Semantics to the Shop Floor: Towards Ontology Modularization and Reuse in the Automation Domain

Christoph Legat ∗ Christian Seitz ** Steffen Lamparter*** Stefan Feldmann

∗ Institute of Automation and Information Systems, Technische Universität München, Garching near Munich, Germany
(e-mail: legat@ais.mw.tum.de, feldmann@ais.mw.tum.de)
** Siemens AG – Industry Sector, Nuremberg, Germany (e-mail: ch.seitz@siemens.com)
*** Siemens AG – Corporate Technology, Munich, Germany (e-mail: steffen.lamparter@siemens.com)

Abstract: The traditional way of engineering production systems requires static information flows within and in between automation software, from field control software to manufacturing operations management. The common understanding of data and its containing information is realized by implicit assumptions on data semantics. In contrast, the vision of future production systems as cyber-physical systems (CPS) focuses on intelligent production facilities, which are characterized by autonomous behavior and dynamic, cooperative interactions. As a consequence, data and information that are stored and exchanged within CPS cannot rely on the assumption that other software systems are aware of data semantics. A means to solve this issue is ontologies – an approach being intensely discussed and applied for enhancing data with semantics. However, reuse of ontologies within the automation domain is hampered as ontologies are developed for specific use cases without having reusability in mind. In this paper, these drawbacks of ontology development are discussed and an approach for maximizing reusability through modularizing ontologies for different fields of the automation domain is presented.

Keywords: Semantic technologies, knowledge-based systems, ontologies, cyber-physical systems

1. INTRODUCTION

The vision of future production systems as cyber-physical systems (CPS) focuses on intelligent production facilities, which are characterized by autonomous behavior and dynamic, cooperative interactions. This novel approach is highly disruptive for the industry domain as in traditional automation systems, components are composed hierarchically from field to enterprise level (IEC, 2003). The physical communication between these components is based on standards, e.g. PROFINET, PROFIBUS. The content exchanged is very often proprietary since a single plant is a closed system.

However, this situation completely changes with the introduction of CPS in the automation domain, as CPS may build a highly connected network. Thus, it may become possible that arbitrary components need to communicate with each other. In such a situation, proprietary communication standards and approaches fail because information sources are heterogeneous and communication standards focusing also on the semantics of the data are not established in the automation domain. A very intensely discussed and also applied approach for enhancing data with semantics and, thus, for describing the semantics of information sources to make the contents understandable for both humans and machines, are ontologies. Ontologies were introduced twenty years ago by Gruber et al. (1993) as an explicit “specification of a conceptualization”. An ontology provides a shared vocabulary, which can be used to model a domain by defining the objects and concepts that exist and specifying the properties and relations between these objects. Using appropriate ontology languages, e.g. Web Ontology Language (OWL), ontologies can be developed modularly and, hence, combined to complex ontologies, which is a major advantage of the approach.

Unfortunately, reality is different. Looking at reusability of existing ontologies in the automation domain, the results of twenty years of ontology development are not satisfying. Today, only a minority of existing ontologies may be – or are – reused. Regarding the results of research projects, novel ontologies are often developed for specific use cases, e.g. commissioning, maintenance or diagnostics, without reusing existing ontologies. For these reasons, in this paper, we propose an approach on how ontologies should be developed in future to guarantee a maximum of reusability. The key idea of the approach is modularization. We suggest fields in the automation domain, for which generic ontologies must be developed in order to achieve a basis to simplify the development of ontology-based applications.

The remainder of this paper is organized as follows: In the next section, the problem of ontology reuse in the
automation domain is discussed and weaknesses of existing approaches are exemplarily highlighted. In section 3, requirements for ontology development in the automation domain to be fulfilled in order to overcome existing reusability drawbacks are derived. Finally, the paper is concluded in section 4.

