Simpler rather than Challenging: Design of Non-Dyadic Human-Robot Collaboration to Mediate Human-Human Concurrent Tasks

Extended Abstract

Francesco Semeraro
Manchester Centre for Robotics and AI, The University of Manchester
Manchester, United Kingdom
francesco.semeraro@manchester.ac.uk

Jon Carberry
BAE Systems (Operations) Ltd.
Warton, United Kingdom
jon.carberry@baesystems.com

Angelo Cangelosi
Manchester Centre for Robotics and AI, The University of Manchester
Manchester, United Kingdom
angelo.cangelosi@manchester.ac.uk

ABSTRACT

Human-robot interaction (HRI) is progressively addressing multi-party scenarios, where a robot interacts with more than one human user at the same time. Conversely, research in this area is still at an early stage for human-robot collaboration (HRC). The intervention of a robot in human collaboration could be helpful to handle mutual disturbances of workers operating at the same time on the same target object. Therefore, this work outlines design methodologies of non-dyadic human-robot collaborations to address concurrent human-human tasks in manufacturing applications. After this, preliminary results regarding a robotic agent's high-level understanding of such scenarios realised through a variational autoencoder trained by means of transfer learning are shown.

KEYWORDS

Non-Dyadic Human-Robot Collaboration; Multi-Party Human-Robot Collaboration; Concurrent Tasks; Multi-User Activity Recognition; Deep Learning; Variational Autoencoder; Transfer Learning

ACM Reference Format:
Francesco Semeraro, Jon Carberry, and Angelo Cangelosi. 2023. Simpler rather than Challenging: Design of Non-Dyadic Human-Robot Collaboration to Mediate Human-Human Concurrent Tasks: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Human-robot interaction research is constantly progressing [1]. Researchers are increasingly addressing scenarios that go beyond the more commonly explored dyadic interaction between a human and a robot [6, 10]. There are works in HRI regarding non-dyadic interactions [12], but they are still outnumbered by the plethora of works related to dyadic HRI. Furthermore, fewer works tackle a human-robot collaboration application in a multi-party scenario (see Figure 1) [11, 12, 20], making this a promising field for scientific investigation. Therefore, this work outlines a collaborative scenario in which the robot works jointly with two human users who perform concurrent tasks. After having contextualised this type of tasks, a case scenario in manufacturing applications is described.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom. © 2023 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

Figure 1: A collaborative robot is assisting two human users working concurrently on a manufacturing object (adapted from [9]).

Besides, by adopting a high-level understanding model of an intelligent agent in such a scenario by means of deep learning models [15], this work demonstrates that it is possible to train such models by using datapoints related to single users to be able to make inference about pairs of users present in the scene at the same time.

2 DESIGN OF NON-DYADIC HRC FOR CONCURRENT TASKS

In collaborative tasks, two agents cooperate to achieve a common goal. There are different possible ways in which this collaboration can unfold. In turn-taking collaborations, the agents mutually switch between activity and inactivity while working in the shared workspace (e.g., collaborative assembly) [16]. In joint tasks, they must act at the same time, due to the constraints of the task itself (e.g., joint load lifting) [8]. However, there are concurrent tasks in which the two agents need to complete a common goal together, but jointly performing different operations, consequently causing mutual disturbances (e.g., tactical vehicle control) [7].

The collaborations described so far consider only two agents in the scene. The inclusion of a third agent is justified if it performs a role that differs from that of the others. If it acts as an entity with its own sub-task towards the same common goal, thus interfering with the other agents, it could be collapsed into one of the agents [19]. Indeed, in this case, from the perception of one of the agents, it could be considered as a component of its counterpart in the collaboration. A third agent could, instead, perform the role of a mediator in the collaboration. Through its intervention in the scene, it would not act on the final goal directly; rather, it would indirectly assist in the collaboration by mitigating the disturbances the two working agents may cause to each other, to improve their joint performance. The other agents would both perceive the third one not as an extension of the counterpart, but as a different standalone agent that supports their actions. This case is the one considered in this work and is further described below in concurrent tasks related to a manufacturing scenario, without excluding other potential applications, such as social robotic interactions [2, 3, 14].

In some production chains, one worker at a time works on the object of manufacturing. Therefore, the productivity of the process would be enhanced by enabling two workers to perform manufacturing operations on a target object at the same time. However, by doing so, there is a chance that workers may interfere with each other while performing operations on the object (e.g., by changing the position and orientation of the object to have it in an optimal configuration for the operation to perform). To simplify this rather challenging scenario, in terms of cognitive load for the workers, a robot (e.g., a robotic manipulator endowed with vision sensors) could act as an intermediary between the two users (see Figure 1). In this way, the workers do not need to interact with each other while working, as queries regarding the state of the target object (e.g., rearrangement of object's orientation) can be handled by the robot (see Figure 2) [11]. Because of this, team collaboration is replaced by the robot's mediation and the interaction and consequent disturbances between the users are reduced [11].

