5-7-2019

Human-Autonomy Teaming - an Evolving Interaction Paradigm: Teaming and Automation

Dale Richards

Bill Kaliardos

HUMAN-AUTONOMY TEAMING - AN EVOLVING INTERACTION PARADIGM: TEAMING AND AUTOMATION

Dale Richards
Coventry University, United Kingdom
Bill Kaliardos
Federal Aviation Administration, United States of America

Intelligent and complex systems are becoming common in our workplace and our homes, providing direct assistance in transport, health and education domains. In many instances the nature of these systems are somewhat ubiquitous, and influence the manner in which we make decisions. Traditionally we understand the benefits of how humans work within teams, and the associated pitfalls and costs when this team fails to work. However, we can view the autonomous agent as a synthetic partner emerging in roles that have traditionally been the bastion of the human alone. Within these new Human-Autonomy Teams we can witness different levels of automation and decision support held within a clear hierarchy of tasks and goals. However, when we start examining the nature of more autonomous systems and software agents we see a partnership that can suggest different constructs of authority depending on the context of the task. This may vary in terms of whether the human or agent is leading the team in order to achieve a goal. This paper examines the nature of HAT composition whilst examining the application of this in aviation and how trust in such stystems can be assessed.

We are surrounded by advanced technologies that are not only pervasive but make us dependent on their use. The relationship we share with intelligent systems traditionally provide us with an increase in efficiency, cost- or time-saving, or simply are technological tokens that we furnish our lives with in order to convince ourselves that we belong to an ever-changing technological world. However, we are amidst a rather different form of technology revolution that now affords us a degree of decision making that can either support, augment or even replace the human component entirely. This may conjure different perceptions depending on the individual and context of the technology being considered, but regardless of your view, there is a perception that such a relationship with this sort of technology is beneficial, efficient and welcomed. We have walked this path in the past, in the form of introducing automation into systems that allow this technology to release the human from conducting dangerous, difficult or dull tasks. The developments in advanced automation, autonomy and Artificial Intelligence (AI) are leading protagonists for this advanced technology, and in many instances we can argue that humankind has already achieved a degree of symbiosis that brings us closer to achieving human-autonomy teaming (HAT). This panel represents a series of papers that discusses some of the main issues with human-autonomy interaction in terms of the nature of a mixed human-autonomy team partnership. The structure of this paper and associated Panel Symposium will focus on the following aspects of HAT:

- Part 1: the nature of HAT composition and the manner in which team roles should be considered in terms of team membership,
- Part 1: how automation within aviation can be used as an example of a precursor of HAT,

- Part 2: the application of Cognitive Engineering as a means to design a HAT that is based on user and system requirements,
- Part 3: an examination of the role trust plays in HAT, with the focus on using neurophysiological measurement.

Human-Autonomy Teaming: An evolving interaction paradigm

While there has been a fair amount of literature that examines how human teams work well together (and when they do not), very little is known in terms of how the new human-autonomy teaming paradigm would fare. The nature of how humans act and interact within a *human-only* team construct is complex and requires a multidisciplinary approach in order to appreciate the different roles and behaviours (Salas et al, 2010). If we consider the seminal work by Belbin (1981) then we may begin to appreciate the importance of what a good team composition would look like, and also understand the nature team roles play within an effective partnership (Belbin, 1993). It would seem obvious that in order for a team to achieve a collective goal then individuals are likely to be defined in such roles in order to maximise efficiency and the likelihood of achieving that goal. Key to this effectiveness is the ability for team members to communicate with each other (Cooke et al., 2001), and unsurprisingly it has been suggested that a HAT would require similar roles and dynamics that we would expect to see in a human only team (Scholtz, 2003). A goal defined by a HAT would need to not only be a shared one (Baxter & Richards, 2010), but would also suggest that the nature of roles may be dynamic (depending of course on the context of the task/goal). This would see a team member sometimes being subservient, but in other cases providing plans or suggest actions to be carried out - which may either be performed by a human or an autonomous agent.

