SMART COLLABORATION BETWEEN HUMANS AND MACHINES
BASED ON MUTUAL UNDERSTANDING

T. Inagaki

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

Abstract: To improve the safety and comfort of a human-machine system, the machine needs to ‘know,’ in a real time manner, the human operator in the system. The machine’s assistance to the human can be fine tuned if the machine is able to sense the human’s state and intent. Related to this point, this paper discusses issues of human trust in automation, automation surprises, responsibility and authority. Examples are given of a driver assistance system for advanced automobile.

Keywords: Authority and responsibility; automation surprises; behavior monitoring; human-centered automation; human/machine interaction; risk and safety

1. INTRODUCTION

A human uses a machine with an expectation that it can extend his/her capability or help him/her to achieve a goal efficiently with fewer burdens. The machine must be designed appropriately so that it may be easy for the human to: (1) understand what the machine can or cannot do, (2) give directives to the machine, (3) monitor what the machine is doing, and (4) intervene in machine control when necessary. The machine thus is required to be an agent that is faithful to the human and is able to perform precisely what it is ordered to do. If the human’s decision and its associated directive to the machine are correct, he/she can obtain a result that matches his/her goal and the situation at the time. In reality, however, the human can fail to give a proper directive to the machine in several ways. One of such cases may be where the human’s understanding of a given situation (and thus his/her decision) is not correct for some reasons, such as inattention or internal distraction. Another case may be where little time is left for him/her to implement a necessary action, such as giving a directive to the machine, although the human’s understanding of a given situation is correct.

The machine may need to implement some control actions, when it determines that the human might be in a condition where he/she is unable to give directives to the machine. In other words, the machine might need to be smart so that it can behave like a human friend (or a teammate) who would try to understand the partner’s psychological/physiological conditions, the situation at the time, what the partner is going or trying to do, and whether the partner’s intent or action matches the situation. In the area of automobile, for instance, various research projects have been conducted world-wide to develop smart machines that provide the drivers with various support functions for enhancing comfort and safety (see, e.g. Witt, 2003; Akamatsu and Sakaguchi, 2003; Saad, 2005; Cacciabue and Hollnagel, 2005; Tango and Montanari, 2005; Amditis et al. 2005; Furugori et al, 2005; Panou et al. 2005). Development of a situation-adaptive Driver Assistance System (DAS) is one of such approaches.

The situation-adaptive DAS was developed by the author and his colleagues in a project supported by the Government of Japan (Inagaki, 2007). The developed DAS provides the driver with multi-layered assist functions (Fig. 1). In the first layer, a driver’s situation recognition is enhanced for proper decisions and actions. It is believed that understanding of the current situation determines what action needs to be done (Hollnagel and Bye, 2000). In the second layer, the DAS monitors the driver’s behavior and traffic condition to evaluate whether his/her intent and behaviors match the traffic condition. When the DAS detects a deviation from normality, it gives the driver an alert or a warning to make him/her come back to normality. In the third layer, the DAS provides the driver with automatic safety control functions, if the deviation from normality still continues to be observed or if little time is left for the driver to cope with a traffic situation. The situation-adaptive DAS adjusts its assist functions dynamically so that they may fit to the human’s intent, psychological/physiological conditions, and the traffic condition. The adjustment of assist functions is made in a machine-initiated manner (Scerbo, 1996; Inagaki, 2003) by inferring intent and
conditions of the human through monitoring his/her behaviors. For instance, the DAS can implement control actions based on its own decisions.

![Diagram of Multi-layered driver assist]

**Fig. 1 Multi-layered driver assist**

This paper presents the benefits of mutual understanding between humans and machines for realizing smart collaboration, as well as the necessity of the machine-initiated (instead of the human-initiated) decision and control in order to assure comfort and safety. Discussions are also made on the issues of trust, automation surprises, responsibility and authority, especially when humans and machines monitor with each other the partner’s behaviors.

