



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *Workshop on Semantic Scene Understanding for Human-Robot Interaction, ACM/IEEE International Conference on Human Robot Interaction*.

Citation for the original published paper:

Bartoli, E., Dogan, F I., Leite, I. (2023)

Contextualized Knowledge Graph Embeddings for Activity Prediction in Service Robotics

In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-324724>

Contextualized Knowledge Graph Embeddings for Activity Prediction in Service Robotics

Ermanno Bartoli¹

Fethiye Irmak Doğan¹

Iolanda Leite¹

Abstract—With the increasing demand for service robots in various environments, it is crucial to develop intelligent robots that can learn and adapt to changing environments. In this paper, we propose an approach that combines contextualized knowledge graph embeddings and continual learning to improve the ability of service robots to assist humans with various tasks. Specifically, we focus on the scenario of a service robot deployed in an office environment to assist employees with scheduling meetings, fetching documents, providing directions to conference rooms, and performing other tasks. By using contextualized knowledge graph embeddings, the robot can understand the current arrangement of entities and connections in a specific environment at a certain point in time, and predict the most appropriate actions to do based on previous actions and the current state of the environment. Continual learning allows the robot to learn and adapt to new activities and changing contexts without forgetting previous knowledge. We also discuss how the data can be collected for the continual learning process. Our proposed approach has the potential to significantly improve the efficiency and effectiveness of service robots in various environments.

I. INTRODUCTION

Modeling the environment is the ability to create a representation of the scene in which robots are placed and extract semantic information from it. This enables robots to comprehend and interact with the surrounding world. However, one of the main challenges in creating these representations is when they include people and their activities, which can be constantly changing as their effect on the environment. In such cases, the robot’s ability to accurately perceive and understand the environment becomes more complex. Although previous research has addressed some related problems concerning scene modeling [6] and prediction [8], little work has been done when these previous assumptions are considered. This paper proposes a novel approach suitable for several applications, with a focus on service robots in office environments. The goal is to assist employees with various tasks such as scheduling meetings, fetching documents, or providing directions to conference rooms. However, the robot needs to perform different activities depending on the dynamic context of the environment. We propose to use contextualized knowledge graph embeddings and continual learning to enable the robot to learn and adapt to new activities and changing contexts. For instance, in a meeting room scenario, the robot’s actions would vary depending on whether it’s empty or occupied.

It can clean the room if it’s empty or deliver refreshments and retrieve forgotten items for meeting participants if it’s occupied. The robot’s ability to learn patterns and regularities in the environment, such as the meeting room being usually empty at a certain time, allows it to plan and execute tasks more efficiently. For example, it can schedule cleaning tasks during lunchtime without disrupting any ongoing meetings or activities. Continual learning improves a robot’s ability to assist employees efficiently in dynamic workplaces by allowing it to learn incrementally without forgetting, enabling it to work in different environments.

II. RELATED WORK

The analysis of the related work is focused on the following two research areas: Scene Modeling and Prediction; and Continual Learning. Although these fields have been studied separately, it’s still challenging and novel to connect them. A robust approach for intelligent systems to operate effectively in complex environments. Combined models can accurately predict actions and adapt to new experiences without forgetting previous knowledge.

A. Scene Modeling and Prediction

Semantic Reasoning is a research area focused on developing methods for encoding and utilizing semantic knowledge, including concepts, facts, ideas, and beliefs about the world. Scene modeling seeks to enable robots to perceive and reason over the objects present in their environments [4]. Various approaches have been proposed to solve this problem, including probabilistic models like Bayesian Networks [8, 10], ontology-based approaches [9], and Boltzmann Machines (BM) [3, 6]. Utilizing Knowledge Graph Embeddings (KGE) is a distinct approach to scene modeling, which has proven to be effective in addressing various downstream tasks, such as question answering [2] and relation extraction. However, KGE models trained on static Knowledge Bases are unable to reason about contextualized events or facts, limiting their effectiveness. Researchers have explored Contextualized KGE to address this limitation, as exemplified in [11].

B. Continual Learning

The problem of catastrophic forgetting, referring to the loss or disruption of previously learned information when new information is acquired, has motivated the trend of Continual Learning (CL) in machine learning. CL aims to develop techniques that enable machine learning models to learn

¹ Ermanno Bartoli, Fethiye Irmak Doğan and Iolanda Leite are with the Division of Robotics, Perception and Learning at KTH Royal Institute of Technology, Stockholm, Sweden {bartoli, fidogan, iolanda}@kth.se

incrementally from new data while retaining knowledge of previously learned information. Little research has investigated the connection between CL and Knowledge Graph Embedding (KGE). A recent attempt to bridge this gap can be found in [5], where CL techniques have been developed specifically for KGE models. The proposed approach aims to reduce the impact of catastrophic forgetting in KGE by continuously learning from new information without losing the knowledge already acquired. The use of CL techniques in KGE can lead to more accurate and robust models that can learn from new data over time, leading to improved performance and generalization, particularly in lifelong learning scenarios. The applicability of CL is broad, including scenarios involving large and evolving knowledge graphs. Balancing the trade-off between learning new information and retaining existing knowledge, CL can lead to improved performance and more efficient use of resources. Although a first attempt on addressing this problem has been proposed in [1], contextualized KGE has not been taken into consideration.

III. PROBLEM DEFINITION

In this work, we propose to address the previously mentioned gap in scene modeling and prediction by using Contextualized Knowledge Graph Embeddings (CKGE) in conjunction with Continual Learning (CL) techniques. Our approach aims to enable a social robot to perceive and reason over entities and relations in a changing environment by incrementally learning and updating its knowledge representation without forgetting previous information, in order to make them able to choose the most appropriate actions to execute.