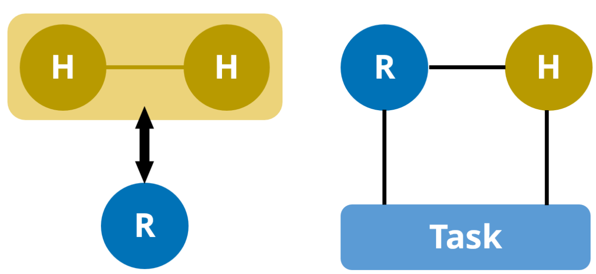

Figure 2: Concurrent non-dyadic human-robot collaboration. Left: Representation in terms of multiplicity [19]. Right: Representation of the roles of humans and robot, in relation to the task and each other [17]. Here, we consider the humans with an active role, while the robot with an adaptive one.

To pinpoint which type of activities would be crucial to recognise in such a scene, it is important to further delineate the role of the robot in the collaboration. More precisely, it needs to know how to handle resources of the setup (e.g., object of manufacturing) between the two users. Additionally, the robot needs to understand which user requires its attention. When the users are both working on the target object, the robot needs to address requirements from both of them. When one of them finishes a sub-step of the manufacturing and gets ready for the next one, the robot can shift all its resources to the user who is currently working and improve their performance. Finally, a user can query the robot's attention through a signal (e.g., specific hand gesture) for the robot to be aware that resources must be reallocated. The description of the unfolding of this collaboration allows us to depict three main activities the robot needs to detect from a single user:

- working on the object of manufacturing;
- preparing for the next task;
- requesting a reallocation of resources from the robot.

Consequently, in the described scenario, a robotic agent should be sensitive to 9 possible cases, related to the possible pairings of the 3 states related to a single user.

3 EXPERIMENTS AND RESULTS

Based upon the case scenario previously described (see Section 2), the possibility of using data recordings related to single subjects to then infer a group activity in this type of scenarios was investigated [13]. Sequences of 3D skeletal poses of single participants performing the three aforementioned states were collected and paired in post-processing. This was done to obtain a training dataset that was relatable to the actual situation of having two people performing joint activities in the scene at the same time. This dataset was used

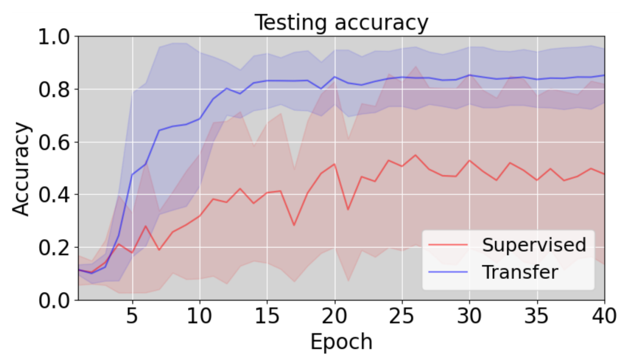

Figure 3: Comparison between supervised and transfer learning during the testing phase.

to train a variational autoencoder (VAE) [4, 5], with spatio-temporal graph convolutional networks (STGCNs) [18] in the encoder, to recognise the nine group activities previously mentioned (see Section 2) by means of transfer learning. Specifically, the VAE was exposed to the training data in an unsupervised way, and then its STGCN layers were appended to a softmax layer to produce the output label [13]. This approach was compared against directly training the same STGCN layers in a supervised way.

The strategy based on transfer learning led to better and more stable accuracy than training the layers on their own (mean: 0.511 and SD: 0.361 for the supervised learning, mean: 0.864 and SD: 0.11 for the transfer learning, see Figure 3). Similarly to this, testing results with datapoints coming from a related multi-user scenario [13] are consistent (mean: 0.337 and SD: 0.0916 for supervised learning, mean: 0.616 and SD: 0.0288 for transfer learning).

Future works regarding this topic are further testing of the proposed inference methodology with different types of deep neural networks and the realisation of an experimental validation in which a robot uses such inference system to handle concurrent tasks between two human users.

ACKNOWLEDGMENTS

Francesco Semeraro's work was supported by the DTP CASE Conversion award "Human-Robot Collaboration for Flexible Manufacturing", funded by UKRI Engineering and Physical Sciences Research Council and BAE Systems Plc. (Ref. 2480772). Angelo Cangelosi's work was partially supported by the H2020 project TRAINCREASE (Ref. 210638710), the UKRI Trustworthy Autonomous Systems Node on Trust (Ref. EP/V026682/1) and the US Air Force project THRIVE++ (Ref. FA9550-19-1-7002).

REFERENCES

[1] Arash Ajoudani, Andrea Maria Zanchettin, Serena Ivaldi, Alin Albu-Schäffer, Kazuhiro Kosuge, and Oussama Khatib. 2018. Progress and prospects of the human–robot collaboration. *Autonomous Robots* 42, 5 (2018), 957–975.

