they are located in, as in a decentralized micro-log. According to this vision smart things are able to continuously enrich their basic descriptive core following events and phenomena they detect and expose a description toward the rest of the world in a self-contained fashion like in a blog.

In order to prove the effectiveness of such a framework, an early experimental testbed has been developed using an *Arduino Uno*[1] board. Preliminary tests and evaluations have been carried out and reported here, allowing to sketch refinement perspectives and future work.

2 Essential Background

In order to enable smart objects to share machine-understandable, unambiguous domain knowledge and to perform automated reasoning, Semantic Web languages and technologies grounded on Description Logics (DLs) [2] have been adopted. Specifically, current framework implementation relies on the *Attributive Language with unqualified Number restrictions* ($\mathcal{ALN}$) DL. It provides adequate expressiveness, while granting polynomial complexity to both standard and non-standard inference services. The reader is assumed to be familiar with DLs and standard reasoning basics.

The proposed approach leverages *semantic matchmaking* [6] for context detection. Given a request R and a resource S as concept expressions satisfiable w.r.t. a common ontology $\mathcal{T}$, the following non-standard inferences have been adopted to support approximate matches [6]:

- *Concept Contraction*: if R and S are not compatible, it determines which part of R is conflicting with S. By retracting conflicting requirements G (for *Give up*) from R, a concept K (for *Keep*) is obtained, representing a contracted version of the original request, such that $K \sqcap S$ is satisfiable in $\mathcal{T}$.
- *Concept Abduction*: when R and S are compatible but S does not subsume R, it determines a concept H (for *Hypothesis*) representing what should be hypothesized in S in order to completely satisfy R.

If R and S are incompatible (partial/disjoint match), one can use Concept Contraction to extract the compatible part K and move to a potential/intersection match. Then Concept Abduction can be used to obtain the required H_K for reaching a full match. Furthermore, *penalty functions* can be computed based on on the structure and number of concepts in G and H_K [6], defining a well-founded *semantic distance* measure, which can be used to rank a set of resources by relevance (*i.e.*, semantic affinity) w.r.t. the request.

[1] https://www.arduino.cc/en/main/arduinoBoardUno

3 Put Semantics in Cyber-Physical Systems: the Object (B)logging approach

Object (b)logging is based on the design, development and optimization of knowledge representation techniques to analyze and interpret continuous streams of heterogeneous data in order to build and exchange a structured and formal representation of an environment, process, subject or any other entity of interest in a cyber-physical system. Each object equipped with an embedded reasoning micro-engine can perform automated inference procedures to derive previously implicit knowledge out of information gathered from the environment or other smart entities. The overall goal is to bridge the semantic gap between low-level observations and high-level detected phenomena, enabling the development of autonomous smart things, each *logging* high-level context-aware descriptions. This is not allowed by conventional paradigms, such as Complex Event Processing (CEP). The major advantages the envisioned framework offers compared to CEP systems are a fully distributed architecture and more flexible reasoning.

The key aspect of object (b)logging is sharing the information produced and learned by all entities involved in the observed area as in a *blog*. The proposed approach envisions the blog as the collection of all the information learned and produced by every involved actor, through which it is possible to achieve a goal. More formally, a blog is composed by the semantic descriptions referred to a domain ontology progressively enriched during the objects' lifetime. It is based on a logical descriptive core and is used for intelligent interpretation of retrieved information.

Each object dipped in a given environment collects data from sensors and processes them in order to produce a high-level annotation of detected events and conditions. By evaluating this descriptive information, implicit knowledge is derived on-the-fly for identifying the task set needed to eventually change the environment state. Then each object automatically infers what useful capabilities it can provide and acts accordingly, in a decentralized and collaborative fashion.

Mobile and pervasive scenarios are featured by severe resource limitations influencing processing, memory, storage and energy consumption. Hence, systems and applications should take into account hardware and software constraints on pervasive object capabilities. The object (b)logging vision enables activity monitoring and recognition without requiring large computational resources. Machine learning algorithms combined with non-standard semantic-based reasoning allow managing approximate matches in order to compensate for possible anomalies in data gathering and communication, so increasing robustness and flexibility in sensing, interpreting and interacting in cyber-physical systems.

The (b)logging activity is a continuous process, composed of the tasks depicted in Figure 1:

- Characterize the context of each smart object in a fully automatic way, starting from a descriptive ontological core and sensed environmental data, generating semantically rich and compact context descriptions. This is achieved

