Developing a Multiple-EDT-Supervision Interface

max kullmann[1], Daniel Pietschmann[2], Benny Liebold[2], Maximilian Eibl[3], Lewis L. Chuang[1]

[1] *Humans and Technology, Chemnitz University of Technology, Germany*
[2] *Institute for Media Research, Chemnitz University of Technology, Germany*
[3] *Media Informatics, Chemnitz University of Technology, Germany*

Abstract

With increasing automation beyond a professional level and into private and public life, one can assume that the ironies of automation explained by Bainbridge have been experienced by most of the members of our hybrid society – one where interaction between humans and embodied digital technology (EDT) becomes the agenda. In order to realise the full potential of hybrid teams we propose taking a new perspective, where EDTs do not operate on a fixed level of automation but become cognisant of their supervisor's capabilities through a shared interface. This enables mutual perception and thus the dynamical adjustment between joint or autonomous activity.

Keywords

Automation supervision, human-automation interaction, embodied digital technology, user interface design

1. Introduction

A world of ubiquitous embodied digital technology (EDT) is about to become a reality both in our private and public realms (e.g., cleaning robots, delivery drones, and so forth). The agenda is for this technology to operate with a fixed level of automation and for the limits to be compensated by control policies, i.e., what cannot be automated is left to humans – a paradigm that brought about the ironies of automation [1]. This approach is contrary to the idea of true collaboration in hybrid teams of diverse autonomous agents. We propose to make a step back and change perspectives: A supervisor of multiple EDT units should not be waiting for one to reach its limitations and be allocated the actual work. Instead, available capacity should be indicated to EDTs in order for such to make use of joint resources, potentially resulting in higher performance than in a mode where a higher level of autonomy is required from EDTs, which is why they then operate exclusively in safe waters.

In a one-supervisor-to-many-EDTs scenario, this not only necessitates to know (and let others know) the state of the supervisor but to know the state and (future) requirements of each EDT and communicate this information to the supervisor. Hence, a bidirectional communication channel must be established. We propose a front end interface that combines both the information on the psychological state of a supervisor (through physiological measures) as well as information on technical necessities from the individual technoid agents.

1.1. Sensing Agents

Prior to going into detail about the implementation of the front end, the back end is explained here briefly. Basically it is supposed to contribute information on the human user and of each EDT. A user study has been designed with the aim of identifying reliable implicit indicators of the human supervisor's readiness to switch supervision from one EDT to another. Traditional physiological indices of attention and orienting are being considered [3]. The first step is to manipulate the number of EDTs (2, 3, or 4) and see how this corresponds with changes in phasic activity (e.g., skin conductance responses; SCR) as an estimation of the strength of an orienting reflex. Tonic responses (e.g. skin conductance level; SCL) are used to explore effects of e.g. fatigue over time. Explicit indicators can be derived from the user's interaction with the front end. Further, to optimally re-direct human attention capacities and establish signal cues in

AutomationXP23: Intervening, Teaming, Delegating - Creating Engaging Automation Experiences, April 23rd, Hamburg, Germany
EMAIL: maximilian.kullmann@phil.tu-chemnitz.de (A. 1); daniel.pietschmann@phil.tu-chemnitz.de (A. 2); benny.liebold@iuz.tu-chemnitz.de (A. 3) ; eibl@cs.tu-chemnitz.de (A. 4) ; lewis.chuang@phil.tu-chemnitz.de (A. 5)
ORCID: 0009-0003-9473-2551 (A. 1); 0000-0001-9174-2234 (A. 2); 0000-0003-2424-730X (A. 3); 0000-0002-9519-2708 (A. 4); 0000-0002-1975-5716 (A. 5)

the interface one must be aware of the state an EDT is in.

1.2. Back end

A computational model – more specifically a baseline model using a recurrent neural network (RNN) based on long short-term memory (LSTM) or gated recurrent units (GRU) – for co-adaptive levels of autonomy that integrates the information on the supervisor's availability and capacity for reorienting attention and intervention as well as the operational capacity of the respective EDTs is imaginable [7] [5]. It serves as the backend to dynamically modify the EDTs' control policies as well as levels of autonomy and to signal the requirement of cues to optimally re-direct human attention capacities.

2. A Co-Adaptive Interface

For now let's assume appropriate data is available from the back end. The aim of developing a co-adaptive interface for human-automation collaboration is to get away from the misconception of humans as fallbacks for automation failure and instead work towards the idea of a platform for humans and EDTs, where both can demonstrate and make use of their current capabilities. The envisaged scenario is a one-to-many supervisory relation; namely one human is responsible for the supervision of a group of EDTs (the effects of different group sizes is to be evaluated). The interface serves to provide a global overview of active EDTs and their current status, i.e., the process step it is currently in, expressed in a universal (multimodal) way that is independent from the various tasks of individual EDTs. This allows to prioritise which unit needs support most urgently on an independent basis. Additionally individual rating can be applied. Details on sensor data and checklists of subtasks as well as explicit requests are forwarded to the interface and can be accessed easily. The utility to visualise the physiological data of the supervisor as biofeedback is to be evaluated.

