Joint Cognition in Automated Driving: Combining Human and Machine Intelligence to Address Novel Problems

David Miller, Wendy Ju
Stanford University Department of Communication, Center for Design Research at Stanford
davebmiller@stanford.edu, wendyju@stanford.edu

Abstract

As in-vehicle automation becomes increasingly prevalent and capable, there will be more opportunity for vehicle drivers to delegate control to automated systems. as well as increased ability for automated systems to intervene to increase road safety. With the decline in how much a driver must be engaged, two problems arise: driver disengagement and reduced ability to act when necessary; and also a likely decrease in active driving, which may reduce the engagement a driver can have for the purpose of enjoyment. As vehicles become more intelligent, they need to work collaboratively with human drivers, in the frame of a joint-cognitive system in order to both extend and backstop human capabilities to optimize safety, comfort, and engagement.

Introduction

Human intelligence is good at solving novel problems, but humans fall short when tasked with maintaining vigilance and exercising supervisory control—especially when cognitively under-loaded (Parasuraman & Riley, 1997; Parasuraman, Sheridan, & Wickens, 2000). Machines can maintain vigilance and react faster than humans, when they are able to properly evaluate the present situation and have a programmed response ready. Humans have the capability to assess and respond to novel situations that may be beyond the capabilities of a computer, thus a dynamic sharing of control will provide the greatest total safety envelope, as well as allow for driver engagement and enjoyment.

This shared control regime represents a joint cognitive system (Woods, 1985), combining the best features of human and machine abilities in order to solve problems that may not be solvable by the entity in control of the vehicle at the time and under the conditions present. Currently, technical systems such as torque vectoring and Electronic Stability Program extend human capabilities, allowing greater control over the vehicle than would be possible without computer assistance. Human and machine abilities in automotive safety critical areas occupy a set of partially overlapping areas, where some human capabilities are currently superior to machine abilities, and in other areas machine abilities are superior (see Figure 1).

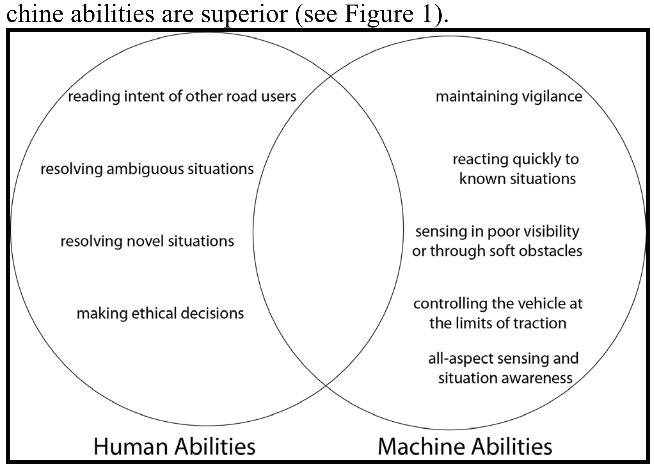

Figure 1. Partially overlapping areas of human and machine abilities in the area of road vehicle systems.

Driving 'Spaces'

Effectively, the space for human controlled driving is 'sandwiched' between two computer controlled regimes: (1) systems to delegate control to, and (2) systems that intervene to provide active safety (see Figure 2). With increasingly capable automation, the driver can delegate control when disengagement from driving is desired. This is not new—cruise control has been available for many years, and increasing sophistication has decreased the attention the driver is required to devote to the ambient environment, Safety has increased with the increase in automated system capabilities through system reactivity to the environment, slowing or stopping in response to obstacles ahead of the vehicle. Systems that intervene in imminent collision situations are in 2015 increasingly available (e.g. Forward Collision Warning and auto-braking), but act only when a collision is nearly certain (Jamson, Lai, & Carsten, 2008). The sensors and computation required are common to both types of systems, independent of intent. The conversion towards fully-automated driving capability will for some time leave a space where human intelligence is required, to make ethical decisions and to resolve ambiguous environmental cues in safety critical situations, and desired to make aesthetic decisions such as where to go and what path to take.

