ARTIFICIAL AGENTS IN SOCIO-TECHNICAL SYSTEM

Setareh Zafari and Sabine T. Koeszegi

E330 - Institute of Management Science at TU Wien

INTRODUCTION

Self-driving cars, smart ambient home systems, industrial and service robots are all examples of cognitive computing systems that are designed to recieve, analyze data and carry out actions without human intervention [1]. Designing these systems goes beyond the technical domain of simply being intelligent and connected, as their repercussions of the social, organizational and user issues become extremely difficult to ignore [2]. While technological advances have made it possible for the passive objects to become agentic, disagreements persist about the notion of machine agency. This conceptual paper provides an initial typology of artificial agents and reflects the socio-cultural consequences encountered in interacting with artificial agents.

FUNDAMENTAL OF THE PROBLEM

Current developments in the field of artificial intelligent suggest an increase in the agency of machines, with artificial agents taking over control that would have been in the hands of human agents. Agency is defined as "the capacity, condition, or state of acting or of exerting power; a person or thing through which power is exerted or an end is achieved" [3]. Indeed, the agency of machines differs from that of the human. While machines have no conscious selection and self-reflectiveness [4,5], some are yet perceived as having agency. Adaptability [6,7], purposeful-looking movement [8], complementary personalities [9], and humanlike appearance [1] are defined features in literatures that contribute to human's tendency in ascribing (social) agency to machines. We argue that among all these features, autonomy in artificial agents induces sufficient ques of the agency. In AI and robotics, autonomy is described as the capacity to perform tasks interdependently from external control [10,11]. Taking this into consideration, we can define machine agency as delegated authority; artificial agents are bound by the goal function set by the human agent but have (relative) autonomy in their operation and interaction to achieve the pre-defined goals. Furthermore, machine agency is not something pre-made that can be put inside an artefact but will be attributed to them.

Taking together, we conceptualize artificial agents in three different types (directed agent, semi-autonomous agent, and quasi-autonomous agent) using several dimensions. Though these artificial agents are all situated and embodied agents, major differences exist among them in how they control the input-output cycle, pursue their goal and the extent to which they can perform tasks independently (Table 1).

	Directed agent	Semi-autonomous agent	Quasi-autonomous agent
Control structure	rule-based, algorithmic	Fuzzy, planner	autonomous control
Goal setting	Achievement of the set goal	Adaption within certain limits	Fundamental change of goal is possible
Interaction with real world	Passive to limited sensory capability	Active data collection and mining	Multisensory perceptions

Autonomous capability	Auto-moving/acting	Auto-learning	Auto-organizing
Response selection	Task and context specific	Task variable, context specific	Task and context variable
Example	Thermostat	Self-driving car	"Survivor" [14]

Table 1: a typology of artificial agents

DISCUSSION AND OUTLOOK

Attributing agency to artificial agents causes a vast amount of discussion about the consequences in practice. Since a key aspect in acceptance of the technology is the social factors [12], understanding the social components of human-agent interaction will provide valuable insights not only for designers but also those who co-work with artificial agents. Therefore, to identify directions for future research, the role of artificial agents in the context of social interaction on different levels need to be analyzed. At the micro level, it is important to consider how human agents perceive the interaction with artificial agents as the subjective readiness of human agents to accept this interaction is more relevant than the objective capabilities of artificial agents [13]. Accordingly, at meso level, greater attention shall be given to social dimension of work and collaboration between the agents. Finally, at the macro level, it is necessary to reflect how the social fabric of society is going to change when we construct and perceive the artificial agents as "other", an abstract notion of what society considers good.

REFERENCES

[1] Lee, J. G., Kim, K. J., Lee, S., & Shin, D. H., International Journal of Human-Computer Interaction, 682-691, 2015

[2] Bolton, A., Goosen, L., & Kritzinger, E., Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists,6, 2016

[3] https://www.merriam-webster.com/dictionary/agency

[4] Engen, V., Pickering, J. B., & Walland, P., International Conference on Human-Computer Interaction, 96-106, 2016

[5] Rammert, W., Agency Without Actors? New Approaches to Collective Action, 98-112, 2012

[6] Harvey, I, Evolutionary Robotics: From Intelligent Robots to Artificial Life, 207-230, 2000

[7] Franklin S. & Graesser, A. Intelligent Agents III Agent Theories, Architectures, and Languages, 21-35, 1997

[8] Scholl, B. J., & Tremoulet, P. D., Trends in Cognitive Sciences, 299-309, 2000

[9] Lee, K. M., Peng, W., Jin, S. A., & Yan, C., Journal of communication, 754-772, 2006

[10] Harbers, M., Peeters, M. M., & Neerincx, M. A. A, World with Robots, 19-33, 2017

[11] Beckey, G.A., Robot ethics: the ethical and social implications of robotics, 2011

[12] Thunberg, S., Thellman, S., & Ziemke, T., Proceedings of the 5th International Conference on Human Agent Interaction, 443-446, 2017

[13] Echterhoff, G., Bohner, G., & Siebler, F., Zeitschrift für Sozialpsychologie, 219-231, 2006

[14] http://www.2045.com/news/33053.html