Human-Autonomy Teaming

Watching the online videos of the foolish things that people have done in "autonomous cars"—such as tricking the wheel detection system with a soda can while driving on the highway, working on their computers, and, in some cases, even sitting in the backseat of the car—leads to two questions:

- How well do these people understand (and appreciate) both the capabilities and limitations of the current technology?
- How aware is the autonomy software of the operator's lack of preparation if something goes wrong?

The possibility of situations that differ greatly from the nominal design conditions result in many design challenges, so the insights in [1] and [2] on how to design systems to function effectively as human-autonomy teams will be of interest to many researchers in the control systems community. Although written by the Office of the Chief Scientist from the perspective of the U.S. Air Force, with a simple mapping of "airman" to "operator," many of the insights on developing autonomous systems to operate *in the wild* should apply to many other domains.

The report discusses both the importance and difficulties of creating autonomy software that is sufficiently robust to function without human intervention. The claim made is that autonomy algorithms implemented to date are too "brittle" and thus will need human intervention at some point during the operation.

Digital Object Identifier 10.1109/MCS.2015.2511940
Date of publication: 17 March 2016

It would be useful as a community to consider the premise that the systems being created are too complex in hardware and software to function reliably.

This brittleness could be a result of the trend for increasingly more complex hardware and software or the result of operating in a challenging environment that is outside the set of conditions envisaged during the original design.

Adaptation and learning algorithms might help to mitigate this brittleness, but adding them often leads to its own set of difficulties because there is a lack of understandability (the models/policies learned lack traceability to the physics of the system), limited verification of the system performance, and limited generalizability of the adapted system to other environments. Adaptation also leads to one of the other major challenges highlighted in the report, which is the difficulty of an operator to develop a level of trust in the autonomy that is appropriately calibrated to the reliability and functionality of the system in various circumstances.

It is well understood that human operators will need to be able to develop a high level of trust in the autonomy software, which includes addressing questions such as what certificates of performance can be made available and what levels of model accuracy can be included in that analysis. Achieving human trust in the autonomy will require that the algorithms and software achieve new levels of transparency and system self-health assessment and that joint operator-autonomy training be performed. For transparency, it is important to address the often-asked question "why is it doing that," which has historically been a very difficult goal to achieve in complex autonomous systems and remains a significant challenge in our field.

One of the more interesting observations in [1] is that, in these mixed-initiative human-autonomy teams, the software may also need to develop a level of trust in the operator. For example, if the driver is not paying enough attention (and possibly not even sitting in the driver's seat) or incapacitated, it seems reasonable that the level of trust of the car's autonomy software in the human operator should go down. The concern is that effective team performance depends on teammates having a shared understanding of how well each member is performing, how the actions of one is impacting the other, and what the teammates are planning to do in the future. To avoid this problem, [1] suggests that the system should be designed to ensure a *shared situational awareness* [3] so that the autonomy and operator can align goals and communicate decisions/desired actions appropriately. This leads to important implications in the design of future two-way communications between the autonomy software and the operator. In particular, this must include communication of the current status and future projections. Achieving this goal requires much more information is shared than the current state and that the information is shared in both directions. This must all be accomplished without operator overload, which might only be accomplished by relying on noninvasive methods.

Although I find the conclusion in [1] that human operators will always be needed a bit overly pessimistic about the capabilities of future autonomy algorithms, I agree that the issues highlighted in the report pose important research challenges for the IEEE Control Systems Society community if it is intended that the advanced algorithms being developed are to be accepted and deployed in future (semi-)autonomous systems. This includes recognizing early on that it is likely that the system being designed will not operate fully autonomously, and, based on the challenges raised in [1], it would be better to build a semi-autonomous system from start that appropriately accounts for the role of, and interfaces with, the operator. In addition to determining what information to communicate and how to communicate it, this includes basic problems such as what constraints and objective function to use when planning future actions (likely not the same as if this was an autonomous system [4]). The stability of systems with multiple controllers (operator and autonomy) and/or the stability of switched systems if there is a handoff between the two is another challenge that must be addressed.

Finally, I think it would be useful as a community to consider the premise that the systems being created are too complex in hardware and software to function reliably. This seems pessimistic, but how complex can a real system be such that it can still be verified? It would be useful to quantify and catalog the capabilities and advances in this area in overcoming both the brittleness (and/or the perception of brittleness) in current systems.

While these advances won't necessarily stop people from doing foolish things, it should help systems based on human-autonomy teams survive events more robustly.

REFERENCES

[1] M. R. Endsley. (2015, Dec. 3). Autonomous horizons: system autonomy in the air force—A path to the future, volume 1: human-autonomy teaming. [Online]. Available: http://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf?timestamp=1435068339702

[2] Secretary of the Air Force Public Affairs. (2015, Dec. 3). AF releases vision for development of autonomous systems. [Online]. Available: http://www.af.mil/News/ArticleDisplay/tabid/223/Article/601443/af-releases-vision-for-development-of-autonomous-systems.aspx

[3] M. R. Endsley and W. M. Jones, "A model of inter- and intrateam situation awareness: Implications for design, training and measurement," in *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments*, M. McNeese, E. Salas, and M. Endsley, Eds. Santa Monica, CA: Human Factors and Ergonomics Society, 2001, pp. 46–67.

[4] A. Clare, S. Andrew, M. Cummings, J. P. How, A. Whitten, K. Andrew, and O. Toupet, "Operator object function guidance for a real-time unmanned vehicle scheduling algorithm," *AIAA J. Aerospace Comput., Inform., Commun.*, vol. 9, no. 4, pp. 161–173, 2012.

Jonathan P. How

Runway Efficiencies

In aircraft control it is already known how to regulate the operation of the individual craft so as to govern the attitude of the craft about three axes to bring about directed, level flight, and to maintain constant airspeed and the elevation of the aircraft. Devices of this nature are referred to as automatic pilots, and they function well as far as control of a single vehicle is concerned. The problem of coordinated control of numerous craft in a sizable area transcends the capabilities of existing automatic pilots, however, and at present is handled by voice communication between the pilots of the various craft and supervisory operators at the control towers of the various airports. … Automatic supervision of the craft in an area so that each, following its schedule, arrives at the destination at such an instant that it can land immediately, without unduly cutting down the capacity of any runway by precautionary delays, especially under overcast weather conditions, is a goal which is difficult of achievement, but one which the present invention accomplishes.

— O. Hugo Schuck, "Air Traffic Control Apparatus," U.S. Patent #2,787,428, April 1957