Bringing a Natural Language-enabled Virtual Assistant to Industrial Mobile Robots for Learning, Training and Assistance of Manufacturing Tasks

LI, Chen; Hansen, Andreas Kornmaaler; Chrysostomou, Dimitrios; Bøgh, Simon; Madsen, Ole

Published in:
The 2022 IEEE/SICE International Symposium on System Integration

DOI (link to publication from Publisher):
10.1109/SII52469.2022.9708757

Publication date:
2022

Document Version
Accepted author manuscript, peer reviewed version

Bringing a Natural Language-enabled Virtual Assistant to Industrial Mobile Robots for Learning, Training and Assistance of Manufacturing Tasks

Chen Li[1], Andreas Kornmaaler Hansen[1], Dimitrios Chrysostomou[1], Simon Bøgh[1], and Ole Madsen[1]

***Abstract*— Nowadays, industrial companies want to enhance their Industry 4.0 competencies. Therefore, they need to help employees master state-of-the-art technologies and gain the necessary knowledge to stay relevant and competitive. As a result, there is a global demand for learning and training tools that assist the employees at all levels. In this paper, we propose a natural language-enabled virtual assistant (VA) integrated with an industrial mobile manipulator to fulfill this target in manufacturing tasks. The latest *Learning, Training, Assistance - Formats, Issues, Tools (LTA-FIT)* model is leveraged to guide the design and development of a pilot version of the VA. To validate its performance, three manufacturing scenarios are analyzed based on the learning, training, and assistance phases, respectively. In our system, the human-robot interaction is achieved through conversation and a dashboard implemented as a web application. This intuitive interaction enables operators of all levels to control a industrial mobile manipulator easier and use it as a complementary tool for developing their competencies. The pilot experiments show that the proposed VA is able to respond to operator commands flexibly within the LTA-FIT model.**

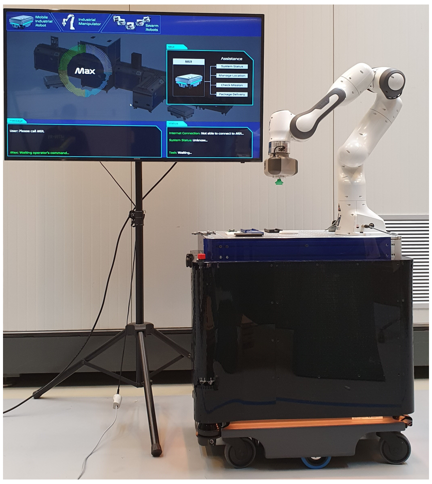

Fig. 1. The autonomous industrial mobile manipulator "Little Helper" where our proposed virtual assistant is integrated.

I. Introduction

The adaptation of Industry 4.0 is a challenging task due to the complexity of manufacturing supply chains and the speed of technological development. Companies must learn new technologies and develop new products, processes, and services with ever-increasing frequency. Furthermore, many of the solutions involve multi-disciplinary activities involving experts, which may not be present in the companies.

As a result, we expect to see changes in manufacturing workers' tasks and the type of knowledge and skills required to engage with new technologies [1]. These include technical competencies (e.g., process understanding and media and technical skills), personal competencies (e.g., problem-solving and motivation to learn), and social competencies (intercultural and teamwork skills) [2]. Organizations willing to move forward with Industry 4.0 should create strategies to enable their workers to develop these competencies in realistic manufacturing tasks [3]. Such strategies could involve work-based learning techniques [4], the use of learning factory environments resembling real production tasks [5] or through intuitive interfaces using natural speech [6].

[1]Chen Li, Andreas Kornmaaler Hansen, Dimitrios Chrysostomou, Simon Bøgh and Ole Madsen are with Department of Materials and Production, Aalborg University, 9220 Aalborg, Denmark {cl, akhan, dimi, sb, om}@mp.aau.dk

A cornerstone of our work is the *Learning*, *Training*, *Assistance - Formats*, *Issues*, *Tools* (LTA-FIT) model proposed by [7]. The model specifies the requirements that can qualify workers and guide companies toward a digital transformation. In this work, we present a natural language-enabled virtual assistant (VA) that provides a digital implementation of the LTA-FIT model in a smart factory. The approach leverages a state-of-the-art (SOTA) pre-trained language model i.e., Bidirectional Encoder Representations from Transformers (BERT) [8] model for natural language understanding, knowledge graphs for information retrieval and integrates them in an autonomous industrial mobile manipulator (Figure 1). The overall system provides instructions to facilitate a basic understanding of the steps of an assembly process, assists the workers through the necessary training process, and then answers workers' task-related questions.