2. ADVANCED SAFETY VEHICLE

Before going into discussion of the situation-adaptive DAS, it would be beneficial to review the Advanced Safety Vehicle Project, one of national projects in Japan. An Advanced Safety Vehicle (ASV) is defined as a vehicle equipped with technology-based support systems that can assist drivers to enhance safety under normal as well as time-critical situations. The ASV Project aims to promote development of new technologies for reducing traffic accidents. The project is carried out through collaboration between vehicle manufacturers, related organizations (such as user associations, insurance companies, dealer associations), academia and government agencies (such as National Police Agency, Ministry of Internal Affairs and Communications, Ministry of Economy, Trade and Industry, and Ministry of Land, Infrastructure and Transport) (MLIT, 2007).

The ASV Project was kicked off in 1991. In ASV-1, the first 5-year phase of the project, technological possibilities and accident reduction effects were investigated. In ASV-2 (1996-2000), ASV design principles and technology development guidelines were established. Demonstrations and exhibitions were also made with 35 ASVs. In ASV-3 (2001-2005), the driver assistance concept was developed, and ASV popularization strategies were examined. Up to ASV-2, driver assistance systems of the onboard self-sensing type had been investigated. Such standalone systems can cope with hazards that are within the driver’s field of view or its equivalent. However, they may fail to detect hazards that are outside or barely within the driver’s field of view. To cope with such hidden hazards, communication technologies were introduced to ASVs. Driver assistance systems of the communication-based type can obtain necessary information through road-to-vehicle and inter-vehicle communications. In 2005, verification tests were made with a total of 17 vehicles (9 passenger cars, 4 heavy vehicles and 4 motorcycles). Now, ASV-4 (2006-2010) is in progress, in which establishment of comprehensive ASV safety strategy as well as promotion of full-scale popularization of ASV technologies are planned. In collaboration with the ITS (Intelligent Transport Systems) Promotion Council, a series of large-scale experiments are conducted now in 2008 on public roads in Tokyo, Yokohama/Yokosuka, Tochigi, Nagoya, and Hiroshima areas in Japan to evaluate and validate efficacy of driver assistance systems that use communication-based and onboard sensing-based technologies.

The ASV Project classifies the forms of assistance by the onboard system into 6 categories: (a) **perception enhancement** that helps the driver to perceive the traffic environment around his/her vehicle, (b) **workload reduction** to reduce the driver fatigue or to let the driver pay more attention to his/her traffic environment, (c) **presentation of information** that may be useful in decision making, (d) **arousing driver’s attention** to encourage the driver to pay attention to the potential risk around his/her vehicle, (e) **providing warnings** to
encourage the driver to make appropriate actions to avoid an accident/incident, and (f) accident avoidance control that is activated when the driver has no action even after being warned or when the driver’s countermeasure action seems to be insufficient. The first two, (a) and (b), are assistance during normal driving conditions, and the latter four, (c) – (f), are support functions mainly for cases in which some hazard may be there around the vehicle. Fig. 2 depicts the ASV Design Principles (MLIT, 2007).

Fig. 2  ASV design principles

Although ASV technologies provide drivers with various assistances, it is assumed that drivers should play the primary role in driving vehicles safely. In other words, the driver is responsible for safe driving. The principle coincides with so-called human-centered automation (Woods, 1989; Billings, 1997) in which the human bears the ultimate responsibility for system safety and, thus, human locus of control is assumed, because authority and responsibility are interconnected (Table 1).

Table 1. Principles of human-centered automation

<table>
<thead>
<tr>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>The human bears the ultimate responsibility for safety of aviation system.</td>
</tr>
<tr>
<td>Therefore:</td>
</tr>
<tr>
<td>◦ The human must be in command.</td>
</tr>
<tr>
<td>◦ To command effectively, the human must be involved.</td>
</tr>
<tr>
<td>◦ To be involved, the human must be informed.</td>
</tr>
<tr>
<td>◦ Functions must be automated only if there is a good reason for doing so.</td>
</tr>
<tr>
<td>◦ The human must be able to monitor the automated system.</td>
</tr>
<tr>
<td>◦ Automated systems must, therefore, be predictable.</td>
</tr>
<tr>
<td>◦ Automated systems must be able to monitor the human operator.</td>
</tr>
<tr>
<td>◦ Each element of the system must have knowledge of the others’ intent.</td>
</tr>
<tr>
<td>◦ Automation must be designed to be simple to learn and operate.</td>
</tr>
</tbody>
</table>


