Special issue on human–automation coagency

T. Inagaki

Published online: 4 October 2011
© Springer-Verlag London Limited 2011

In human supervisory control, machines control the process according to directives given by human operators, where the humans have the final authority over the machines or automation. There are many human–machine systems in our society that are modeled nicely by the human supervisory control model. A glass-cockpit passenger aircraft is a typical example of such a system. However, even in transportation, human–automation relation for technologically advanced automobile may be quite different from that for a glass-cockpit aircraft. In automobile, it is only the automation that may be able to help the driver by backing up or replacing him/her to assure safety: Neither air traffic controller-like function nor copilot-like colleague is available there (Inagaki 2010).

An advanced driver assistance system (ADAS) is a machine to assist a human to drive a car in a dynamic environment. Some functions of ADAS may include: (a) perception enhancement that helps the driver to perceive the traffic environment around his/her vehicle, (b) arousing attention of the driver to encourage paying attention to potential risks around his/her vehicle, (c) setting off a warning to urge the driver to take a specific action, and (d) automatic safety control that is activated when the driver takes no action even after being warned or when the driver's control action seems to be insufficient.

The first two functions, (a) and (b), are to help the driver to recognize or understand the situation. Understanding of the current situation determines what action needs to be done (Hollnagel and Bye 2000). Once situation diagnostic decision is made, action selection is usually straightforward

(Klein 1993). However, the driver may not be successful in action selection decision. Function (c) is to help the driver in such a circumstance. Any ADAS that uses only the three functions, (a)–(c), is completely compatible with the *human-centered automation* principle (Billings 1997) in which the human is assumed to have the final authority over the automation.

Suppose an ADAS contains the forth function, (d). Then, the ADAS may not be fully compatible with the human-centered automation, because the automation can implement a safety control action without any human intervention. Should we ban such an ADAS just because it can implement an action that is not ordered by the driver?

It is well known that highly automated machines sometimes bring negative effects, such as the out-of-the-loop performance problem, loss of situation awareness, complacency or overtrust, automation surprises. However, humans have limited capabilities, and they might fail to understand the situation, select a right action and implement it appropriately, especially when available time and information are quite limited. Today's machine can sense and analyze a situation, decide what must be done, and implement control actions. Cannot such a smart machine help the humans in more a positive manner as a teammate to the humans?

The systems in which humans and technology collaborate together to achieve a common goal are called joint cognitive systems (Hollnagel and Woods 2005), in which human–automation coagency is central for realizing sensible human–automation partnership.

This special issue of Cognition Technology and Work intends to discuss human–automation coagency for collaborative control of automobile. The first topic is authority and responsibility. The first three papers argue the need of situation- and context-dependent sharing of authority between the human and the automation, without assuming

T. Inagaki (✉)
Department of Risk Engineering,
University of Tsukuba, Tsukuba 305-8573, Japan
e-mail: inagaki@risk.tsukuba.ac.jp

which agent, human or automation, should always be superior to the other. The paper by Flemisch, Heesen, Hesse, Kelsch, Schieben, and Beller investigates relations among four cornerstone concepts (ability, authority, control, and responsibility) for implementing shared and cooperative control. A graphical tool is presented for visualizing and designing a dynamic balance between humans and machines. The paper by Abbink, Mulder, and Boer discusses haptic shared control as promising human–machine interface that can prevent some of the negative effects of automation. The paper presents two case studies for automotive steering guidance to prove the efficacy of haptic shared control. The paper by Inagaki and Sheridan investigate cases in which the computer detected mismatches between the driver’s action and a given traffic situation. The paper proves mathematically that a machine-initiated trading of authority is useful to assure the driver’s safety, even if the computer’s judgment may be wrong.

If the automation behaves nicely and intelligently as a partner, the human’s trust in automation may progress over time. It is well documented that the notion of trust has some dimensions; see, e.g., Lee and Moray (1992). An intelligent ADAS is often complex, and it is thus possible for the driver to overrate one or more of the dimensions of trust. Inappropriate trust in the ADAS may yield inappropriate use of it (Parasuraman and Riley 1997). The second topic of this special issue is trust in and reliance on automation and two papers deal with the topic. The paper by Ghazizadeh, Lee, and Boyle points out that approaches by cognitive engineering (CE) and information systems (IS) are different and independent when investigating users’ acceptance of automation. By finding the complementary relations between the two approaches, the paper proposes an integrated framework for assessing automation adoption that incorporates micro-level concepts of human–technology coagency as well as effects of macro-level factors. The paper by Itoh discusses the need of a theoretical framework to represent how human operator’s trust can become excessive and gives a model of trust so that overtrust can be defined rigorously. The model enables to distinguish three types of overtrust. The paper discusses experimental results

exhibiting two types of overtrust that were observed in simulated drivings with an adaptive cruise control (ACC).

The third topic of this special issue is the need of learning to make it possible for human–automation coagency to evolve over time. It is vital for both of the human and automation to understand their capabilities and limitations in order to rely on and compensate with each other appropriately. This special issue’s last paper by Vanderhaegen discusses the need of cooperation and learning to realize meaningful coagency between the human and the automation, especially for the automation to predict actions by the human. An idea of cooperative learning-based ADAS is presented for a new and future ADAS.

The special issue was proposed partly based on the panel discussions, “Human-automation coagency for collaborative control: Issues and perspectives for advanced driver assistance systems,” held in the IFAC HMS 2010 at Valenciennes, but was enriched by inviting distinguished researchers after the conference. Human–automation coagency is a fascinating research area with full of topics for further investigations. It would be nice if this special issue could be successful in attracting keen interests from the readers.

References

- Billings CE (1997) Aviation automation—the search for a human-centered approach. LEA, Mahwah
- Hollnagel E, Bye A (2000) Principles for modeling function allocation. *Int J Hum Comput Stud* 52:253–265
- Hollnagel E, Woods DD (2005) Joint cognitive systems: foundations of cognitive systems engineering. CRC Press, Boca Raton
- Inagaki T (2010) Traffic systems as joint cognitive systems: issues to be solved for realizing human-technology coagency. *Cogn Tech Work* 12:153–162
- Klein G (1993) A recognition-primed decision (RPD) model of rapid decision making. In: Klein G et al (eds) *Decision making in action*. Ablex, New York, pp 138–147
- Lee JD, Moray N (1992) Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35(10): 1243–1270
- Parasuraman R, Riley V (1997) Humans and automation: use, misuse, disuse, abuse. *Hum Factors* 39(2):230–253