In the remainder of this paper we present related works in Section II and describe the system architecture of the VA in Section III while we present the pilot studies in Section IV and discuss findings and future work in Section V.

II. Related Work

Work-based learning is a broad term for activities related to learning in a workplace environment or learning generated directly through workplace considerations and concerns. It has been proven more effective than traditional classroom teaching, mainly when centered around real work issues and mediated through the workplace [4]. Based on work-based learning principles, learning factories have proven effective in manufacturing in recent years [5]. They are usually formed as close collaborations between industry and universities to advance engineering education [9], transfer Industry 4.0 competencies to companies, and introduce artificial intelligence (AI) techniques [10]. However, learning factories have often been criticized for representing an idealized industrial environment with a narrow focus on specific tasks. To address this concern, the LTA qualification model was developed by Rehe et al. [7].

In recent years, voice-enabled interaction became an innovative and intuitive way to engage humans with everyday tasks. Powered by natural language processing (NLP) techniques, various commercial language-enabled VAs have been developed, e.g., Amazon's Alexa, Google's Assistant, and Apple's Siri, to support a more natural and interactive linguistic interaction in domestic environments. Additionally, in the manufacturing domain, several researchers have proposed chatbots for the training of employees [6], have built language-enabled VAs to support the training process of industrial plant operators [11] and enabled operators to complete the ramp-up process of an assembly system [12].

Many researchers also utilize verbal cues to facilitate Human-Robot Collaboration (HRC). For example, several models for voice control, including logical and semantic networks, have been to control an industrial robot in the context of collaborative assembly [13]. Besides, deep neural networks have been evaluated for the classification of commands in a natural speech recognition system for the interactive control of a robot manipulator in industrial collaborative tasks [14]. Finally, a language-enabled VA has been presented to assist the workers in handling a variety of complex manufacturing scenarios [15] while a novel language-enabled VA has been developed to assist shop floor workers with internal logistics [16].

In our work, we bring an intelligent virtual assistant closer to work-based learning. It is powered by natural language processing combining two disciplines for the first time, namely a Question Answering system [17] (to assist with a learning scenario) and a Task-Oriented Dialogue System [17] (to support training and assistance scenarios). Our long-term aim is to enable easier HRC in real industrial cases.

III. System Architecture

In this work, we design and develop a VA, called *"Max"*, as a web-based application based on the Flask web framework[1]. It involves three main actors: 1) the *Max Client*, devoted to the translation of the operator's verbal commands, showing the robot's status, displaying the response, and controlling the shop floor robots; 2) the *Max Server*, committed to serve the interpretation of the operator's requests, ground verbal commands to robot's actions and generate the corresponding response; 3) the *robotic platform*. Figure 2 shows the overview of the system architecture.

[1]https://flask.palletsprojects.com/en/2.0.x/

A. Max Client

The Max Client is implemented through three submodules: the web interface, designed to display real-time information, the cognitive voice service, and the robot control service.

Web interface: Max can verbally respond to the operator's questions or commands while showing the text-based response or related system status simultaneously to enhance the operator's experience during training. Using a web-based interface, allows the operators to browse Max's services on any web browser supported device, e.g., mobile phone, tablet, and PC. Figure 3 displays the main interface on a tablet, together with the two robot service menus. The *Dialogue Panel* shows the response from Max according to the operator's questions or commands. The *Learning* scenario heavily relies on this panel to provide instructions and the basic knowledge to the operator. A picture of the robot and a list of the robot services are wrapped into the *Robot Service Panel*, which provides additional information and indicates what services are currently supported for the desired robot platform. The *System Status Panel* is mainly used in the *Training* and *Assistance* scenarios, where it displays the real-time information of the tasks or system, e.g., network connection, which task is running.