[2] Filippo Cavallo, Francesco Semeraro, Gianmaria Mancioppi, Stefano Betti, and Laura Fiorini. 2021. Mood classification through physiological parameters. *Journal of Ambient Intelligence and Humanized Computing* 12, 4 (2021), 4471–4484.

[3] Laura Fiorini, Francesco Semeraro, Gianmaria Mancioppi, Stefano Betti, Luca Santarelli, and Filippo Cavallo. 2018. Physiological sensor system for the detection of human moods towards internet of robotic things applications. In *New Trends in Intelligent Software Methodologies, Tools and Techniques*. IOS Press, 967–980.

[4] Ali Ghadirzadeh, Xi Chen, Wenjie Yin, Zhengrong Yi, Mårten Björkman, and Danica Kragic. 2020. Human-centered collaborative robots with deep reinforcement learning. *IEEE Robotics and Automation Letters* 6, 2 (2020), 566–571.

[5] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. 2013. Stochastic variational inference. *Journal of Machine Learning Research* (2013).

[6] Eva Hornecker, Antonia Krummheuer, Andreas Bischof, and Matthias Rehm. 2022. Beyond dyadic HRI: building robots for society. *interactions* 29, 3 (2022), 48–53.

[7] William J Horrey and Daniel J Simons. 2007. Examining cognitive interference and adaptive safety behaviours in tactical vehicle control. *Ergonomics* 50, 8 (2007), 1340–1350.

[8] Loris Roveda, Jeyhoon Maskani, Paolo Franceschi, Arash Abdi, Francesco Braghin, Lorenzo Molinari Tosatti, and Nicola Pedrocchi. 2020. Model-based reinforcement learning variable impedance control for human-robot collaboration. *Journal of Intelligent & Robotic Systems* 100, 2 (2020), 417–433.

[9] Corey Ryan. 2020. The Fine Line between Industrial and Collaborative Robots: It's Smaller than You Think. Article. Retrieved October 5, 2022 from https://www.qualitymag.com/articles/96119-the-fine-line-between-industrial-and-collaborative-robots-its-smaller-than-you-think

[10] Eike Schneiders. 2022. Non-Dyadic Human-Robot Interaction: Concepts and Interaction Techniques. In *HRI'22: ACM/IEEE International Conference on Human-Robot Interaction*. IEEE Press, 1176–1178.

[11] Eike Schneiders and Stanley Celestin. 2022. Non-Dyadic Entrainment for Industrial Tasks. In *Workshop on Joint Action, Adaptation, and Entrainment in Human-Robot Interaction*.

[12] Eike Schneiders, EunJeong Cheon, Jesper Kjeldskov, Matthias Rehm, and Mikael B Skov. 2022. Non-dyadic interaction: A literature review of 15 years of human-robot interaction conference publications. *ACM Transactions on Human-Robot Interaction (THRI)* 11, 2 (2022), 1–32.

[13] Francesco Semeraro, Jon Carberry, and Angelo Cangelosi. 2023. Towards Multi-User Activity Recognition through Facilitated Training Data and Deep Learning for Human-Robot Collaboration Applications. https://doi.org/10.48550/ARXIV.2302.05763

[14] Francesco Semeraro, Laura Fiorini, Stefano Betti, Gianmaria Mancioppi, Luca Santarelli, and Filippo Cavallo. 2018. Physiological wireless sensor network for the detection of human moods to enhance human-robot interaction. In *Italian Forum of Ambient Assisted Living*. Springer, 361–376.

[15] Francesco Semeraro, Alexander Griffiths, and Angelo Cangelosi. 2023. Human–robot collaboration and machine learning: A systematic review of recent research. *Robotics and Computer-Integrated Manufacturing* 79 (2023), 102432.

[16] Samuele Vinanzi, Angelo Cangelosi, and Christian Goerick. 2020. The Role of Social Cues for Goal Disambiguation in Human-Robot Cooperation. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 971–977.

[17] Xi Vincent Wang, Zsolt Kemény, József Váncza, and Lihui Wang. 2017. Human–robot collaborative assembly in cyber-physical production: Classification framework and implementation. *CIRP annals* 66, 1 (2017), 5–8.

[18] Sijie Yan, Yuanjun Xiong, and Dahua Lin. 2018. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Thirty-second AAAI conference on artificial intelligence*.

[19] Holly A Yanco and Jill Drury. 2004. Classifying human-robot interaction: an updated taxonomy. In *2004 IEEE international conference on systems, man and cybernetics (IEEE Cat. No. 04CH37583)*, Vol. 3. IEEE, 2841–2846.

[20] Mohammad Samin Yasar and Tariq Iqbal. 2021. A scalable approach to predict multi-agent motion for human-robot collaboration. *IEEE Robotics and Automation Letters* 6, 2 (2021), 1686–1693.