Fig. 3 illustrates the ASV concept more concretely. For each of the interactions, (1) - (6) in Fig. 3, between the driver and the assistance system, the following assumption is made, respectively (MLIT, 2007):

1. The system should act in line with intent of the driver.
2. The system should assist the driver to perform safe driving and steady operation.
3. The driver should monitor operations of the assistance system when it is in action.
4. The system should not cause overconfidence or over-trust of the driver.
5. The system, when it is in action, should allow the driver’s intervention to override its operation.
6. The system’s control should be smoothly passed over to the driver when the situation goes beyond the range of the system.
Humans working with intelligent machines sometimes suffer from negative consequences of automation, such as the out-of-the-loop performance problem, loss of situation awareness, complacency or lack of vigilance, over-trust in automation, automation surprises (see, e.g., Woods 1989; Wickens 1994; Endsley and Kiris 1995; Sarter and Woods 1995; Parasuraman and Riley 1997; Sarter et al. 1997; Inagaki and Stahre 2004; Hollnagel and Woods 2005). The above requirements (1) - (6) have been set based on such concerns and lessons learned from previous incidents and accidents in human-machine systems. The human interface must, thus, be designed appropriately so that it may be easy for the human to (a) share the situation awareness with the automation, (b) recognize the intent of the automation, (c) understand the rationale of the automation's judgment, and (d) perceive limitations of functional abilities of automation (Inagaki, 2006).

A large number of driver assistance systems have been developed so far (and some of them are already put to practical use). Such developed systems include: Full speed range adaptive cruise control system with brake control, Lane keeping assistance system, Nighttime pedestrian monitoring system, Forward obstacle collision prevention support system, Lane departure prevention support system, Crossing collision prevention advisory system. As has been noted earlier, driver assistance systems in the ASV project basically assume the driver's locus of control, unless the systems are designed for highly time-critical situations like a Pre-crash safety system.

A pre-crash safety system is a driver assistance system to reduce the damage caused by a collision into a forward vehicle (Fig. 4). There are several types of pre-crash safety systems. Common functions among them are to tighten the seat belts and add a warning to urge the driver to hit the brake pedal. Some systems can do more: When the system determines that the driver is late in depressing the brake pedal, it applies the brakes automatically. This type of pre-crash safety system goes beyond the borders of human-centered automation, because the automatic brake is not implemented based on the driver’s directive, but based on the decision made by the system. Whether such a machine-initiated automation invocation may be allowed or not is a difficult question, and conclusion may be domain-dependent (viz., the best solution for aircraft may not be applicable to cases of automobile). As a matter of fact, in cases of automobile, machine-initiated trading of authority from human to machine is sometimes indispensable for attaining safety of the driver (Inagaki, 2006).
3. WHY IS UNDERSTANDING OF DRIVER STATE NECESSARY?

The pre-crash safety system in the previous section applies the automatic brakes upon detecting the delay in action. In other words, the driver assistance function is activated when it detects the fact that the driver’s braking is absolutely late. Note here that, if the driver assistance system could predict that the driver might be late in braking, it would be able to apply the automatic brakes a bit earlier for a better result. To make such a prediction of action delay feasible, it is necessary to note that driving is a continuous process of perception (or recognition), decision, and action. For instance, when the driver is inattentive to the behavior of the lead vehicle, or when the driver has an inappropriate assumption in mind that, “The adaptive cruise control system would cope with every deceleration of the lead vehicle,” the driver would never try to anticipate or make him/herself ready to a possible rapid deceleration of the lead vehicle. His/her braking would therefore be late.

In the classic and important study, Treat et al. (1979) investigated causes of the traffic accidents in a systematic manner. They say that, “the human errors and deficiencies which caused accidents primarily involved recognition errors (intended to include both perception and comprehension problems), and decision errors” (p. 40), and that, “Based on probable cause results, recognition and decision errors were identified with nearly the same frequency, according to both on-site and in-depth results. However, based on definite cause results, the in-depth team identified recognition errors somewhat more frequently than decision errors (41.4% vs. 28.6%)” (p. 39).