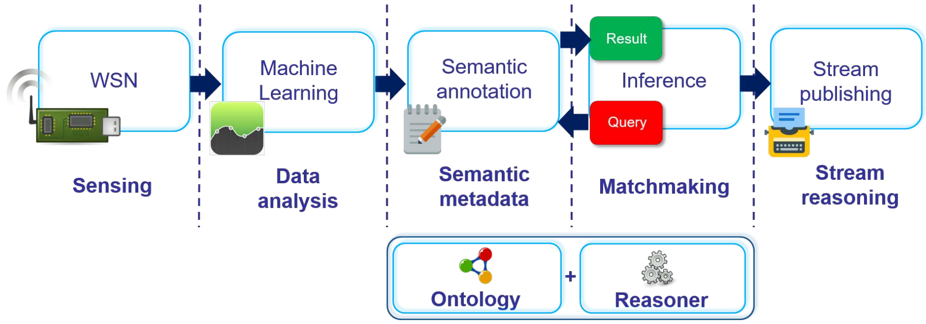

Fig. 1. Block diagram of the proposed approach for object (b)logging

by exploiting the integration of standard supervised machine learning techniques with non-standard inferences described in Section 2.
– Manage annotated data as a continuous stream of information by means of stream reasoning techniques [3, 4], providing the means to harness the flow of semantically annotated updates progressively improved during the object's lifetime.
– Achieve objects cooperation through the blog for triggering actions, taking decisions or making interventions on the environment by leveraging semantic matchmaking between available capabilities and required goals and tasks.

The proposed vision results as a general-purpose, cross-domain semantic-based facilitator for analyzing raw environmental data gathered via heterogeneous sensing devices in an environment, interpreting information and associating high-level characterization suitable for planning proper feedback in cyber-physical systems.

4 Preliminary Results and Conclusion

An early Java-based prototype interacting with Arduino Uno board has been developed and tested, basically devoted to assess feasibility of the proposed ideas. The semantic matchmaking on $\mathcal{ALN}$ DL (b)log entities grounding knowledge extraction is in charge to *Mini-ME* reasoning engine [6], which is suitable for computationally constrained nodes. Preliminary performance evaluation was focused on fire risk monitoring in confined environments. Data were collected through Arduino board equipped with additional sensors in 10 s observation windows and 50% window overlap degree. Each experiment involved tasks described in Section 3 carried out in a continuous process through 200 iterations. Time was measured embedding timestamp instructions in the source code. Figure 2 reports on turnaround time results per iteration. As expected, there are peaks in the first iterations, due to preliminary access to configuration files and data structure initialization. Subsequently, processing required less than 100 ms for each 10 s data window.

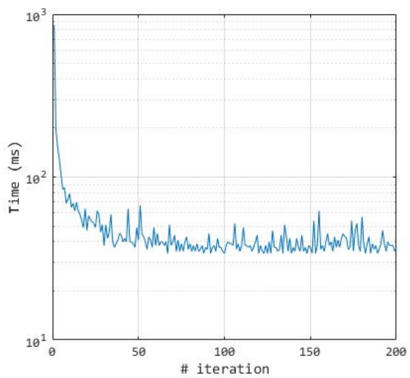

Fig. 2. Turnaround time test

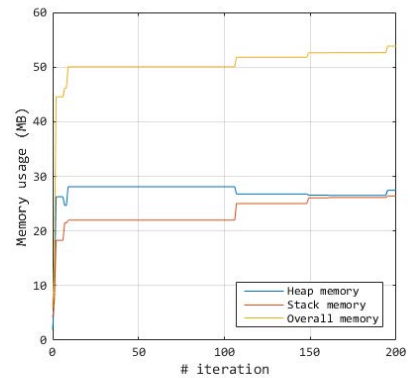

Fig. 3. Memory usage test

For memory usage analysis, methods of the *MemoryMXBean* Java Platform class were used: results are shown in Figure 3. Experimental results demonstrate that time performance is well suited to real-time monitoring needed to detect events and trigger actions on-the-fly, while memory load needs further optimization to fit properly the target application scenarios. Moreover, despite the hardware limitations, current events and conditions were accurately detected and their high-level machine-understandable annotations could be shared.

Future work directions concern further performance optimization, deployment and comparison with state-of-the-art approaches. Several future perspectives are open for an extension of the testbed and framework integration into software platforms for robots and drones (ROS – Robot Operating System, NuttX).

References

1. Atzori, L., Iera, A., Morabito, G.: From "smart objects" to "social objects": The next evolutionary step of the internet of things. IEEE Communications Magazine 52(1), 97–105 (2014)
2. Baader, F., Calvanese, D., Mc Guinness, D., Nardi, D., Patel-Schneider, P.: The Description Logic Handbook. Cambridge University Press (2002)
3. Della Valle, E., Ceri, S., van Harmelen, F., Fensel, D.: It's a Streaming World! Reasoning upon Rapidly Changing Information. Intelligent Systems, IEEE 24(6), 83–89 (2009)
4. Ruta, M., Colucci, S., Scioscia, F., Di Sciascio, E., Donini, F.M.: Finding commonalities in rfid semantic streams. Procedia Computer Science 5, 857–864 (2011)
5. Ruta, M., Scioscia, F., Di Sciascio, E.: Enabling the semantic web of things: Framework and architecture. In: Semantic Computing (ICSC), 2012 IEEE Sixth International Conference on. pp. 345–347. IEEE (2012)
6. Scioscia, F., Ruta, M., Loseto, G., Gramegna, F., Ieva, S., Pinto, A., Di Sciascio, E.: A Mobile Matchmaker for the Ubiquitous Semantic Web. International Journal on Semantic Web and Information Systems (IJSWIS) 10(4), 77–100 (2014)