Cognitive voice service: In order to provide a natural communication environment, two Microsoft cognitive services, speech-to-text (STT) and text-to-speech (TTS), are introduced into Max. The audio clips of the operator's spoken commands are collected through microphone and streamed to the STT services. The translated transcript is sent to the Max server as the input of the language service. Besides displaying the text response in the dialogue panel, the TTS service is leveraged to generate near-human voice responses to the operator.

Robot control module: In our application, the robot control module (RCM) controls a robot arm, Franka Emika Panda[2] and a mobile robot, MiR200[3]. RCM extracts commands and responses from Max's server, invokes predefined Python scripts, and routes function calls through REST APIs to the robot controllers. The backend threads of the webserver of Franka Emika and MiR listen for commands on the in-going TCP/IP socket as shown in Figure 4.

B. Max Server

The Max server combines three functionalities, 1) *language service*, devoted to recognizing the operator's intent and the keywords of the utterance, 2) *grounding*, grounds each phrase tuple extracted by language service to a robot operation action or entity/relationship, and 3) the *response generation*.

Language service: Our language-enabled VA supports a hybrid method, the BERT model and rule-based keyword extraction, to achieve the recognition of the operator's intent. *Learning* and *Training* scenarios mainly rely on the

[2]https://www.franka.de/robot-system

[3]https://www.mobile-industrial-robots.com/

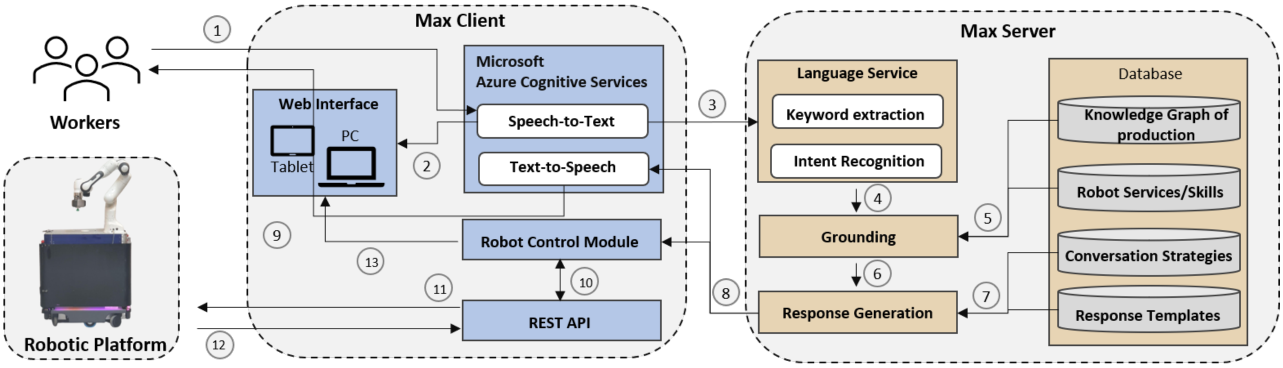

Fig. 2. An overview of the system architecture of the proposed virtual assistant.