Typical factors that can cause delay in recognition are identified as follows (Treat et al. 1979, p. 198): (R1) Inattention, in which attention is not paid, for instance, to traffic stopped, position of other cars, road features, merging or intersecting traffic. (R2) Internal distraction, such as conversation with passenger, sick person in the car, adjusting an audio device or air-conditioner. (R3) External distraction, such as other traffic, driver’s activity to look for a street or a house, fire or accident outside the host vehicle. (R4) Improper lookout that may occur when changing lanes, passing a vehicle, entering into an intersection, or pulling out from a parking place. Also, the followings are examples among many that are identified as contributing factors to decision errors (p. 199): (D1) Misjudgment, in which the distance or the relative speed to another vehicle (e.g. the lead vehicle) is judged inappropriately. (D2) False assumption, a typical example of which is to assume that the other driver has to stop at the intersection.

Is it then possible to estimate whether the driver is in a psychological/physiological condition that his/her recognition may be delayed or its associated decision may be improper? The answer is affirmative, as is discussed in the next section.

4. REAL TIME SENSING OF DRIVER STATE

The author conducted a research project, ‘Situation and Intent Recognition for Risk Finding and Avoidance’ with the support of Government of Japan during the period of July 2004 – March 2007.
aim of the project was to develop proactive safety technologies to realize a driver assistance system that provides the driver with various support functions in a situation-adaptive and context-dependent manner (Inagaki, 2007). The idea of the project was: Although it is not possible to ‘see’ inside of a driver’s mind to determine whether a driver’s situation recognition is correct or not, monitoring the driver’s action as well as the traffic environment (including positions and speeds of other vehicles around the host vehicle) may make it possible to guess: (a) whether the driver has lost situation awareness, (b) whether the driver’s interpretation of the traffic environment is proper, and (c) whether the driver is inactive psychologically or physiologically (Fig. 5).

![Fig. 5 Assessing driver’s situational recognition](image)

The following sensing data are collected (Itoh and Inagaki, 2007); see also Fig. 6: (1) **Eye fixation point**, monitored with eye trackers. The time length of a fixation is also measured. It is regarded as a fixation when driver’s gaze point stays within a region of around one degree, both horizontally and vertically for at least 100 ms. (2) **Head position** as well as **pitch**, **yaw** and **roll angles of the head**, measured with the head tracker. By combining the data obtained with the eye tracker, it is possible to identify which direction a driver is looking. Frequency of blinks is also calculated. (3) **Ear plethysmogram**, which gives the blood flow at the ear lobe. The **maximum Lyapunov exponent**, one of characteristics to represent chaotic dynamics, is calculated every second for estimating the driver’s mental workload. (4) **Nose tip temperature**, measured every one second with an infrared thermal imaging camera. It is known that an increase in mental workload may cause a decrease in the tissue blood volume and, thus, in the temperature at the nose-tip (Miyake et al. 2003; Veltman and Vos, 2005). (5) **Pressure distributions on the seat cushion and the back**, measured with sensor sheets at 2 Hz. Driving posture and frequency of body movement can be speculated with the pressure distributions (Itoh, 2008). (6) **Steering angle**, **strokes of gas and brake pedals**, which are recorded at 60 Hz. (7) **Facial expression** and **driving posture** are recorded with video cameras, which are used to interpret data obtained in (1) – (6).
What can be known with the data listed above? The followings are some examples.

**Case 1:** Suppose a driver is following a lead vehicle with the use of the adaptive cruise control (ACC) system. The driver here assumed that, “The lead vehicle shall never make a steep deceleration around here. Should there be such a case, the ACC would take care of it appropriately.” Based on such a false assumption, the driver put his foot far away from the brake pedal. A pattern recognition algorithm is available to identify the driver’s posture by analysing the pressure distribution on the seat; Fig. 7 gives examples of the pressure distribution on the seat cushion. The algorithm can distinguish whether the driver is ready to brake when necessary or he puts both feet on the floor (Itoh, 2008).