While many in the public would welcome fully automated vehicles (at least in concept), these will not be forthcoming in the near future. The regime of shared control, encompassing both computer assisted human driving (e.g. antilock braking), and human-assisted computer controlled driving (e.g. a driver choosing to pass another vehicle by overriding cruise control then returning to the set speed once the other vehicle has been passed). These shared-control relationships define a space for a joint-cognitive system, where both a human and computer share control dynamically, not in an either-or relationship. True human-only control exists only in very limited situations, such as total electrical failure; and complete computer control without human input would be a frightening thought—for example, a vehicle locking the doors and driving to a police station.

A Case Study Example

A rules-based system is fundamentally constrained by the fundamental inability to create rules in response to situations that have not been previously encountered or imagined by the designers or engineers creating the system. While artificial intelligence can extrapolate to some degree, truly novel situations are likely to confuse a purely technical system. An example of a situation that a computer would likely find ambiguous and difficult to resolve, but where a human would have little trouble in assessing the situation would be a fallen cyclist. A cyclist is an environmental feature that a computer can easily identify, plot a future path for, and knows to protect. A cyclist that has fallen at speed, becomes a very short, fast moving object with a variable and likely difficult to characterize radar and visual signature. Would the computer characterize this object as a human, as an animal, or as an inanimate object such as a ball? If the computer mischaracterizes the cyclist, it may choose to hit him or her, in preference to causing an alternate collision, as it could determine the optimal human protection scheme incorrectly. A human driver would instantly know to connect the cyclist that was previously upright and the present state of the cyclist now skidding across the road, and would (hopefully) plot a solution that would protect the vulnerable road user, potentially choosing to suffer vehicle damage or some risk of injury in preference to causing another road user serious injury or death.

