IOT Semantic Interoperability Workshop 2016

Position paper: Semantic Interoperability in Open Connectivity Foundation (OCF) – formerly Open Interconnect Consortium (OIC)

Author(s): Ravi Subramaniam (ravi.subramaniam@intel.com)

Submitted as nominated representative of OCF/OIC

Preamble

Semantic interoperability is important to the formation of an ecosystem of heterogeneous devices at any scale and is critical to realizing the full potential of the Wearable Computing and Internet of Things (WIOT).

The key aspects around semantic interoperability are:

- Clear definition of the semantics
- Unambiguous mapping between the semantics of the parties – within an ecosystem and across ecosystems
- Ability to implement easily to enable interoperability – both static or pre-defined interoperability and dynamic or late-binding interoperability

The author believes that there is a strong distinction between the "Internet of Things" (IOT) and "Things on the Internet" (TOTI). Many of the current use cases can be solved with the TOTI view for example cloud based data analytics where devices passively provide data that are aggregated and processed in the cloud. In such cases, solving the primary requirements of connecting these devices (securely?) and describing the data on these devices consistently is sufficient. On the other, IOT requires a tighter integration of multiple systems into a consistent and coordinated higher order system i.e. a "system of systems" view. While IOT view can solve all the TOTI use cases, there are additional use cases that require orchestration, control, coordinated interaction, determinism that require definition of and coordination of semantics at multiple levels.

The Open Connectivity Foundation (OCF) (formerly the Open Interconnect Consortium (OIC)) has defined principles, concept and framework to comprehensively support these semantic and "IOT view" objectives and is developing specifications that enables an interoperable ecosystem of IOT and Wearable devices.

This proposal does not aim to describe the entire spectrum of work done in the OCF but will focus on describing the parts that are relevant to a discussion on semantics and semantic interoperability.

Introduction

OCF Organization

The OCF is an organization of leading IOT companies and research institutions from around the world. In addition OCF works closely with a sister open-source project in the Linux Foundation called IoTivity to develop/contribute a reference implementation of the specifications. The OCF mission and focus is on broad system interoperability even though the "Connectivity" in the name may suggest a narrower focus.

OCF is organized as a Core working area (to build concepts and interoperability independent of the domain of application), vertical domain specific work areas (to build domain specific concepts, descriptions of "objects" and profiles), security work area (to build security models that support Core

principles and models), certification work area (to build the certification plan and test suite required to certify OCF interoperability).
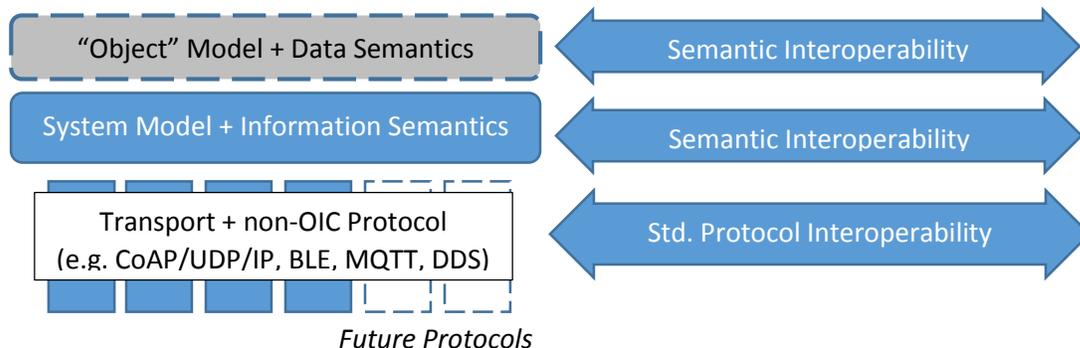
Overview

The OCF approach is Resource-oriented with a peer to peer RESTful architecture. The approach also follows a declarative paradigm which requires the explicit definition of information, data, semantics and objectives – these declarative statements are bound to imperative actions in a late-binding manner.