2.1. Nature of the Interface

Teleoperation of vehicles dates back to the early 20th century, but it was not widely used until 1970's and today it is common for ground, underwater, air, and space vehicles. Fong and Thorpe divide teleoperator interfaces into four categories: direct, multimodal/multisensor, supervisory control, and novel, where the latter is admittedly very relative and includes web based interfaces but also hands-free controls via brainwaves or gestures [4]. The challenge in the design of the interface in our scenario is the diverse nature of the group of EDTs and hence the same accounts for the interface itself. Therefore it makes sense to split the interface in a macro and several micro levels. The macro level will be something like a dashboard. However, this must not mean that a desktop application is the only way to go. In an ideal world supervision is not a full-time job but a task among many, which one can than either actively allocate dedicated time slots or be on-call. In the former case the interface should be integrable into individual workflows. In the latter case the interface should be integrable into everyday life and eventually vanish until it is needed. Only then you go into active mode, where the access point is the dashboard. From here there is the possibility to zoom in on an EDT's local perspective to gain or maintain situational awareness. So both the macro level (i.e. group) and the micro level (i.e. individual) are observable for the supervisor to anticipate events that may have a fatal impact on performance.

2.2. Mutual Perception

This work will extend an EDT's ability to perceive its immediate physical environment to include an assessment of a supervisor's dynamic resources. As such it becomes cognisant to a supervisor's current accountability for each EDTs' behaviour, and in response adapts its LoA. Let's go through a possible scenario where a supervisor announces (implicitly or explicitly) free resources to its supervisees: The unit that is at the top of the priority list requests focal attention. As soon as this is ensured (by measurements and/or manually) the unit adjusts its control policy in such a way that it now carries out tasks that afford e.g. readiness to take over control. Depending on the exact actions the supervisor could enter the micro level of that unit. The more diverse the group of EDTs under supervision is, the more mixed the interface gets in respect of the aforementioned categories. A careful design is necessary for the principles of different micro "worlds" to at least resemble each other. A study of McGovern [6] identified shortcomings (difficulties to detect obstacles, loss of situational awareness) of direct or so-called inside-out controls via a simple monitor streaming the static view of a camera. We have been familiar with this limited perception at least since human-human interaction was reduced to screens due to the pandemic. On the contrary everyone who already experienced some remote environment through a tracked, stereoscopic system, e.g. a virtual environment via a head mounted display or the like, knows the phenomenon of immersion [2]. EDT which is equipped with depth cameras and lidar devices can provide not only a two but a three dimensional spacial representation of its surrounding. Not only does this improve spatial and hence situational awareness, but it also allows one to literally take different perspectives, e.g. to move around in the proximity of a remote agent.

3. Contribution

This project will provide answers to the fundamental requirements for co-adaptive levels of autonomy of EDTs in one-to-many supervision scenarios and implement them in simulated and physical demonstrators. This entails in particular:

- A concept for the human supervisor's resource allocation and its implementation in an iteratively designed user interface.
- An identified set of reliable and effective markers for estimating a human supervisor's availability for intervention as well as an open dataset containing the gathered data.
- A computational model of co-adaptive models of autonomy guiding EDT decision-making policies and the human supervisor's situational awareness.

We provide a perspective to increase the deployment of EDTs in hybrid societies by mitigating shortcomings of EDT automation and thus promoting smooth cooperation. The expected results allow for a new quality of one-to-many EDT supervision and are highly applicable in EDT design.

Acknowledgment

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – CRC 1410.

References

1. Bainbridge, L. Ironies of automation. Autom. 19 (1982): 775-779.
2. Berkman, M.I., Akan, E. Presence and Immersion in Virtual Reality (2019). In: Lee, N. (eds) Encyclopedia of Computer Graphics and Games. Springer, Cham. https://doi.org/10.1007/978-3-319-08234-9_162-1
3. Bradley, M. M. Natural selective attention: Orienting and emotion (2009)
4. Fong, T., Thorpe, C. Vehicle Teleoperation Interfaces. Autonomous Robots 11, 9–18 (2001). https://doi.org/10.1023/A:1011295826834
5. Liu, M., Xu, X., Hu, J., & Jiang, Q. Real time detection of driver fatigue based on CNN−LSTM. IET Image Processing (2022), 16(2), 576–595. https://doi.org/10.1049/ipr2.12373
6. McGovern, D. Experiences and results in teleoperation of land vehicles (1990). Sandia National Laboratories, Albuquerque, NM, Technical Report SAND 90-0299.
7. Wöllmer, M., Blaschke, C., Schindl T., Schuller, B., Färber, B., Mayer, S., Trefflich, B. Online Driver Distraction Detection Using Long Short-Term Memory. IEEE Transactions on Intelligent Transportation Systems 12 (2011): 574-582.