2. EXEMPLARY APPLICATION PROBLEMS OF ONTOLOGIES IN AUTOMATION

Various applications of ontologies in the automation domain were proposed. Dibowski et al. (2010) propose a device model for the building automation domain to discover devices that provide adequate interfaces and capabilities. In Legat et al. (2013b), an ontological model is applied to reason about reconfiguration capabilities of a system in case of re-engineering. Intensive research on dynamic reconfiguration, i.e. the adaptation of control behavior during a plant’s operation, was conducted in recent years. In Alsafi and Vyatkin (2010), an ontology for reconfiguring mechatronic systems was proposed. Feldmann et al. (2013) utilize ontologies to identify adequate functionality of a manufacturing system in order to realize a given production process. To manage the communication between holonic agents for an agile manufacturing control, an ontology is applied in Leitão and Restivo (2006). For system diagnostics, various approaches relying on formal semantic models were proposed, e.g. Jirkovsky et al. (2012); Hubauer et al. (2011). Detailed overviews on ontologies in the automation domain are available, e.g. Lastra and Delamer (2009); Legat et al. (2013a). All these approaches apply semantic technologies successfully for a specific application. However, reuse of existing ontologies is rarely addressed and only a few application examples are available. Orozco and Lastra (2006) already identified the need for modularity of ontologies in automation to simplify their application and proposed an ontology to describe mechatronic devices covering hardware and software aspects. Nevertheless, its reuse is hindered as the ontology is neither publicly available nor documented in detail. OntoCAPE (Morbach et al., 2007) is an example of a set of modular ontologies for chemical process engineering with intensive documentation. In the remainder of this section, challenges and drawbacks when reusing existing ontologies and standards are highlighted exemplarily.

2.1 Example for semantical equivalence with different vocabulary and relations

CAEX (IEC, 2008) is a standardized XML-based data format for storing and exchanging plant information. It enables the description of functional and structural aspects on a very abstract level. Functional aspects by means of roles enable to describe what a physical equipment is able to provide and which functionality is expected to be provided. Structural information are described in two different ways: Containments, i.e. which device is part of another one, can be expressed by means of an entity hierarchy. Furthermore, connections between devices through interfaces to model e.g. wiring information can also be expressed. In Abele et al. (2013), formal semantics of CAEX were proposed using OWL in order to enable automatic validation of the CAEX plant model. It was identified, that it is rather hard to capture all semantic aspects of the standard correctly as information are only available as informal textual standard and not described in detail precisely enough.

In Lohse et al. (2006), an ontology-based approach for modular reconfigurable assembly systems was presented. The proposed ontology covers structural aspects by distinguishing, analogously to CAEX, between containments and connections in between interfaces. Thus, the information content is equivalent to CAEX but is modeled using different vocabulary and relations. Similarly, Feldmann et al. (2014) combine Systems Modeling Language (SysML) and ontologies for ensuring compatibility among system components by identifying structurally incompatible interfaces. The information content is similar to CAEX but modeled using a SysML-based vocabulary.

2.2 Example for semantical inconsistency of different vocabularies

Functional aspects proposed by Lohse et al. (2006) are expressed by tasks which might have subfunctions, namely operations, which in turn might have subfunctions, called actions. An equipment taxonomy containing systems, cells, units, etc. whose semantics is defined by the capabilities to perform a limited number of functionality, e.g. units have at least one operation but cannot perform tasks, was proposed. The IEC 61512 (IEC, 1997) standard defines models and a terminology for batch control and is widely applied in industry. An ontological model which relies on this standard for monitoring and order sequencing was for example proposed in Lamparter et al. (2011). IEC 61512 defines structural aspects by means of a containment hierarchy for structuring physical equipment and defines, among others, units which can performs process stages, process operations and process actions.

Combining the ontology for reconfigurable assembly systems according to Lohse et al. (2006) with an ontology compliant to IEC 61512 is not possible because the semantics of a unit with respect to its functionality in the two ontologies is not consistent to each other. The reason is that both ontologies are tailored towards different application fields. They work well for the application field they were designed for but, however, did not consider reusability.

2.3 Example for terminological equivalence with different semantics

A variety of ontologies for device descriptions were developed inside and outside the automation domain. For the vision of CPS, a variety of ontological device descriptions in the research field of smart environments (Internet of Things, Internet of Everything) were proposed. For the automation domain, it was identified in Fantana et al. (2013) that the Internet of Things is of high potential for manufacturing. The ontological modeling of physical devices such as sensors was conducted intensively, e.g. the sensor net ontology (Compton et al., 2012), which was developed to standardize the semantic description of sensors. The use of ontologies for sensors and sensor data in the automation domain was e.g. proposed in Legat et al.
(2011) to improve the adaptability of system diagnostics. Most of these works focusing on physical devices define units as a structural entity.