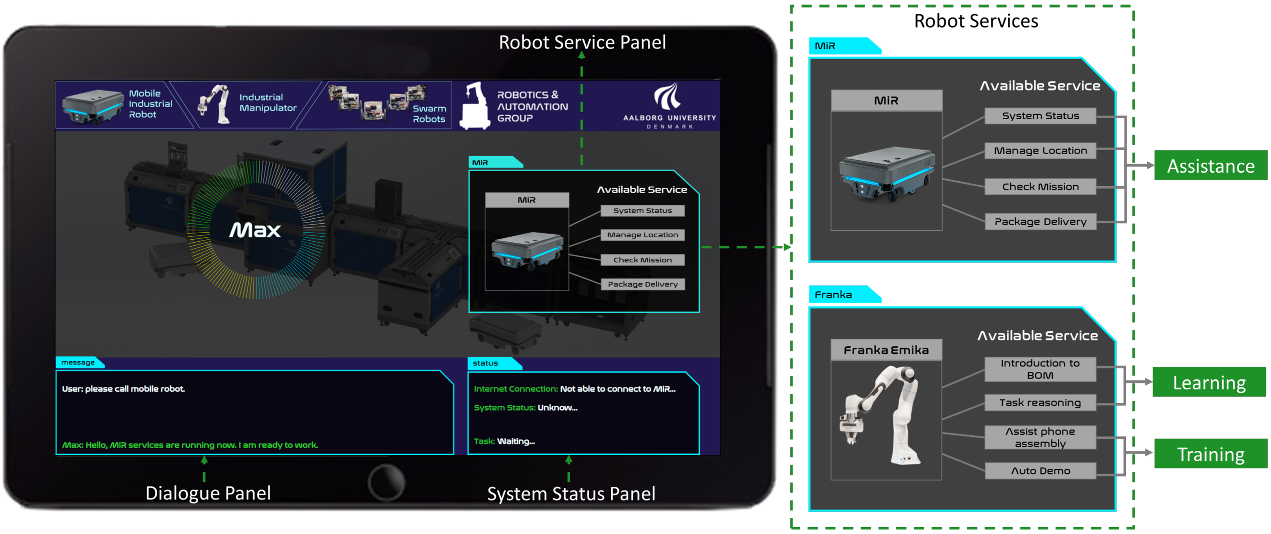

Fig. 3. The web-based user interface of the Max Client.

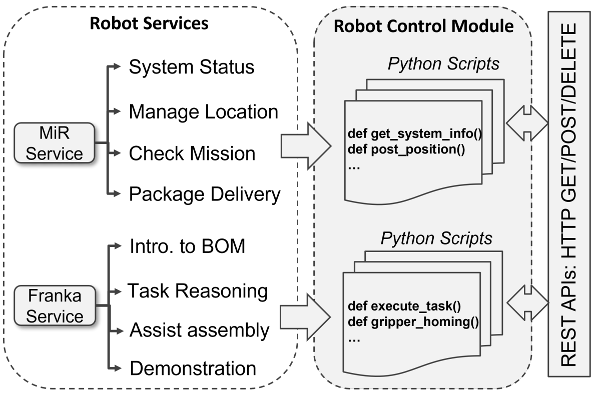

Fig. 4. A robot control module with predefined Python functions for calling the robot services through REST APIs.

keyword extraction method due to the high simplicity and low response time. However, for the *Assistance* scenario, Max should be able to accurately understand the operator's intent and commands so it can assist with daily logistic tasks. Therefore, we fine-tuned the BERT model to encode all the intents, slots and slot values (annotated with Inside-Outside-Beginning (IOB) tags) of the current operator's utterance into an embedded representation. The model architecture is illustrated in Figure 5. The operator's utterance, U_t is concatenated with two special tokens, [CLS] (standing for the classification and indicating the beginning the sequence) and [SEP] (denoting the separator of a sequence), resulting in I_t which serves as input into the utterance encoder:

$$I_t = [CLS] \oplus U_t \oplus [SEP] \tag{1}$$

The output of the BERT model is the predicted intent class label and the IOB tags of the given input. The BERT model is a masked language model that randomly masks 15% of input tokens and predicts the next sentence. It is primarily used in NLP tasks, e.g., intent classification, slot filling. After training on our dataset, the model achieves 0.977 and 0.968 on intent accuracy and slot F1 score respectively.

Natural language grounding: Natural language grounding plays an essential role in matching a word to the correct representation in terms of terminology and relationship, for the *Learning* and *Training* scenarios, and assigning a verbal command to the corresponding robot's action representation for the *Assistance* scenario.

To support *Learning* scenarios, a knowledge graph (KG), which stores interlinked descriptions of phone production,

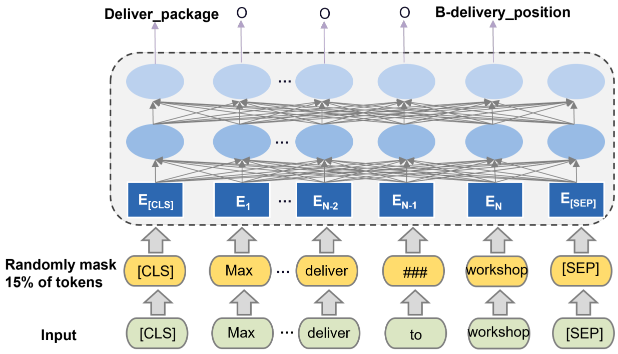

Fig. 5. The overall architecture of the fine-tuned BERT with exemplary inputs from a package delivery task.


```
{"services":[
        {
          "service name":"MiR",
         "skills":{"greeting", "missioncheck", "positioncheck",
                   "batterycheck", "askhelp", "statestop", "staterun",
                   "delivery"}
          },
          ...
        ]
}
```


