Proposal of a panel session at the IFAC HMS 2010

HUMAN-AUTOMATION COAGENCY FOR COLLABORATIVE CONTROL

1. Panel Organizer

Toshiyuki INAGAKI
Department of Risk Engineering, University of Tsukuba
Tsukuba 305-8573 JAPAN
inagaki@risk.tsukuba.ac.jp

2. Panel Topic

Driving a car requires a continuous process of perception, cognition, action selection, and action implementation. Various automated functions are implemented in an Advanced Driver Assistance System (ADAS) to assist a human to drive a car in a dynamic environment. Such functions 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 encourage the driver to take a specific action to avoid an incident or accident, and (d) 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 (Inagaki 2008).

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 & Bye 2000). Once situation diagnostic decision is made, action selection is usually straightforward, as has been suggested by recognition-primed decision making research (Klein 1993). However, the driver may sometimes feel difficulty 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 (Woods 1989; 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 always be fully compatible with the human-centered automation, because the system can implement an action that is not ordered by the driver explicitly. Some automatic safety control functions have been already implemented in the real world. The fact that the driver may not always be kept as the final authority over the automation in such ADAS does not mean that those designs should be avoided. On the contrary, the automatic safety control functions are effective especially for automobile and indispensable for attaining driver safety, by taking into account the domain-dependence of human-centered automation (Inagaki 2006).

However, the regulatory authorities sometimes take a cautious stance on putting into practical use ADAS with automatic safety control functions. There are two big issues behind such an attitude. The first is the issue of authority and responsibility, as has already been mentioned. The Convention on Road Traffic (1968) states that "Every driver of a vehicle shall in all circumstances have his vehicle under control so as to be able to exercise due and proper care and to be at all times in a position to perform all manoeuvres required of him" (Article 13.1). Therefore, it is usually assumed that the driver must be in charge and that the functions provided by the automation are to assist the driver, instead of trying to replace the driver. The idea is quite similar to the principles of human-centered automation in aviation.

The second is the issue of the drivers' possible overtrust in and overreliance on an intelligent and autonomous ADAS. The following question is frequently asked: "When the ADAS is capable of coping with the situation automatically without any intervention of a driver, is not it possible for the driver to be overly reliant on the system and give up active involvement in driving?" However, discussions regarding overtrust and overreliance have not been rigorous enough yet until this point. For example, the terms 'overtrust' and 'overreliance' are sometimes treated as if they are synonyms, which is of course incorrect. Moreover, countermeasures to prevent overtrust are quite different from those to prevent overreliance.

This panel session is organized to: (1) make it clear what questions to be asked to look for a wise, sensible, and comfortable relations among human drivers and automation, (2) identify and develop necessary viewpoints and methodologies to attack the problems, and (3) propose solutions to the problems.

3. Organization of panel members

The panel consists of 5 members.

Prof. Toshiyuki Inagaki, Department of Risk Engineering, University of Tsukuba, Tsukuba, Japan.

Dr.-Ing. Frank Flemisch, DLR Institute of Transportation Systems, Braunschweig, Germany.

Prof. Makoto Itoh, Department of Risk Engineering, University of Tsukuba, Tsukuba, Japan.

Prof. Guy A. Boy, Center for Interaction Design, Florida Institute of Technology, Melbourne, USA.

Prof. Erik Hollnagel, Industrial Safety Chair, École des Mines de Paris - Pôle Cindyniques, Sophia Antipolis Cedex, France

4. Structure of the panel session

Each panel member delivers a talk according to the plan described below. Short Q&A time follows each presentation just for conveniences of the audience. Time for discussions among panelists as well as between the audience and the panelists are set as the last part of the session.

(1) Prof. Toshiyuki Inagaki: Issues to be solved for realizing human-automation coagency for collaborative control.

This talk is planned as an introduction to the audience at the panel session. The author tries to explain to the audience what the topics of the session are, why the topics must be discussed, what kinds of viewpoints are needed in the discussions, and where human-automation systems are going (i.e., what kind of solutions may be expected)

(2) Dr. Frank Flemisch: Towards a dynamic balance between human and automation: Authority and responsibility in cooperative control situations.

A generic framework of cooperative control of movement is sketched to show that the cooperation is not static, but can shift between human and automation. A strong double-bind between authority and responsibility is stressed. Practical applications of theoretical thoughts are discussed, with the H-Mode project as well as the HAVEit project as examples.

(3) Prof. Makoto Itoh: Toward "over-trust"-free advanced driver assistance systems.

It is argued at first that over-trust and over-reliance have not been clearly distinguished. A unified theoretical framework is sketched for analysis and evaluation of over-trust and over-reliance. The speaker will present a research result that shows a process for the driver to reach the state of placing over-trust in an ADAS.

(4) Prof. Guy A. Boy: Orchestra organizational automation.

Authority is defined from two main perspectives, i.e., control in the engineering sense, and accountability in the legal sense. An orchestra model is proposed as an alternative to the traditional army-type model. It is argued that interaction models are useful to support the way cognitive functions are implemented in complex software.

(5) Prof. Erik Hollnagel: Intractability, planning, and automation.

It is argued that growing intractability of systems does on the one hand explain why there are limits to the use of automation in practice, and why the 'optimistic' application of automation can create risks and traps. The same growth in intractability, however, also becomes a reason why automation is necessary to an increasing degree.

(6) Discussions among panelists and between the audience and the panelists.

(7) Wrap-up by Prof. Toshiyuki Inagaki

References

Billings, C.E. (1997). *Aviation automation – The search for a human-centered approach*. LEA.

Convention on Road Traffic (1968). 1993 version & amendments in 2006.

Hollnagel, E. & Bye, A. (2000). Principles for modeling function allocation. *Int. J. Human-Computer Studies*, 52, 253-265.

Inagaki, T. (2008). Smart collaborations between humans and machines based on mutual understanding. *Annual Reviews in Control*, vol. 32, pp. 253-261.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. Klein, et al (Eds.), *Decision making in action* (pp. 138-147). Ablex.

Woods, D. (1989). The effects of automation on human's role: Experience from non-aviation industries. In S. Norman & H. Orlady (Eds.). *Flight deck automation: Promises and realities* (NASA CR-10036, pp.61-85). NASA-Ames Research Center.