![Fig. 7 Pressure distributions on the seat](image)

(a) ready to brake (b) both foot on the floor

**Case 2:** Suppose that the host vehicle \(H\) is following vehicle \(F\), and that vehicle \(A\) is about to cut in just in front of vehicle \(F\) (Fig. 8). It may be possible for vehicle \(F\) to decelerate rapidly to avoid a collision to vehicle \(A\). If the driver of the host vehicle does not move his right foot near to the brake pedal, it may suggest that the driver is inattentive to a possible deceleration of vehicle \(F\). The driver’s inattentive state can be found from the pressure distribution on the seat.

![Fig. 8 Is a rapid deceleration anticipated?](image)

**Case 3:** Suppose the driver begins to allocate his attention to some non-driving tasks, because the cruise has been quite peaceful for quite a long time. He may turn around to take something to drink from behind, or may begin to manipulate the audio device, or may try to pick up something on the floor (Fig. 9). Such driver’s behaviors can be identified via the pattern recognition technique applied to the pressure distribution on the seat (Itoh, 2008).
Case 4: Suppose the driver is thinking hard for a solution to a problem that needs to be solved as soon as possible. Even when he looks ahead, his situation awareness might be poor due to high mental workload. Detecting whether the driver is distracted internally is not straightforward and requires a sensor fusion technique that combines multiple data, such as nose tip temperature, maximum Lyapunov exponent, head movement, and eye movement, by taking into account characteristics of individual drivers. For instance, a decrease in the temperature at the nose-tip is a good measure for some drivers to detect that they are under high mental workload. However, a temperature drop can also be observed when the drivers feel tense. In order to distinguish whether the driver feels tense or his/her mental workload is actually high, some other parameters need to be monitored (Itoh et al. 2006; Itoh, 2007; Itoh and Inagaki, 2007a, 2007b).

Case 5: Passing a slower vehicle is not an easy task for a large truck with a speed governor. Truck drivers usually try to keep their speed at the maximum allowable level on expressways. They tend to pass a slower vehicle even when speed difference is quite small. Suppose the traffic is not sparse and thus the driver needs to glance at the side mirror to a right time to pass a lead vehicle. The driver’s intent to change lanes can be detected through monitoring eye movement. Fig. 10 depicts a model that tries to relate the driver’s intent of making a lane change (where four levels, from “very low” to “high,” are distinguished to express strength of intent) and the frequency of glancing the side mirror (which is measured as the number of times that the side mirror has been checked during the last 10 seconds). The model has been developed by analysing behaviors of truck drivers (Zhou et al. 2006). An algorithm in (Zhou et al. 2006) can detect the driver’s intent of making a lane change a couple of seconds prior to the actual execution of the lane change maneuver.

5. WHAT IF A DEVIATION FROM NORMALITY IS DETECTED?

The situation-adaptive DAS developed in the ‘Situation and intent recognition for risk finding and avoidance’ project monitors the driver behavior and outside traffic environment to evaluate whether his/her intent and behaviors match the traffic condition. When it detects a deviation from normality (such as undesirable conditions or behaviors in Cases 1 – 4, or driver’s intent of an inappropriate action as in Case 5), the DAS gives the driver an alert or a warning to let him/her come back to normality. If the deviation from normality still continues to be observed, or if little time is left for the driver to cope with the situation, the DAS executes a safety control action autonomously based on its decision (Inagaki, 2007).
Case 6: Suppose the driver of the host vehicle $H$ determined to make a lane change, since the lead vehicle $A$ is slow. The DAS detects the driver’s intent through monitoring eye movement. The driver saw vehicle $B$ almost passed him on the left, and was about to begin steering the wheel to the left, failing to notice that a faster vehicle $C$ is coming from behind on the left lane (Fig. 11). What should the DAS do, if it knows the approach of vehicle $C$ through monitoring backward with a camera? Giving a warning may not be effective to avoid a collision when the speed of vehicle $C$ is extremely high. Then the DAS may have to take a safety control action immediately. Such a safety control action may be to make the steering wheel either slightly heavy to steer (soft protection) or extremely heavy to steer (hard protection). The soft protection is for correcting the driver’s interpretation of the traffic environment, while the hard protection is for preventing a collision from occurring (Inagaki & Itoh, 2007; Inagaki et al. 2007). Note here that activation of protection function is machine-initiated, and that the driver is not maintained as the final authority over the DAS.

Fig. 11 May authority be exercised by a machine?