Fig. 6. Exemplary structure of robot services and skills in JSON format.

is designed based on Neo4j[4] system. It defines the entities (e.g., materials of producing a phone, operation process), attributes (e.g., color, size) and relationships (e.g., pair, need). The grounding process can be defined as: given a corpus of operator's utterance U which contains n extracted keywords $K = \{k_1, ..., k_n\}$, find a partial function f to map the extracted keywords from operator's utterance to their representations in the graph, $f : K \rightarrow KG$.

Furthermore, a predefined set of categories of robot services and skills is stored as a JSON file in the database. It lists the supported robot services and skills for the *Training* and *Assistance* scenarios. Although it currently only focuses on the control of the Franka Emika and MiR, it can be extended to other robots or even a production line [15]. Figure 6 lists the different skills that are used to compose commands for controlling the MiR robot. The output of keywords extraction/BERT model, predicted operator's intent (i.e., requested robot service and skills), is used to match to the listed robot service and skills.

Response generation: In our case, template-based natural language generation is leveraged to map the non-linguistic input (i.e., extracted or predicted keywords from the language service) directly to the linguistic surface structure. To provide a more natural interaction with robots, Max also supports two human-to-human conversation strategies, namely **Lexical semantic strategy** and **General diversion strategy** [18]. By following the *Do not repeat yourself* rule of lexical-semantic strategy, Max can respond in different ways even it is given the same questions. Furthermore, Max can also initiate activities and switch topic during the interaction to enhance user engagement. By applying such strategies, Max provides diverse responses for creating a humanized dialogue environment and increases user favorability.

C. Robotic Platform

We deploy our VA system on the latest iteration of the autonomous industrial mobile manipulator Little Helper (LH) presented in [16]. Little Helper is a family of industrial mobile manipulators combining industrial robotic arms with mobile platforms to offer flexibility and versatility to industrial shopfloors [19]. It has been designed and built to work alongside operators and enable them to accomplish difficult and dangerous tasks such as bin picking [20] and tele-operation in safety critical applications [21].

The latest generation of LH consists of a Franka Emika collaborative robotic arm mounted on a mobile platform (MiR) (Figure 1).In our case, the Franka Emika robot arm is mainly used in the *Training* scenario to demonstrate how to assemble the mock-up phone and guide the operator in hands-on, stepwise assembly using grasping actions. At the same, the *Assistance* part is demonstrated based on the MiR mobile robot where location management for daily internal logistic tasks is considered.

IV. Pilot Studies

The pilot studies are set up at a Smart Factory, a learning factory platform for companies and researchers to collaborate on a real, physical production setup located at Aalborg University [22]. Six tasks are identified (shown in Table I) within three scenarios. Based on the definition of the LTA-FIT model we set up scenarios for *Learning*, *Training* and *Assistance*.

In the *Learning* scenario, the VA can teach the production terminology and the process of the assembly of mock-up phones to the worker. In the *Training* scenario, the VA performs as a Task-Oriented Dialogue System. Based on the dialogue with the operator, the VA controls the industrial mobile manipulator and demonstrates the required steps for the assembly of the mock-up phone. Finally, in the *Assistance* scenario, we evaluate the performance of the VA when it assists the operators' daily work and specifically during internal logistics procedures. Links to video demonstrations for all tasks and scenarios are provided in the respective footnotes.

A. Learning: introduction to product

In this scenario, Max needs to give an introduction to a phone assembly task. The information includes product definition, bill of materials (BOM), and production processes. For this scenario, two tasks (task 1 and 2 of Table I) are identified and tested: *1) learning basic terminology*[5] and *2) identify operational relationships*[6], based on a mock-up mobile phone assembly. The phone consists of five components; a top cover, a bottom cover, two fuses, and a printed circuit board (PCB), as Figure 7 depicts. The set up of the

[4]https://neo4j.com/

[5]Video: Task 1 - Learning basic terminology (Learning)

[6]Video: Task 2 - Identify operational relationships (Learning)

TABLE I
IDENTIFIED TASKS FOR THE THREE PROPOSED SCENARIOS.