The NASA \(^2\) developed a set of modular ontologies to model various aspects in the domain of earth and environmental observation (Raskin and Pan, 2005). They contain among others an ontology of units, which describes units of measurement, e.g. according to the SI standard, to describe the physical quantity of sensor measurements.

Both, units of measurement as well as units as structural element according to Lohse et al. (2006) or IEC 61512, are referred to as units. Here, the need for semantics is obvious to clarify terminology when reusing it in the automation domain.

3. REQUIREMENTS FOR ONTOLOGY DEVELOPMENT IN THE AUTOMATION DOMAIN

Ontology development in the automation domain is challenging due to heterogeneous application fields (e.g. industry, building, grid automation) and due to often strong dependencies between vendor-specific hardware devices and tooling. As exemplified in the previous section, the characteristics of the automation domain often lead to the development of highly specific ontologies addressing a certain use case, system structure and even device setup. While the developed ontologies typically fit well for few selected use cases the ontology engineer had in mind during the ontology development, they are difficult to reuse within other contexts. However, without extensive reuse of existing models, development of ontologies is extremely resource-intensive and expensive, which hampers business use cases of ontology-based automation solutions.

Therefore, a core challenge of future research in the area of ontology-based automation solutions is to reduce the modeling effort spent for realizing a specific solution by enabling reuse of models from an available model library. From our point of view, four requirements are crucial to reach this goal.

Requirement 1: A common standard language for describing ontologies in the automation domain

In order to combine models from different resources, they have to be expressed in the same language. The transformation of a model to another language is often very difficult – particularly if the source model is not formally described and the meaning of terms is therefore ambiguous. This is also the reason why automated methods for lifting models to a coherent common formalism are extremely difficult to realize. Avoiding the problem of heterogeneous modeling formalism by leveraging a standard modeling language seems to be a natural way to deal with the problem. This is also the approach taken by the World Wide Web consortium (W3C) in the context of their semantic web initiatives, which feature language standards such as the Resource Description Framework (RDF), the RDF Schema (RDFS) and the Web Ontology Language (OWL). The semantics of these languages is formally defined, which thus facilitates the definition of mappings between different ontologies. In addition, by leveraging Unique Resource Identifiers (URIs) with unique namespaces, modularization is natively supported by these ontology languages. This feature is particularly used by linked data initiatives such as Bizer et al. (2009) and could prove as a very useful property also in the automation domain. However, while RDF, RDFS and OWL provide the technical means for expressing automation modules in a standardized language, they do not guide the modularization of a domain.

Requirement 2: Consistent modularization of ontologies enabling effective reuse, refinement and extension

An effective modularization of a domain is a key for enabling reuse of existing ontologies. The goal is to avoid high overlaps between ontologies that have to be aligned while of course the domain should be covered as completely as possible. Ontology modules that are too comprehensive are difficult to align as contradictions and inconsistencies may arise. Furthermore, they are ineffective to refine/extend by the community. A too fine-grained modularization increases the complexity and makes the management of the models more complicated. Therefore, it is important for the automation community to identify the suitable module structure for the domain. Technologies and best practices regarding ontology modularization are already available and used in various other domains (cf. Stuckenschmidt et al. 2009). If a module structure becomes an agreed and common understanding in the community, ontology development could be streamlined and, hence, development costs could be reduced.

In addition to an effective module structure, reusability will improve if ontologies focus on contextualization of relevant data independently from use cases, i.e. modeling requirements should be derived solely from a data producers’ point of view and should not reflect any use-case-specific requirements regarding data provisioning. If this independency from the use case can be achieved, the reusability of ontology modules will be considerably higher. In the following, we exemplify an excerpt of a possible high-level module structure that is independent from specific use cases. However, the structure should be considered only as a starting point for further discussions in the corresponding academic as well as industry initiatives and standardization bodies.

(1) Description of physical objects is required to describe e.g. the equipment installed in a production plant, the parts available in stock or objects used for assembling a product. The description of physical objects consists basically of a taxonomy of devices, e.g. sensor, actuator, PLC, etc. Attributes, properties, and functionalities of a device are described by utilizing other ontology modules. For specifying compositions of physical objects, separate ontology modules should exist.

