The Complex Human: How to Enhance a Control System ICRA 2022 Workshop *Perspectives on Achieving Research Convergence for Robotics-Enabled Future of Work*

Emily C. Collins[1[0000−0001−9396−536X]], David Cameron[2[0000−0001−8923−5591]], and Thomas L.M. Piercy[3[0000−0001−6769−6645]]

Institute of Experiential Robotics, Northeastern University, Boston, USA
e.collins@northeastern.edu
Information School, The University of Sheffield, Sheffield, UK
d.s.cameron@sheffield.ac.uk
Department of Electronic and Electrical Engineering, The University of Manchester, Manchester, UK
thomas.piercy@manchester.ac.uk

Abstract. An ideal control system designed to improve the human experience whilst performing a task with a robotic agent, is one that accounts for the human as a complex agent and active partner in the control system's objectives. However, how does one build the complexity of a human, and the infinite choices humans make, into such a richly targeted control system? One approach, here described, views the robot as the primary agent in the human-robot interaction providing the necessary inputs to the human to guide the task. Designing within a constrained set of conditions can enhance the human-robot collaboration, without the need to fully define all human decision possibilities. We here give a brief example to explain this way of accommodating for The Complex Human in a control system.

Keywords: Human Robot Interaction · Human-in-the-Loop · Control

1 Introduction

In its most basic form, a control loop is a cycle of measurement and adjustment that relies on a finite set of possible inputs, and outputs. These loops can be more complex, including multivariate inputs and outputs, hidden layers of processing, and learning cycles. Examples include, control loops used in extremely bounded conditions such as automated manufacturing, and more nuanced control loops such as those employed by stock market algorithms.

In situations where a human and a robot are involved in a shared activity, the ideal control model should be capable of allowing a robot to reliably and consistently act towards a human by including a model of a human that

provides reliable and consistent human-action estimates. Including such sophisticated models of humans in a control loop provides the most accurate and reliable control system for use in HRI as it can fully account for a human as a complex agent, and an active partner in the control system's objectives.

However, how does one build the complexity of a human, and the infinite choices humans make, into a control system? One approach is to not aspire to building a robot with artificial general intelligence, capable of human-level observational nuance. Rather, we propose an approach that is constrained by task parameters: high risk environments with unknown specifics, but controlled, known locations, which also maintain a human operator as the first and last actor in play. Here the robot is the primary agent in the interaction with the human, providing the needed inputs to support the control system's objectives.

2 The Complex Human

2.1 Defining human complexity

To build a complete model of a human into a system, we would need to account for all given permutations of a single observed action. Humans are adept at observing others' behaviours and extracting a variety of data to reach conclusions about the action and intention (e.g. see Theory of Mind [8]). This is often done adeptly despite conflicting intentions behind identically observed behaviours. For example, imagine two people (A & B) interacting independently with a dog, whose observed interactions appear identical: playing with it, laughing, stroking, etc. However, A and B have different intentions driving identical observable actions. A dislikes spending time with other humans and plays with the dog for companionship; conversely, B, a gregarious individual, plays with the dog because they enjoy any and all social interaction.

This simple example highlights the issue of translating human complexity to binary understanding. We must ask, which features would we extract from each observation such that a robot co-worker could discern the *true intent* of a human? Human observers, consciously or not, draw from a variety of data to infer the true intentions behind observed actions. This 'mind-reading' allows humans to appropriately engage with others by understanding what observed actions have revealed about individual personalities [1].

2.2 Modelling human complexity

Bounding prediction models to specific contexts is one way to maintain the richness of human action, within the current limited ability of robotics. For example, human driving actions can be accurately categorised to the extent that, given a sequence of control steps around steering, acceleration and breaking patterns, a driver's intended action within a set-boundary can be classified [2]. However, extending the automatic recognition of people's intended action from classical methods (e.g. Markov Dynamic Models as in [2]) cannot account for

behaviour that is *not obvious*. Although humans are not consistently random, and are diverse in their knowledge and abilities (which can thus be tracked and categorised), they *are* both rational and emotional in their decision-making [3]. The psychology and root of emotion is so complex that to reduce emotional decision-making to prediction trees is incalculable given current robotics.