Category	Task Description	Intent/slot Recognition	Conversation Strategy
Learning	1: Explain terminology	Keywords extraction	None
Learning	2: Reason operation relationships	Keywords extraction	None
Training	3: Demonstrate auto phone assembly	Keywords extraction	None
Training	4: Guide operator to assemble a phone	Keywords extraction	None
Assistance	5: Assist package delivery	Fine-tuned BERT Model	None
Assistance	6: Check system status	Fine-tuned BERT Model	Lexical semantic & General diversion strategy

Fig. 7. The parts of the mock-up mobile phone.

environment for the pilot experiment is shown in Figure 1. To provide appropriate instruction and education for the learning purpose, a knowledge graph is built for storing the master data of phone assembly, i.e., definition and attribute of each component and relationships among components. Figure 8 shows the graphical representation of the knowledge graph for the mock-up mobile phone in our case by using Neo4j platform. The relationships identified here are *needs*, *on* and *pair*. For example, relationship between *Top Cover* and *Bottom Cover* is defined as a *pair*.

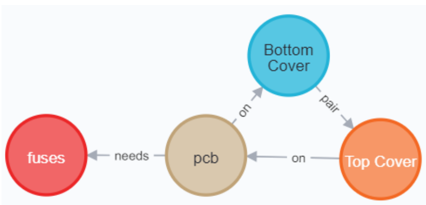

Fig. 8. Designed knowledge graph of the mock-up phone

A Question Answering example is shown in Figure 9, where the user requests an explanation of the term *housing*. The grounding result of the term *housing* is *Bottom Cover*, and its definition is retrieved from the knowledge graph. In Figure 9, the operator also queries the operational relationship between *Fuse* and phone *Cover*. According to the knowledge graph, *fuses* do not have a direct relationship with either the *Bottom Cover* or *Top Cover*. Therefore, *Unmatched* is returned as the grounding result. The text response is generated and sent back to the Max client. The dialogue panel in the web-based dashboard shows the user's question and the system's text response. The system status panel displays the category of the user's question, e.g., Learning BOM: Bottom Cover.

B. Training: phone assembly

To deepen the obtained knowledge from the above experiments of the *Learning* scenario, two types of tests are organized in *Training* scenario to train the operator to assemble a phone. *A demonstration of automatic phone assembly*[7] and a robot guided *hands-on exercise*[8] are included (see tasks 3 and 4 of Table I). The robot arm is programmed to assist the tasks such as grasping and moving objects.

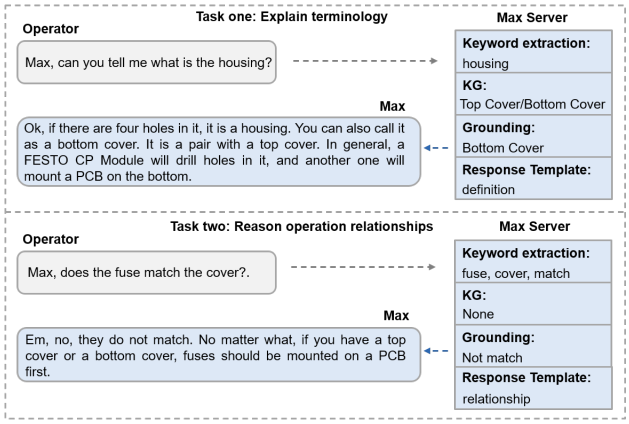

Fig. 9. Examples of two used cases: question the definition of the term *Housing* and operational relationship of *Fuse* and *Cover*.

Max leverages the defined robot service/skill to ground the user's utterance in a different fashion from the *Learning* scenario. The robot service is grounded as *Franka* in both tests, and the *demonstrate* and *guide* instructions, identify the required skills respectively.