Case 7: Suppose the driver of the host vehicle $H$ wants to make a lane change to pass a slower vehicle $A$ in Fig. 11. Glancing at the rear view mirror, the driver notices that faster vehicles, $C$ and $D$, are coming from behind on the left lane. Suppose the driver is seeking a time moment to make a lane change, by taking several looks at the side mirror. The DAS detected the driver’s intent to change lanes through monitoring the driver’s eye movement. The system can also infer that the driver’s attention to the forward vehicle $A$ may not be perfect. The DAS then put its safety control function into its armed position. If vehicle $A$ does not make any deceleration before the host vehicle’s driver completes a lane change, the DAS will never activate the safety control function and will put it back into a normal standby position. If a rapid deceleration of vehicle $A$ is detected, on the other hand, the DAS may apply an emergency brake immediately to avoid time delays that are inevitable when the driver with imperfect attention to the lead vehicle is supposed to apply the brakes himself.

6. CONCLUDING REMARKS

This paper has discussed the need and importance for the machine to ‘know’ the human operator in the system. If the machine is able to sense whether the human is in a good condition, what he/she is trying to do, whether he/she can accomplish the aim alone, then the machine’s support to the human can be fine tuned. We have seen that some sensing technologies and related methods are available for that purpose. However, no ‘universal sensing methods’ have been developed. It is necessary to tune sensing methods themselves at first so that they can extract individual characteristics properly and efficiently.

The next questions then would be how the machine assistance may be tuned, and who decides the tuning. A right type of assistance has to be chosen among alternatives, such as, perception enhancement, workload reduction, presenting information to aid decisions, arousing attention, directing a necessary action via a warning, and control implementation. Which type must be chosen is highly dependent on the state (or condition) of the human and a situation at the time. Moreover, a selection decision and a control implementation may have to be made by the machine. The need for such a machine-initiated decision and control has been discussed in this paper with the situation-adaptive DAS. Whether decision and control may be machine-initiated or has to be human-initiated is a crucial research issue in adaptive automation (Rouse, 1988; Parasuraman et al. 1992; Scerbo 1996; Inagaki, 2003). A purely human-centered automation point of view may not be willing to accept machine-initiated decision and control. However, humans may not always be powerful enough to cope with any given situation. Rouse (1988) pointed out many years ago that, “when an aid is most needed, it is likely that humans will have few resources to devote to interacting with the aid.”
It would be appropriate to distinguish the following two types of authority: (1) authority for a machine to decide and act when a human is unable to do, and (2) authority for a machine to prevent a human from doing what he/she wants to do (Inagaki and Sheridan, 2008). No serious concerns exist regarding the machine-initiated design for the first type. Whether to allow the second type may be controversial. Even when we may admit the second type, the following question arises: Between soft protection and hard protection, which is better? The hard protection can cause automation surprises, while the soft protection may not be free from an undesirable event in which the machine’s correct aid is cancelled out by an improper decision of the human operator with a false assumption in his/her mind (Inagaki and Itoh, 2007; Inagaki et al. 2007).

When discussing the authority issue, it also should be noted what kind of human we are talking about. Billings (1997) says, “The user almost always has more knowledge of the world state and its implications than the machine” (p. 228). However, the argument can be highly domain-dependent. For instance, an airline pilot with high competence and proficiency should not be considered equal with an ordinary car driver who does not usually receive any continual education program or training after acquiring a license. Human capabilities can also change as time passes by (e.g. some of the capabilities may degrade as a human becomes older). Moreover, psychological/physiological conditions may change hourly or daily within each individual. Technological progresses should also be counted in the discussion. In other words, a machine may not be what it used to be, and a machine in the near future may not be what it is now.

ACKNOWLEDGMENTS

This work has been partially based on the results of a research project, ‘Situation and intent recognition for risk finding and avoidance,’ conducted during the period of 2004-2007 with the support by MEXT, the Ministry of Education, Culture, Sports, Science and Technology, Government of Japan. The author expresses his thanks to members of the project for their great efforts. Thanks are extended to MLIT, the Ministry of Land, Infrastructure and Transport, Government of Japan, for their permission to reproduce figures of the ASV project.

REFERENCES