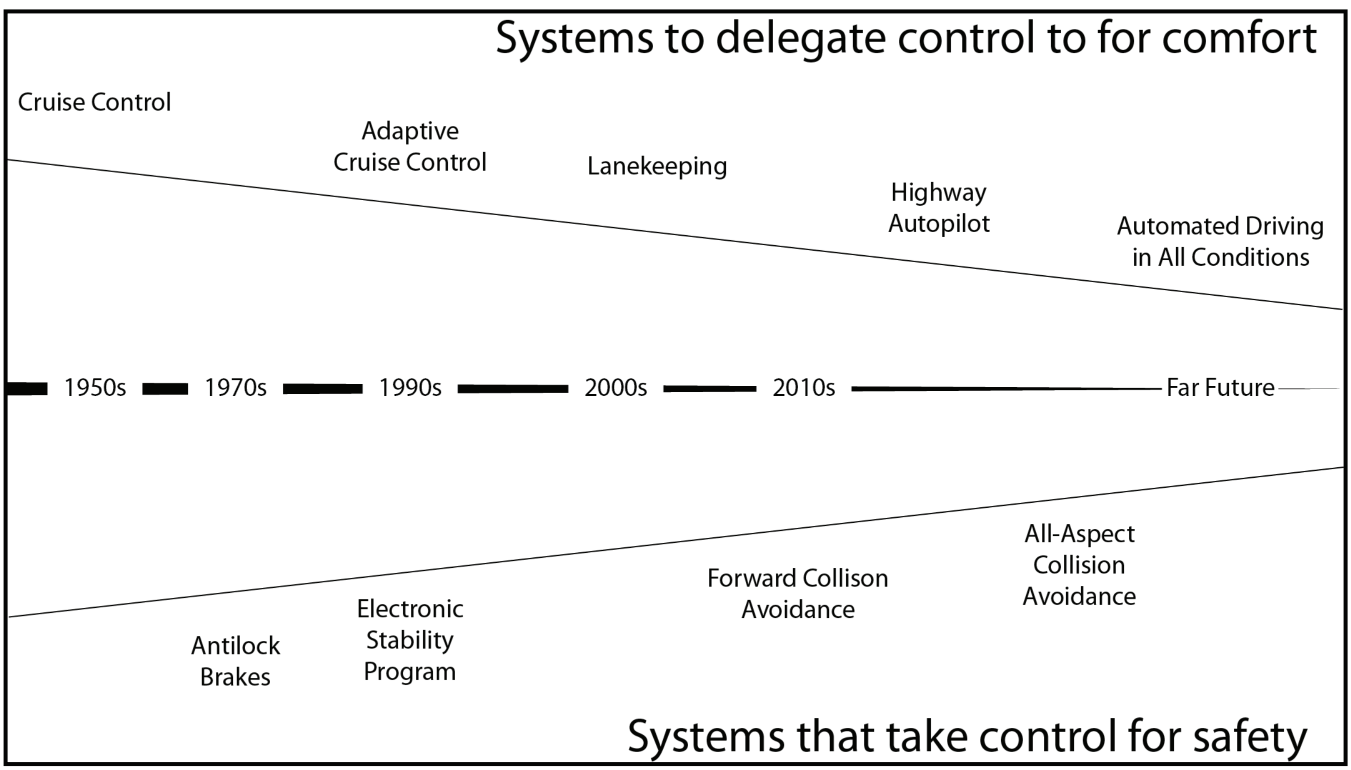

Figure 2. Vehicle automation systems can be divided into systems that drivers can delegate control to, and systems that take control from the driver. As systems become more capable, the space for driver control will narrow, providing improved safety and increased comfort.

A partnership between driver and computer

If the computer were to engage the driver in the decision making process, a more optimal solution may be achievable. Dependent on the parameters of the situation (time to collision, evasion paths available, following vehicle time to collision, engagement state of the driver, etc.) the computer can decide whether to delegate the decision to the driver, to incorporate the driver's input into the decision-making process, or to maintain or take control and to make all decisions independently. This decision of how and whether to engage the driver would require the use of extensive sensing inside the vehicle to determine driver state, and to process driver reactions to the situation, engaging the driver as part of the joint cognitive system.

Complex accident situations are seldom 'monolithic' in terms of occurring so rapidly that there is only one decision made that determines outcome—in such a time sensitive situation, such as a deer jumping out into the road, the computer would have to act without consulting the driver, the decision being unambiguous. In a slower-unfolding situation with ambiguity as to what to do or what the parameters of the situation are, the computer could delegate control to the driver, and if the driver does not respond in a way that the computer considers sufficiently safe, the computer can then intervene, as is the case with current collision avoidance systems that intervene at the last moment to avoid or reduce the severity of an impact. With the human supervising and interacting with the automated driving system, and automation 'backing up' the human driver, greater overall safety is likely achievable.

A joint-cognitive system would use the sensing capabilities of the vehicle (RADAR, LIDAR, machine vision) to evaluate the ambient environment, and cue the driver using a parallax-corrected heads-up display, to pay attention to potential hazards. The system would also need to assess the driver's state—is she or he attending to the critical features in the environment, is she or he capable of responding, and what response should be appropriate. This would have to be a collaborative decision, arrived at as a result of both human and technical signals.

Effective communication between the human and technical components of the vehicle system will not be limited to 'traditional' vehicle controls of throttle, brake, steering wheel, and even voice control; these modalities may not be fast enough in order to allow for the use of the driver as part of the computerized decision-making system. Sensors for evaluating driver physiological outputs are rapidly advancing, as is the understanding of how to use physiological data to assess driver state—which is essential to building a model in the computer of the driver's capabilities and to use the driver as part of the decision-making system. If the computer is to rely on the driver to make good decisions, or to know if the driver can fully evaluate the ambient environment, the computer must know how alert the driver is, what she or he is looking at, and is the driver aroused or concerned by a feature in the environment that the computer may not be able to evaluate. Research programs such as HAVEit (Rauch, Kaussner, Krüger, Boverie, & Flemisch, 2009) have highlighted the importance of assessing driver state in highly automated vehicles, as until automated systems can handle all situations, human intelligence will be required to resolve ambiguous situations.