C. Assistance: internal logistics

The *Assistance* module is designed for assisting the operator in a higher-level logistic task. Two tasks i.e., delivering package[9] and tracking robot status[10], are identified and tested in this scenario (tasks 5 and 6 of Table I).

The MiR 200 mobile robot (bottom part of the LH robot seen in Figure 1) is used as a robot entity to perform the transportation task. The VA, Max, continuously acquires the task related information from the operator and invokes robot control commands. In task 5, Max should deliver an object to an operator who works at a place in the smart factory. In this task, a 2D digital map is built on the MiR web server representing on the smart factory's layout. Places, e.g., warehouse, marked on the digital map are consistent with physical locations of the environment.

[7]Video: Task 3 - Automatic phone assembly (Training)
[8]Video: Task 4 - Hands-on exercise (Training)
[9]Video: Task 5 - Delivery of object (Assistance)
[10]Video: Task 6 - Check system status (Assistance)

In particular, the Max server invokes the BERT model, instead of extractig keywords, to predict the requested operator's intent and key slot values with a high accuracy. Max uses two turns dialogue to obtain all the required slots, i.e., object, recipient, destination, for performing the task. The identified intent and slots are grounded as the corresponding robot skill and parameters for API calls respectively. Besides, the operator may need to be aware of whether the robot can perform a task with the current system status if low battery is detected, or which mission is running at this moment. This is the motivation behind the sixth task. Instead of rejecting or committing to the task, Max can inform the operator what it is capable of doing at that specific instance. It utilizes *general diversion strategy: switch topic* to provide suggested tasks to the operator. Such suggestions are generated based on the results of the battery level checking running in the background. Therefore, with the automatic system status checking, the operator can schedule tasks more efficiently.

V. Discussion & Conclusions

In this work, we presented a natural language-enabled VA integrated with the LH robot platform to perform several Learning, Teaching, and Assistance scenarios in a real manufacturing context. The pilot experiments showcase that it can achieve natural and flexible interaction during task execution. The encouraging results based on the two embedded conversation strategies prove that Max can support an active interaction during various manufacturing tasks. It provides task-related suggestions, and forms a diverse and thoughtful dialogue to improve user engagement in HRI for industrial robots. Unfortunately, an extensive user usability study was postponed due to COVID-19 related restrictions; however, it remains a central part of our future work to collect comments and impressions about the user satisfaction and usability of the VA from shop floor operators who will be invited to participate in the learning/training/assistance scenarios.

Acknowledgments

The authors would like to acknowledge support by EU's SMART EUREKA programme S0218-chARmER and the H2020-WIDESPREAD project no. 857061 "Networking for Research and Development of Human Interactive and Sensitive Robotics Taking Advantage of Additive Manufacturing – R2P2".

References

[1] C. Adam, C. Aringer-Walch, and K. Bengler, "Digitalization in manufacturing–employees, do you want to work there?" in *Congress of the Intern. Ergonomics Association*. Springer, 2018, pp. 267–275.

[2] L. B. Liboni, L. O. Cezarino, C. J. C. Jabbour, and N. O. Stefanelli, "Smart industry and the pathways to HRM 4.0: implications for SCM," *Supply Chain Management*, vol. 24, no. 1, pp. 124–146, 2019.

[3] R. Belinski, A. M. M. Peixe, G. F. Frederico, and J. A. Garza-reyes, "Organizational learning and industry 4.0 : findings from a systematic literature review and research agenda," *Benchmarking: An Intern. Journal*, vol. 27, no. 8, p. 23, 2020. [Online]. Available: https://doi.org/10.1108/BIJ-04-2020-0158

[4] S. Lester and C. Costley, "Work-based learning at higher education level: value, practice and critique," *Studies in Higher Education*, vol. 35, no. 5, pp. 561–575, 2010. [Online]. Available: https://doi.org/10.1080/03075070903216635

[5] J. S. Lamancusa, J. E. Jorgensen, and J. L. Zayas-Castro, "The learning factory—a new approach to integrating design and manufacturing into the engineering curriculum," *Journal of engineering Education*, vol. 86, no. 2, pp. 103–112, 1997.

[6] M. Casillo, F. Colace, L. Fabbri, M. Lombardi, A. Romano, and D. Santaniello, "Chatbot in industry 4.0: An approach for training new employees," in *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 2020, pp. 371–376.

