Grounding of the Functional Object-Oriented Network in Industrial Tasks*

Rafik Ayari, Matteo Pantano† and David Paulius

Abstract—In this preliminary work, we propose to design an activity recognition system that is suitable for Industrie 4.0 (I4.0) applications, especially focusing on Learning from Demonstration (LfD) in collaborative robot tasks. More precisely, we focus on the issue of data exchange between an activity recognition system and a collaborative robotic system. We propose an activity recognition system with linked data using functional object-oriented network (FOON) to facilitate industrial use cases. Initially, we drafted a FOON for our use case. Afterwards, an action is estimated by using object and hand recognition systems coupled with a recurrent neural network, which refers to FOON objects and states. Finally, the detected action is shared via a context broker using an existing linked data model, thus enabling the robotic system to interpret the action and execute it afterwards. Our initial results show that FOON can be used for an industrial use case and that we can use existing linked data models in LfD applications.

I. INTRODUCTION

With the a priori defined 4th Industrial revolution (I4.0) the concept of connected factories is de facto standard for competitive advantages of European manufacturing [1], [2]. The concept envisages that the Factories of the Future (FoF) will be able to share information regarding processes and products that will enable life-cycle tracking of development, production, and services [3]. However, the realization of this vision is still far from reaching modern companies [4]. Automation will play an important role in the domain of FoF, and the National Institute of Standards and Technology (NIST) has outlined the main challenges in the manufacturing interoperability program of NIST [5]. Among those, the most important to our work is the misinterpretation of definitions or meaning of terms when exchanging information, which is mainly due to the lack of common data models that can be easily shared among factories [4]. Therefore, in this work, we would like to address this issue by analysing an I4.0 related industrial use case where a collaborative robot system needs to communicate with an activity recognition system to learn about a new task using Learning from Demonstration (LfD). To accomplish this, we will use the functional object-oriented network (abbreviated as FOON) [6], [7], [8] together with a semantically linked data model.

The paper is organised as follows. In Section [II] we describe prior work in this area and discuss the need for a semantically linked data model to improve interoperability considering a case use in LfD. Afterwards, in Section [III] we explain the addressed use case and the methodology used for the action recognition together with its grounded FOON data model. Finally, in Section [IV] we discuss our preliminary results and the future steps of our research.

II. BACKGROUND

A. Functional Object-Oriented Network

FOON is a symbolic knowledge representation for robots that takes the form of a semantic graph. Motivated by the theory of affordance [9], a FOON can be used to describe objects, actions, and their corresponding effects on those objects or on the robot’s environment. This is achieved through its bipartite network structure comprising of two types of nodes: object nodes and motion nodes. Object nodes describe any sort of object that is used in manipulation or activities, such as tools, ingredients or components, while motion nodes describe actions using human action labels, which can be either manipulation actions (e.g., pouring or cutting) or non-manipulation actions (e.g., baking or frying).

As a bipartite graph, object nodes are connected to motion nodes, while motion nodes are connected to object nodes. This innately allows a FOON graph to describe an object’s state change(s) as actions are executed, as a motion node’s incoming edge indicates a precondition, while a motion node’s outgoing edge indicates an effect. In FOON terminology, an action is described as a functional unit, which comprises of input and output objects (describing precondition and effects) and a motion node. A chain of functional units describing a single activity is referred to as a FOON subgraph (example shown in Figure [I]), and a FOON graph that is made from two or more subgraphs is referred to as a universal FOON.

B. Interoperability for the Factories of the Future

As previously mentioned, interoperability is the field addressing the perfect exchange of information between I4.0 actors (e.g., machine, human), where perfect means that accurate information exchange is guaranteed during the data life-cycle (i.e., origin, production, and use) [10]. To address this issue in the manufacturing domain, the so-called reference architectures (RAs) have been introduced by several standardization bodies. RAs are blueprints that contain and track all the information regarding processes’ life-cycles, thereby enabling faster reuse of information [11]. The most known RAs are: Reference Architecture Model...
Fig. 1: Illustrations of FOON subgraphs for: a) industrial assembly and b) disassembly of parts (best viewed in colour). Here, object nodes are denoted by circles, while motion nodes are denoted by squares. Regular object nodes are denoted in lime green, while intended goal nodes for each subgraph’s task is denoted in blue. For the assembly task, the goal is to attach a bracket to a strut profile and secure it with a T-bolt and flange nut, while for the disassembly task, the goal is to dismantle the secured bracket and strut profile.

III. USE CASE AND METHODOLOGY

A. Use Case

The use case analyses the assembly and disassembly of strut profiles. Strut profiles were chosen as they are conventionally used in the industry for building workcells (e.g., robotic workcell). The use case envisions an operator in front of a workbench with all the parts that needs to be manipulated as depicted in Figure 2. Considering the numbering in the Figure, the parts are: strut profile 1, bracket 2, T-bolt 3, and flange nut 4.