To achieve this goal, we plan to develop a CKGE model that can take into account the context of the scene in which the robot is situated. Specifically, we aim to train our model to encode not only the objects and relations in the scene but also the context in which they appear, such as the time, location, and previous events. We believe that this could enable our model to reason more effectively about the scene and make more accurate predictions about the next stages of the environment and therefore how to assist users in these dynamic environments.

Furthermore, we plan to incorporate CL techniques into our approach to allow our model to learn from new scenes without forgetting previously learned knowledge. Specifically, algorithms such as Elastic Weight Consolidation (EWC) or other regularization techniques [7] could be used to selectively protect important weights in our model while allowing others to change. By doing so, we hope to maintain the accuracy of our model while adapting to new scenes.

In the context of evaluating a reasoning model's predictions, user feedback serves as a crucial component. Furthermore, users can aid the robot in minimizing the uncertainty surrounding the extraction of context, thereby enhancing the accuracy of its predictions.

To evaluate the effectiveness of our approach, we plan to conduct experiments using a publicly available dataset of scenes and objects. We will compare the performance of

our CKGE model with and without CL techniques to other state-of-the-art approaches in scene modeling and prediction. Additionally, we will perform ablation studies to analyze the contribution of each component in our approach. Overall, our approach aims to contribute to the development of robots that can perceive, reason, and interact more effectively with humans in a changing environment.

IV. DATA ACQUISITION

Deploying a robot in a real-world office environment to collect data is slow and resource-intensive. To speed up the process, simulations or publicly available datasets can be used.

Simulations are a useful way to generate large quantities of diverse data and create complex scenarios that may be difficult to recreate in the real world. This can be especially useful for testing the robot's behavior in emergency situations or rare events, which may not frequently occur in a real-world office environment. Additionally, simulations provide a safe and controlled environment for testing the robot's behavior without risking damage to the robot or its surroundings.

Publicly available datasets can also be used to augment the data acquired through simulations and real-world deployments. These datasets provide a diverse range of scenarios and can help to overcome the problem of data scarcity that is often encountered in machine learning applications. However, it should be noted that there may be limitations to these datasets, especially if they are not activity-focused.

Crowdsourcing can also be used to collect data from a wide range of users in different environments and scenarios. This can be especially useful for collecting data on specific activities or tasks that may be difficult to simulate or obtain through real-world deployments. Crowdsourcing can also provide a diverse range of data and enable the collection of feedback from a large number of users. A combination of these approaches may be necessary to build a service robot that can adapt to a wide range of environments and scenarios while addressing the limitations of each approach.

V. CONCLUSION AND FUTURE WORKS

In this paper, we proposed a novel approach using contextualized knowledge graph embeddings to improve service robots' performance in office environments, we discussed the challenges faced by current service robots and how our approach can address them. Specifically, our approach enables robots to understand the current context and predict the next activity based on previous actions and the current state of the environment. By incorporating continual learning techniques, our approach can adapt to new scenes without forgetting previously learned knowledge. We expect that our approach will result in more accurate and efficient service robots, leading to increased productivity and improved user experience in office environments.

VI. ACKNOWLEDGMENT

This work was supported by the PerCorSo project funded Knut and Alice Wallenberg Foundation.

REFERENCES

- [1] Ermanno Bartoli, Francesco Argenziano, V Suriani, and D Nardi. “Knowledge Acquisition and Completion for Long-Term Human-Robot Interactions using Knowledge Graph Embedding”. In: *arXiv preprint arXiv:2301.06834* (2023).
- [2] Antoine Bordes, Sumit Chopra, and Jason Weston. “Question answering with subgraph embeddings”. In: *arXiv preprint arXiv:1406.3676* (2014).
- [3] Hande Celikkanat, Güner Orhan, Nicolas Pugeault, Frank Guerin, Erol Şahin, and Sinan Kalkan. “Learning context on a humanoid robot using incremental latent dirichlet allocation”. In: *IEEE Transactions on Cognitive and Developmental Systems* 8.1 (2015), pp. 42–59.
- [4] Kang Chen, Yu-Kun Lai, and Shi-Min Hu. “3D indoor scene modeling from RGB-D data: a survey”. In: *Computational Visual Media* 1 (2015), pp. 267–278.
- [5] Angel Daruna, Mehul Gupta, Mohan Sridharan, and Sonia Chernova. “Continual learning of knowledge graph embeddings”. In: *IEEE Robotics and Automation Letters* 6.2 (2021), pp. 1128–1135.
- [6] Fethiye Irmak Doğan, Hande Celikkanat, and Sinan Kalkan. “A deep incremental boltzmann machine for modeling context in robots”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 2411–2416.
- [7] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. “Overcoming catastrophic forgetting in neural networks”. In: *Proceedings of the national academy of sciences* 114.13 (2017), pp. 3521–3526.
- [8] Yaser Sheikh and Mubarak Shah. “Bayesian modeling of dynamic scenes for object detection”. In: *IEEE transactions on pattern analysis and machine intelligence* 27.11 (2005), pp. 1778–1792.
- [9] Moritz Tenorth and Michael Beetz. “KnowRob—knowledge processing for autonomous personal robots”. In: *2009 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 2009, pp. 4261–4266.
- [10] Jonas Tjomsland, Sinan Kalkan, and Hatice Gunes. “Mind your manners! a dataset and a continual learning approach for assessing social appropriateness of robot actions”. In: *Frontiers in Robotics and AI* 9 (2022), p. 4.
- [11] Quan Wang, Pingping Huang, Haifeng Wang, Songtai Dai, Wenbin Jiang, Jing Liu, Yajuan Lyu, Yong Zhu, and Hua Wu. “Coke: Contextualized knowledge graph embedding”. In: *arXiv preprint arXiv:1911.02168* (2019).