OCF architecture is organized logically into 3 areas of concern as shown in Figure: transports/non-OCF protocols, framework (includes information model) and object descriptions (includes data model) – each of these areas have concepts and support to make semantics explicit.

The framework abstracts the definition, description and interactions in a transport agnostic manner but maps to appropriate transports and non-OCF protocols based on the context and transaction needs (transport BOM, optimal transport for specific transaction, etc). In some cases OCF protocols are tunneled over other standard protocols. Furthermore the framework and defined abstractions allows for a rich information model and wider interoperability without "application" knowledge of protocols, maintains consistent semantics and allows for optimization. The framework could choose, from an available set, the transport or protocol that provides the richest set capabilities to map the semantics required for a specific interaction. The OCF approach spans multiple established protocols and transports (for example: CoAP/UDP/IP, BLE, XMPP, DDS and proprietary transports) with a goal of establishing a pan-transport ecosystem with well-defined abstractions and semantics.

OCF defines both interoperability and the semantics required for that interoperability at multiple levels – protocols, transports, framework model, data model and conceptual model.

Future Protocols

A broader overview and additional details on OCF (nee OIC) may be found at: http://openconnectivity.org/wp-content/uploads/2016/01/OIC_Specification_Overview_201501131.pdf

Semantics and Interoperability

The OCF framework defines the concepts, patterns and *formal semantics* required for establishing and describing context and interactions. The data model defines the data schemes and *lexical semantics* required to define and understand artefacts in the context. The conceptual model builds structures using the concepts and constructs established in the framework to describe the *conceptual semantics* between artefacts and contexts. OCF, therefore, considers "semantic interoperability" to exist at multiple levels. (NOTE: These forms of semantics don't match exactly with their definition in traditional linguistics but derive from them – formal semantics allows for logical reasoning on the behavior or outcomes in a system, lexical semantics use a lexicon-oriented approach with defined terms and

composition to describe objects/devices and conceptual semantics use structures (e.g. graphs) to capture meaning.)

While parts of the general description of OCF and aspects of semantics are still aspirational there has been strong progress towards these goals in many areas like the framework and data models as described in the OCF specifications. See http://openconnectivity.org/wp-content/uploads/2015/09/OIC_Specifications_1.0.zip and http://oneiota.org/documents?filter%5Bmedia_type%5D=application%2Framl%2Byaml

OCF Framework and Formal Semantics

The OCF framework is resource-oriented. The elements of formal semantics in the framework include denotational elements like Resources (and Resource Types), Interfaces, Collections, Links, (no formal logic but through spec descriptions), operational elements defined through patterns for example for discovery, introspection, notification etc. and other elements like Rules. The formal semantics that describe the model of computing are captured in normative text but some aspect of the semantics is also specified through a normative use of a meta-language like RESTful API Modelling Language (RAML).

With more of the formal semantics captured explicitly in a meta-language, OCF sees the possibility of semantics discovery and exchange. In this case, device can host OCF Resources that expose meta-descriptions which can be discovered using any of the discovery patterns described in the specification. This allows for late binding and dynamic adaption of devices for interoperability without a priori description of all aspects in text and code. This is also allows for backward and forward interoperability and minimizes the impact of "legacy".

To build interoperability with other ecosystems, OCF believes that there is sufficient flexibility in the formal semantics of the OCF framework to map the semantics of key elements of the "other" ecosystem into the OCF semantic model. This allows the OCF implementation to take on the "personality" of the other ecosystem without requiring in all cases the need for distinct boundaries between the two ecosystems and "bridge" systems. This is not a trivial exercise in most cases and may also require the lexical and conceptual semantic mappings too. For example mapping to an ecosystem based on an imperative model - Resources could be mapped to APIs, Collections to Services and Links to a logical bus and methods to RPCs. In addition, new media types can be enabled to indicate the expectations of the respective ecosystems and to communicate in a way to preserve the common semantics.