(2) Descriptions of structures contain information about how physical elements are composed, i.e. how physical objects are assembled or arranged. This information can be used for each assembled physical thing, e.g. production systems (e.g. construction), products (e.g. assembling) or logistics (e.g. palletization). As identified in the previous section, two major approaches can be distinguished. On the one hand, interface-based composition fa-

\(^2\) http://sweet.jpl.nasa.gov/ontology/
(3) **Functionality descriptions** are needed to concretize what is done, what is required to do or what can be done by a physical device or organizational unit. For example, the description of production facility tasks or tasks required to be executed to produce a specific good as well as for both, their specific realizations (e.g. referring to process descriptions) and properties, parameters and constraints (e.g. referring to auxiliary aspects like material or time). Classification of tasks into hierarchies allows matching descriptions specified on different abstraction levels.

(4) **Process descriptions** are necessary to model compositional aspects of functionality. They contain different process composition models and workflow patterns like different kind of sequences and parallelism (compare e.g. van der Aalst et al. (2003)). They are used to describe functional (required or existing) dependencies. In addition, some annotations like temporal aspects are possible by reusing required auxiliary ontology modules. Furthermore, process descriptions are required to overcome varying granularity of functionality descriptions by aggregation and splitting descriptions. Consequently, process descriptions are highly related to functionality descriptions but should be separated from it to ease reuse of different functionality and process models.

(5) **Observations and measurements** are essential to exchange information observed by (cyber-physical) systems about the physical world. In contrast to the ontology modules described previously, ontology modules for observation and measurement are related to data produced during a system’s operation. Traditionally, it is assumed in the automation domain that an engineered automation system and its applications remain more or less static and, consequently, no need for semantic description of observed data exist. For the vision of CPS, this assumption is not valid any more. For this reason, it is increasingly important to describe the semantics of observations and measurements. To describe observations, information about where and what is or will be observed as well as how it has been measured are required. In case of a consistent modularization, these aspects of observations and measurements are covered by a variety of other ontology modules such as physical properties and units, geometrical aspects, process descriptions, and so on. Since a holistic description of observations and measurements is very complex – compare e.g. the effort to standardize sensor-related information by the OGC Sensor Web Enablement (Botts et al., 2008) – modularization would facilitate effective development of ontology-based applications in the automation domain.

(6) When modularizing ontologies, **auxiliary ontology modules** play an essential role to describe processes, observations, structures (especially interfaces) and functionality. In the following, some important examples of auxiliary ontology modules are described to exemplify our vision without any claim to completeness.

(6.1) **Material descriptions** are required to organize material and, thus, must contain taxonomies of material as well as their characteristics. They might be reused for the description of a material a physical object (e.g. device or product) is or should be built of. In addition, for describing restrictions of a physical object’s functionality, material taxonomies might be relevant. For example, a specific milling cutter (i.e. device) can process (i.e. functionality) only wood, but not iron (i.e. functionality restriction).

(6.2) **Descriptions of physical quantities, dimensions and units** are required by nearly all ontology modules, e.g. description of material properties, spatial extent, physical phenomena a sensor is observing or an actuator is using. Properties and restrictions on functionalities might also reuse such ontology modules for describing e.g. the mass or dimensions of a product as well as the maximum possible mass that can be handled by a conveyor. Furthermore, a detailed model of physical quantities and their relations is necessary in the automation domain to describe relationships between functionalities and processes as well as agreements on specific units (e.g. kilogram or gram) or dimensions (e.g. length). In addition, support of composite units and dimensions is required to represent the correlations of physical quantities, dimensions and units. Ontology modules describing different temporal aspects, e.g. time points or intervals, belong also to this group of ontology modules because time is also defined as physical dimension with corresponding units.