The computer can understand the driver by using measures of arousal (e.g. pupilometry (Laeng, Sirois, & Gredebäck, 2012); sensing of heart rate variation (Appelhans & Luecken, 2006)); and emotion (e.g. FACS coding (Cohn, Zlochower, Lien, & Kanade, 1999; Ekman & Rosenberg, 1997)). Eye and head tracking can determine what the driver is attending to—is the driver attending to a concern inside the vehicle (e.g. media, passengers), the critical situation at hand, or other, non-critical external environmental features, and thus decide as to what the best driver engagement strategy would be.

The driver's senses of vision, hearing, and proprioception can also be valuable inputs to use to supplement technical sensors, and of course are the primary interaction channels between the driver and the world. The driver's ability to detect and filter stimuli from the environment, and to communicate that information to the vehicle either directly through the vehicle controls or mediated through the in-vehicle sensing layer, adds a powerful extra dimension of capability to the automated system, extending the human capabilities that are the foundation for driving under human control.

If the driver is to effectively share control with the computer, the system will have to be designed so to be understandable to a relatively untrained driver, and to effectively communicate with the driver. The driver will need to have a reasonably accurate and comprehensive mental model of the system's operation, and a clear view of the computer's assessment of the ambient environment, and of what it may do in response. The necessity for the computer to hold a model of the driver, and for the driver to hold a model of the computer presents a design challenge—designing understandable systems and feedback mechanisms so that the two entities can truly share control. With sensors and machine intelligence enhancing the capabilities of the driver, and backstopping human failings, and with human intelligence expanding the capabilities of the automated systems, the two can be considered to extend or expand each other's capabilities.

Conclusion

This arrangement of a human-computer 'crew' considers the two actors to be working together for one purpose, and to be in communication with each other, through both explicit and implicit channels, each oversees and interacts with the other. This partnership extends human capabilities by first providing extended information to the driver, by allowing the computer to step in when human frailties compromise safety, and further by leveraging human capabilities to be used to solve problems that humans are uniquely suited to solving at this time. While no system, even one that achieves a much larger total safety envelope by combining human and machine abilities, will be able to eliminate all accidents, this paradigm of shared control will likely lead to vastly improved safety, as well provide for greater driver engagement and enjoyment.

References

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, *10*(3), 229–240. doi:10.1037/1089-2680.10.3.229

Cohn, J. F., Zlochower, A. J., Lien, J., & Kanade, T. (1999). Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology*, *36*(01), 35–43.

Ekman, P., & Rosenberg, E. L. (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press.

Jamson, A. H., Lai, F. C. H., & Carsten, O. M. J. (2008). Potential benefits of an adaptive forward collision warning system. *Transportation Research Part C: Emerging Technologies*, *16*(4), 471–484. doi:10.1016/j.trc.2007.09.003

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. doi:10.1177/1745691611427305

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. doi:10.1518/001872097778543886

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286–297. doi:10.1109/3468.844354

Rauch, N., Kaussner, A., Krüger, H.-P., Boverie, S., & Flemisch, F. (2009). The importance of driver state assessment within highly automated vehicles. In *Conference Proceedings of the 16th World Congress on ITS* (Vol. 21). Stockholm, Sweden: HAVEit. Retrieved from http://haveit-eu.org/LH2Uploads/ItemsContent/25/3117-FULL-PAPER-THE-IMPORTANCE.pdf

Woods, D. D. (1985). Cognitive Technologies: The Design of Joint Human-Machine Cognitive Systems. *AI Magazine*, *6*(4), 86.