[7] G. Rehe, R. Schneider, and D. D. Thimm, "Transformation towards Human Workforce 4.0 through the Qualification Model LTA-FIT," in *Proceedings of the 53rd HCISS Conference*, no. January, 2020. [Online]. Available: https://hdl.handle.net/10125/64305

[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[9] A.-L. Andersen, T. D. Brunoe, and K. Nielsen, "Engineering education in changeable and reconfigurable manufacturing: Using problem-based learning in a learning factory environment," *Procedia Cirp*, vol. 81, pp. 7–12, 2019.

[10] W. Zhang, W. Cai, J. Min, J. Fleischer, C. Ehrmann, C. Prinz, and D. Kreimeier, "5g and ai technology application in the amtc learning factory," *Procedia manufacturing*, vol. 45, pp. 66–71, 2020.

[11] A. da Silva Barbosa, F. P. Silva, L. R. dos Santos Crestani, and R. B. Otto, "Virtual assistant to real time training on industrial environment," in *Transdisciplinary Engineering Methods for Social Innovation of Industry 4.0: Proceedings of the 25th ISPE Inc. International Conference on Transdisciplinary Engineering, July 3–6, 2018*, vol. 7. IOS Press, 2018, p. 33.

[12] M. Zimmer, A. Al-Yacoub, P. Ferreira, and N. Lohse, "Towards human-chatbot interaction: a virtual assistant for the ramp-up process," Apr 2020. [Online]. Available: https://hdl.handle.net/2134/12091137.v1

[13] S. Maksymova, R. Matarneh, V. Lyashenko, and N. Belova, "Voice control for an industrial robot as a combination of various robotic assembly process models," *Journal of Computer and Communications*, 2017.

[14] M. C. Bingol and O. Aydogmus, "Performing predefined tasks using the human–robot interaction on speech recognition for an industrial robot," *Engineering Applications of Artificial Intelligence*, vol. 95, p. 103903, 2020.

[15] C. Li and H. J. Yang, "Bot-x: An ai-based virtual assistant for intelligent manufacturing," *Multiagent and Grid Systems*, vol. 17, no. 1, pp. 1–14, 2021.

[16] C. Li, J. Park, H. Kim, and D. Chrysostomou, "How can i help you? an intelligent virtual assistant for industrial robots," ser. HRI '21 Companion. New York, NY, USA: Association for Computing Machinery, 2021, p. 220–224. [Online]. Available: https://doi.org/10.1145/3434074.3447163

[17] D. Jurafsky and J. H. Martin, *Speech and Language Processing (2nd Edition)*. USA: Prentice-Hall, Inc., 2009.

[18] Z. Yu, Z. Xu, A. W. Black, and A. Rudnicky, "Strategy and policy learning for non-task-oriented conversational systems," in *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Los Angeles: Association for Computational Linguistics, Sept. 2016, pp. 404–412. [Online]. Available: https://www.aclweb.org/anthology/W16-3649

[19] R. E. Andersen, E. B. Hansen, D. Cerny, S. Madsen, B. Pulendralingam, S. Bøgh, and D. Chrysostomou, "Integration of a skill-based collaborative mobile robot in a smart cyber-physical environment," *Procedia Manufacturing*, vol. 11, pp. 114–123, 2017.

[20] A. S. Olesen, B. B. Gergaly, E. A. Ryberg, M. R. Thomsen, and D. Chrysostomou, "A collaborative robot cell for random bin-picking based on deep learning policies and a multi-gripper switching strategy," *Procedia Manufacturing*, vol. 51, pp. 3–10, 2020.

[21] B. Bejczy, R. Bozyil, E. Vaičekauskas, S. B. K. Petersen, S. Bøgh, S. S. Hjorth, and E. B. Hansen, "Mixed reality interface for improving mobile manipulator teleoperation in contamination critical applications," *Procedia Manufacturing*, vol. 51, pp. 620–626, 2020.

[22] M. Nardello, O. Madsen, and C. Møller, "The smart production laboratory: A learning factory for industry 4.0 concepts," in *CEUR Workshop Proceedings*, vol. 1898. CEUR Workshop Proceedings, 2017.