A more detailed discussion on the modularization and language aspects of automation models can be found, e.g., in Abele and Grimm (2013). Once a suitable structure with complementing ontology modules is available, still work has to be done to select, integrate and maybe even extend modules relevant to a specific use case.
Fig. 2. Map of informational coverage of selected related work compared to proposed module structure

Requirement 3: A common core automation vocabulary that facilitates alignment of ontology modules to fit use case-specific requirements

To support the alignment of ontology modules, a common modeling paradigm is extremely helpful, which may e.g. be provided via a common top-level ontology. A top-level ontology covers a set of general terms that are domain-independent and used across many different applications including process, measurement and unit. Examples are top-level ontologies such as DOLCE (Gangemi et al., 2005). A first approach on aligning an existing ontology for manufacturing with DOLCE is described in Borgo and Leitão (2007). Although mappings based on formal ontologies could also be constructed without relying on top-level ontologies, a common conceptual basis of the different models largely facilitates the alignment process (Mascardi et al., 2010). That means if a standard automation top-level ontology was introduced as a basis for the different ontology modules, assembling a common use-case specific ontology would be much simpler and cost-effective. The assembling process is shown in Fig. 1.

Requirement 4: Automation ontology repository

In order to support the reuse of models, it is important to find the models that are relevant to a given setting and to select as well as align them appropriately. Therefore, tool support is required that includes on the one hand an automation ontology repository and on the other hand further tools to update and align these ontologies. The technology for ontology repositories has been developed and several repositories are publicly available. The W3C³ and the University of Manchester ⁴ provide overviews.

Our vision is not one of an ontology covering all aspects of the world but one of an ontology simplifying the reuse of ontologies developed all over the world whose fusion might cover most relevant information required for applications. To ease the development of applications based on ontologies, a global ontology repository plays an essential role. Instead of developing a novel ontology for the application in mind, an engineer first has to analyze and identify informational aspects essentially required for the application to be developed (cp. figure 1). Based on these required aspects, relevant existing vocabularies and their informational coverage can be automatically identified with the help of a global automation ontology repository as exemplarily depicted in Fig. 2. Multiple ontologies covering probably identical aspects is beneficial because each vocabulary inherits different strengths and weaknesses. If no reasonable semantics is required, a lightweight vocabulary is more suitable than one with high axiomatization, but for other applications, it might be essential. For example, for some applications, it is sufficient to use an ontology for units of measurement without any axioms only to refer to specific units on vocabulary level, but if it is essential to reason about consistency of such units, an ontology with high axiomatization level is beneficial. Consequently, the ontology to be reused for an application to be developed depends directly on the requirements on the application. For this reason, a standardization of ontologies, especially on fine-grained level, is neither required nor beneficial.

Conclusion of the requirements

A common standard language for ontologies in the automation domain (requirement 1) lays the foundation for combining these ontologies. Ensuring the consistency of different ontologies by an informational aspect-oriented modularization (cp. requirement 2) ensures the compatibility between vocabularies. Grounding each ontology on a common core vocabulary (requirement 3) facilitates the alignment of the modular ontologies. Appropriate tool support enables retrieval, alignment and update of the ontologies available (cp. requirement 4). Finally, the informational and terminological relationships between ontologies utilized for various applications will facilitate to launch novel applications of CPS with reduced effort since required operation data for example is directly available without additional development.

4. CONCLUSION

Today, there is a wide agreement that semantic modeling plays an increasing role for future automation systems. Twenty years of research in the ontology area emphasize this fact. Various ontology languages were developed and some are already standardized. Sophisticated reasoning mechanisms were developed to prove ontologies for consistency. Repositories to discover adequate ontologies and query languages to perform requests to a huge amount of semantically enriched data exist. Various ontologies were developed in the automation domain, but unfortunately – aside some exceptions – with limited reusability.

In this paper, we highlighted the lack of reuse in ontology development using exemplary and typical problems of ontology alignment in the automation domain. We proposed basic requirements needed to be fulfilled for better

³ http://www.w3.org/wiki/OntologyRepositories
⁴ http://owl.cs.manchester.ac.uk/tools/repositories/
reusability of ontologies during application development. At first, we recommended a common ontology language for avoiding unnecessary transformations between different languages. Furthermore, we proposed a modularization of ontologies providing ontology modules that can be combined for arbitrary applications. The combination of a minimal vocabulary with a top-level ontology was proposed to reach this goal. Finally, we suggested to establish a repository for ontologies in the automation domain for simplifying the identification of matching ontologies.

Using the proposed requirements, we encourage the community to start developing ontologies not specialized to certain use cases but with having reusability in mind.

REFERENCES