Data Models and Lexical Semantics

All Resources in OCF conform to Resource Types. A Resource Type represents a class of Resource but is also a model for capturing the lexical semantics that are relevant to that class.

OCF defines Resource Types for vertical application/business domains like Smart Home, Industrial, Healthcare, Automotive and others. The Resource Types capture the lexical semantics for both the fundamental capabilities required (*domain agnostic*) and the aggregation of these fundamental capabilities into higher-order functional units like appliances, devices etc (*domain specific*).

The lexical semantics are captured through a combination of normative text (specific to a domain) and the use of jSON Schema as a meta-description of the syntax required. The terms or lexical units used in the Schema are defined in normative text or is implied by context/association or natural language meaning *(This is an open area that needs more formalization as in a lexicon; this is not the same as but is complementary to the work OCF has done on data models and data model repositories like OneIota.org.)*

Within the OCF ecosystem the lexical semantics are conveyed by the Resource Types. As of this writing, the Resource Types are statics and predefined and the semantics are prescribed and understood from

normative text. OCF is also working toward a model of discovery and late-binding where manufacturers, developers or users can define the semantics of any artifact/concept/entity and make that available for discovery and use without prior registration with a centralized authority or clearing house. The semantics is defined by composing terms from a well-defined lexicon and declaring the composition as Resource Types which can be discovered.

(*The clear definition of terms and widely-accepted lexicon is an area for further OCF work – preferably done in collaboration with other standards organizations and is required for interoperability; additionally the compositional semantics specified in text need to be explicitly captured in a meta-language. One option is to adopt JSON-LD not as meta-data but as the (meta?)-meta-language for describing the meta-data that defines the OCF Resource Types. Processing of rich meta-language may not be possible by devices with resource-constraints below a threshold and so methods to manage this while providing rich late-binding is an area of work*).

For interoperability between ecosystems, OCF expects work to be done to define both explicit mappings between semantics and a coordinated effort for harmonizing semantics. Efforts are already underway and progress has been made to define this for UPnP ⇔ OCF and for OCF ⇔ OneM2M *(parts of this work is described in other workshop submissions by OCF)*

Information Model and Conceptual Semantics

OCF has defined Collection and Links which when combined with the lexical semantics of Resource Types are flexible to allow expression of conceptual semantics of various forms. A Link has a reference/target, a relationship and parameters to describe the link. A Link is unidirectional. A Collection is a Resource that contains Links.

The conceptual semantics that can be expressed using Resource Types, Collections and Links can define linked data, conceptual graphs of concepts (ontology, concept map), abstract representations of real/physical structure, a mapping or transformation, a declarative program etc. – these could be called conceptual semantics modes.

One of the first steps towards semantics interoperability across ecosystems is to establish equivalence between the two ecosystems. As interesting use of this OCF flexibility in conceptual semantic modes is to explicitly define the equivalence so that this mapping can be used computationally. This is done by using Collection/Links/Resource Types to frame the semantics in the interfaced ecosystems to a consistent model and then describe the semantic mappings from the OCF ecosystem to the other ecosystem and then, optionally, describing the transformations required as this mapping is traversed. This allows for general patterns for mapping between ecosystems rather than describing the mapping for every data object in the two ecosystems (tedious and potential brittle).

Examples of conceptual semantics of different can be provided/discussed at the workshop if required.

Ideas for cross industry collaboration

1. Need a broad effort to develop a common lexicon for capabilities and devices in IOT and Wearable computing (similar to a dictionary for natural languages) – the lexicon approach should be lightweight and easily deployed from the resource constrained to Cloud. For this approach to be successful it needs to reasonably comprehensive and widely adopted. Furthermore, the lexicon should be flexible and allow composability based on the needs of the usage context.
2. An approach that extends jSON-Schema using JSON-LD i.e. use JSON-LD to generate a jSON Schema. (NOTE: This is not the same as supporting jSON Schema in JSON-LD).