2.3 Ethical Considerations

i) Modelling a human as a system with simplified inputs and outputs is necessarily reductionist and could be reminiscent of Taylorist work environments [4]. Moreover, reducing people to characteristics necessitates stereotyping, leading to racism and discrimination [5].

ii) If our goal is to make better predictive models of human behaviour, are we then trying to encourage predictable human behaviour? This could lead to deception and coercion on the part of the robot, in its attempts to elicit actions required for its goal. Is this a 'good' or a 'bad' thing [6]? Especially if influencing a human to act a certain way is done for the benefit of task completion, e.g., by reducing cognitive load for the user.

Building better control models must be mindful of these issues.

3 Example: A Human-Accommodating Control System

In nuclear decommissioning robotics, a human-robot team is deployed in a high risk environment with unknown specifics, but which exists in a controlled operational area, and must maintain a human operator as the first and last actor in play. Due to the risks associated with catastrophic failure, and the complexity of the tasks, nuclear decommissioning is unsuitable for full-automation. Operator actions are a relatively simple starting point from where to consider ways to monitor or influence the system to reduce risk. Alarm fatigue is a good example of one such risk. Too-frequent alarms have been implicated in UAV mishaps, as alert saturation results in human complacency, an unexpected, undesirable consequence which the system fails to factor for [7].

State of the Art robotics can factor for this via the introduction of a programmed agent capable of monitoring the alerts and messages displayed to the operator, which then optimises the alerts and warnings that the operators receive to make them appear only when they can be addressed easily, and not when they may distract, resulting in reduced risk of human error. This proposed agent learns from the operator, noting: when alerts are looked at, when they are ignored until later, and in what circumstances they are ignored entirely, and uses this knowledge to intelligently provide alerts and warnings. In this example operator data is obtained from such as body tracking, eye tracking, and biomonitors, utilising supervised Machine Learning to build task classifiers for task estimation and task prediction. Operator attention is monitored to build a simple state-machine model of attention per task (e.g., with task-concentration being high, medium or low), and when paired with task estimation and prediction,

may also be used to reduce the risk of alarm fatigue, decrease the possibility of distractions, and increase the likelihood of alarms being addressed.

This is a small, but tangible step towards building the complexities of a human into a control system by using the robot as the primary agent providing necessary inputs to guide the human at their task.

Acknowledgements We would like to extend our appreciation to Prof Guido Herrmann, at The University of Manchester.

References

1. Collins, E.C., 2019. Drawing parallels in human–other interactions: a transdisciplinary approach to developing human–robot interaction methodologies. Philosophical Transactions of the Royal Society B, 374(1771), p.20180433.
2. Pentland, A., Liu, A.: Modeling and prediction of human behavior. Neural computation, 11(1), pp.229-242. (1999)
3. Kennedy W.G.: Modelling Human Behaviour in Agent-Based Models. In: Heppenstall A., Crooks A., See L., Batty M. (eds) Agent-Based Models of Geographical Systems. Springer, Dordrecht. (2012)
4. Taylor, F.W. (1911). The Principles of Scientific Management. NY: Dover Publications.
5. Dovidio, J.F. and Jones, J.M., 2019. Prejudice, stereotyping, and discrimination. Advanced social psychology: the state of the science, 2.
6. Collins, E.C. "Vulnerable users: deceptive robotics." Connection Science 29.3 (2017): 223-229.
7. Sibley, C., Coyne, J., Sherwood, S.: Research Considerations and Tools for Evaluating Human-Automation Interaction with Future Unmanned Systems. In Autonomy and Artificial Intelligence: A Threat or Savior? (pp. 157-178). Springer, Cham. (2017).
8. Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? The Behavioural and Brain Sciences, 4, 515–526.