B. Creating a FOON for Industrial Tasks

FOON has been applied mainly to cooking activities. In this work, we apply the FOON representation to industrial tasks, specifically assembly and disassembly of parts required for preparing a robotic workcell. Such parts include

[https://opcfoundation.org/developer-tools/specifications-opc-ua-information-models]
Fig. 2: Image of the workbench with the operator and the parts used for the assembly process. Going from the left to the right, the parts are as follows: strut profile 1, bracket 2, T-bolt 3, and flange nut 4.

strut profiles, flange nuts, T-bolts, and brackets. For the time being, we constructed subgraphs for each task, and these are shown as Figures [a] and [b] respectively. These subgraphs describe several high-level skills or actions required to assemble components, such as picking-and-placing, screwing, and unscrewing. Skill definitions were taken from [22], [23]. Furthermore, to make such graphs suitable for action recognition, the notion of hands or end-effectors were included to better identify objects of focus when using hand recognition, which will be addressed in Section IV.

C. Modelling of the Ontology

To address the identified need of an activity description aligned to an industrial-like infrastructure, we opted to use a semantically linked activity information model. Such decision was driven by the possibility of easily mapping a semantically linked model to the OPC UA information model as outlined by [24]. Considering that we are using the FOON comprised of several functional units, our modelling proposes an approach to build and connect functional units. For our modelling, we start from the Task data model and Resources data models created by the Smart Human Oriented Platform for Connected Factories (SHOP4CF) consortium [25]. These information models were selected because they are based on linked data and they represent manufacturing use cases; therefore, they satisfy the requirements for mapping into OPC UA. The Task data model was chosen due to its similar definition to a motion node: a Task is intended to be a manufacturing operation that may contain sub-steps, and it is associated to resources (i.e., persons, devices, materials, and assets) and locations. Moreover, the Resources data models were chosen due to their definition of abstractly representing objects that are located on the shop floor. Accordingly, the mapping between functional unit, Task, and Resource attributes can be summarized as follows. Input and output objects are mapped to the Resource data models, as they can represent the state of the objects (i.e., state change). Motion nodes are represented using the Task data models, as they can specify which object is involved through the involves property, and they can integrate the chain of events by using the isDefinedBy property.

IV. WORK IN PROGRESS

A. Designing an Action Recognition Model

The aim of action recognition is to identify the action presently being performed by a human demonstrator or worker, which can be useful for collaborative robot (cobot) scenarios, where a robot needs to assist the human in completing a joint goal or task. For the time being, we plan to employ a deep neural network framework to process hand-object information extracted by a YOLO-based [26] pipeline to recognize actions as functional units in our FOON representation. From a functional unit, we will know which objects are being used and in which state(s) they may currently be in as it relates to the task. Afterwards, with this input information in the format of an object-to-object distance matrix, object-to-hand distance matrix, and image, we plan to recognize the actions by using a recurrent neural network due to the sequential nature of the activities in our use case. The pipeline can be visualized in Figure 3.

B. Connecting FOON to an Ontology

As stated in Section III-C we are using the linked information models Task and Resources. These models have been created using the NGSI-LD [27] specifications and have

![Fig. 3: Pipeline for the action recognition model. Initially, the classified images are scaled down, and the distances between object to object and object to hands are calculated and inserted into a matrix. Secondly, with the initial state of the scene, this matrix is fed to the recurrent neural network as input, which predicts the Task associated with it.](https://shop4cf.github.io/data-models/task.html)https://shop4cf.github.io/data-models/
been integrated within the FIWARE open-source platform. Therefore, the connection between the activity recognition and the information models will be integrated through this framework. Once the predicted unit is estimated via action recognition, the following information is added to the linked model. First, the unit is created by a new Task model. Second, the objects related to the Task are added to involves property. Third, the Task attribute isDefinedBy is updated with a sequential ID referring to the modelled FOON above. Finally, the object state is updated according to the perceived state through the attribute state and the Task with Resources are published on the FIWARE context broker. This approach is described by Algorithm 1.

Algorithm 1 Connection of the FOON to the ontology

1: procedure FOON2ONT(PredictedUnit)
2: Resources = []
3: Task = []
4: for each object obj in PredictedUnit do
5: Resource[state] = obj[state]
6: Resources.append(Resource)
7: Task[involves] = obj[id]
8: end for
9: Task[isDefinedby] = PredictedUnit[id]
10: publish(Task)
11: publish(Resources)
12: end procedure

C. Future Work

Once we have established a basic connection between the FOON representation and the ontology via FIWARE, we will further investigate ways to construct new graphs for other types of parts and unseen tasks. This can be achieved by treating object detection and motion recognition as separate pipelines. Moreover, a definition of the set of activities will follow in order to ensure a common representation of the set of activities both from the operator and the robotic system. This can be achieved by a common representation of the isDefinedBy, which will point to a new entity Task Definition with a pre-defined and explained vocabulary (e.g., JSON schema vocabulary, SOMA ontology) of actions.

REFERENCES

[1] Forschungsunion and Acatech. Recommendations for implementing the strategic initiative INDUSTRIE 4.0 securing the future of German manufacturing industry: Final report of the Industrie 4.0 working group.
[2] Mario Herrmann, Tobias Pentek, and Boris Otto. Design Principles for Industrie 4.0 Scenarios. In 2016 49th Hawaii International Conference on System Sciences (HICSS), pages 3928–3937. IEEE, 2016.
[3] Deutsches Institut für Normung. Referenzarchitekturmödell Industrie 4.0 (RAMI4.0), 2016.04.
[4] Christoph Legat, Christian Seitz, Steffen Lamparter, and Stefan Feldmann. Semantics to the Shop Floor: Towards Ontology Modularization and Reuse in the Automation Domain. IFAC Proceedings Volumes, 47(3):3444–3449, 2014.
[5] Sharon J. Kemmerer. Manufacturing Interoperability Program, a Syn- opsis. NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, Gaithersburg, MD, 2009-02-24 2009.
[6] D. Paulius, Y. Huang, R. Milton, W. D. Buchanan, J. Sam, and Y. Sun. Functional Object-Oriented Network for Manipulation Learning. In 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 2655–2662. IEEE, 2016.
[7] D. Paulius, A. B Jelodar, and Y. Sun. Functional Object-Oriented Network: Construction and Expansion. In 2018 IEEE International Conference on Robotics and Automation (ICRA), pages 5935–5941. IEEE, 2018.
[8] D. Paulius, K. S. P. Dong, and Y. Sun. Task Planning with a Weighted Functional Object-Oriented Network. In 2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021.
[9] J.J. Gibson. The theory of affordances. In R. Shaw and J. Bransford, editors, Perceiving, Acting and Knowing. Hillsdale, NJ: Erlbaum, 1977.
[10] Thomas C. Ford, John M. Colombi, Scott R. Graham, and David R. Jacques. Survey on interoperability measurement. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH.
[11] Soo-hung Han. A review of smart manufacturing reference models based on the skeleton meta-model. Journal of Computational Design and Engineering, 7(3):323–336, 2020.
[12] International Organization for Standardization. ISO/TR 23087:2018: Automation systems and integration — The Big Picture of standards, 2018.
[13] Yan Lu, K. C. Morris, and Simon Frechette. Current Standards Landscape for Smart Manufacturing Systems.
[14] Luis Alberto Cruz Salazar, Daria Ryashentseva, Arndt Lüder, and Birgit Vogel-Heuser. Cyber-physical production systems architecture based on multi-agent’s design pattern—comparison of selected approaches mapping four agent patterns. The International Journal of Advanced Manufacturing Technology, 105(9):4005–4034, 2019.
[15] Stefan Profanter, Ayhan Tekat, Kirill Dorofeev, and Markus Rickert, editors. OPC UA versus ROS, DDS, and MQTT: Performance Evaluation of Industry 4.0 Protocols, 2019.
[16] Hirekumar Nakawala, Paulo JS Goncalves, Paolo Fiorini, Giancarlo Ferringo, and Elena De Momi. Approaches for action sequence representation in robotics: a review. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 5666–5671. IEEE, 2018.
[17] Sumaira Manzoor, Yuri Goncalves Rocha, Sung-Hyeon Joo, Sang-Hyeon Bae, Eun-Jin Kim, Kyeong-Jin Joo, and Tae-Yong Kuc. Ontology-based knowledge representation in robotic systems: A survey oriented toward applications. Applied Sciences, 11(10):4532, 2021.
[18] Moritz Tenorth and Michael Beetz. KnowRob: A knowledge processing infrastructure for cognition-enabled robots. The International Journal of Robotics Research, 32(5):566–590, 2013.
[19] Michael Beetz, Moritz Tenorth, and Jan Winkler. Open-EASE. In Robotics and Automation (ICRA), 2015 IEEE International Conference on, pages 1983–1990. IEEE, 2015.
[20] Markus Waibel, Michael Beetz, Javier Civera, Raffaello d’Andrea, Jos Elfring, Dorian Galvez-Lopez, Kai Häussermann, Rob Janssens, JMM Montiel, Alexander Perzylo, et al. RoboEarth. IEEE Robotics & Automation Magazine, 18(2):69–82, 2011.
[21] Mina Saghafian, Karin Laumann, and Martin Rasmussen Skogstad. Stagewise Overview of Issues Influencing Organizational Technology Adoption and Use. Frontiers in psychology, 12:630145, 2021.
[22] J. Backhaus and G. Reinhart. Digital description of products, processes and resources for task-oriented programming of assembly systems. Journal of Intelligent Manufacturing, 28(8):1787–1800, 2017.
[23] Roman Froschauer and Rene Lindorfer. Workflow-based programming of human-robot interaction for collaborative assembly stations.
[24] Rainer Schiekofer, Stephan Grimm, Maja Milicic Brandt, and Michael Weyrich. A formal mapping between OPC UA and the Semantic Web. In 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), pages 33–40. IEEE, 2019.
[25] Michal Zimniewicz. Deliverable 3.2 - SHOP4CF Architecture.
[26] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You Only Look Once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
[27] European Telecommunications Standardization Institute. ETSI GS CIM 009 V1.2.2: Context Information Management (CIM), 2020.