There are many different ways by which we can imagine a human interacting with an agent-based system, and indeed the role a human team member may play within a mixed team. Richards (2017) suggests this may present a number of alternative team constructs of HAT which begins to blur the boundaries of HAT composition. The construct of HAT may therefore not be as straight forward as we would traditionally perceive it, with the agent component not always being subservient to a human supervisor. In some instances we may see the agent sharing tasks in paralel to humans, and in some instances even providing commands to human team members. This raises a number of important aspects of HAT worth considering; namely that of acceptance and trust. Richards (2017) further stresses that in order to achieve an effective HAT it is important to define a framework of control that allows the delegation of control between the human and the agent. This is not new, as we have seen the demand for understanding the nature of how a human can interact with higher levels of automation and the manner in which authority may be delegated from human to machine (Sheridan & Verplank, 1978). Key to this interaction design is the manner by which information is passed between team members, ensuring that the intent of the team member conducting the task is visible.

Further to the composition and dynamic of HAT, the manner in which the augmentation of support is presented will dictate the effectiveness of the HAT. Within Aviation Psychology we are only too aware of the ineffective manifestations of human-automation that relate to losing situation awareness due to the human taking a back seat to the automation (Endlsey, 2016), the human not keeping up with automation in terms of understanding (Lyons, 2013) or even the human mistrusting the automation altogether (Lyons & Stokes, 2012). The next sections will examine these issues in a bit more detail.

Understanding Automation Behavior in Aviation: A practical limit to automation as a team member

The HAT concept differs from traditional human-computer interaction, or human-systems integration, in that it suggests some degree of human-like qualities for the "autonomous" agent (computer). This is in contrast to viewing the computer as, say, a processor that follows predetermined rules, and suggests the technology can offer intelligent and adaptive solutions (Cox, 2013). With recent advances in artificial intelligence, notably machine learning, the HAT concept can be reasonable, and advances in human-computer interface technologies can make HAT compelling. The issue in aviation—whether for pilots flying aircraft, or air traffic controllers (ATC) separating traffic—is that HAT might be in conflict with one of the core human-automation principles: the operator must sufficiently *understand* automated system behaviors. HAT concepts suggest a level of complexity or intelligence that make automation behavior challenging to understand and predict from the human operator perspective. Furthermore, when addressing off-nominal conditions (e.g., system degradation and failures, environmental conditions that are beyond the design limits), the challenge of understanding behavior of an autonomous agent is likely to be compounded.

That argument is a practical one, primarily based on lessons learned from decades of aviation research and practice in human-automation interaction (Brown, 2016). There is not a theoretical basis for what automation is capable of, given access to the right information, and given that it is operating within its design envelope. But even with proper design, clear human-automation design tradeoffs have emerged, and are well known (e.g. Hoffman & Woods, 2011; Woods & Branlat, 2011). For example, human supervisory control reduces excessive workload, but can lead to loss of situation awareness. Human-computer interfaces need to provide sufficient information to the human and be transparent and provide mode awareness, but information overload and display clutter can limit this (Selkowitz et al., 2017). History has demonstrated there are a number of inherent design tradeoffs with "traditional" human-automation interaction that suggest practical limits with respect to understanding by the human operator, even with substantial training and standard operating procedures. These limits are not just important; they are drivers for safety-critical systems in aviation and are explored in the following two examples.

Example 1: Flight Deck

Looking back, there are many well-known examples from aircraft accidents in which a primary factor was an off-nominal condition and a subsequent lack of understanding of that condition and the resulting system behaviors. The crash in 2009 of Turkish Airlines Flight 1951 is one of many examples. In that example the crew were clearly dependent on the auto-throttle to maintain velocity and height as they ran through their landing checklists. Unfortunately, a malfunctioning radio altimeter affected the auto-throttle configuration into a different flight mode (Dekker, 2009). In such examples, even when procedures have been developed for these off-nominal conditions, pilots can fail to first make sense of the situation, including the resulting automation behavior, in order to apply the procedures appropriately. Furthermore, it is impractical to develop procedures for all possible combinations of off-nominal conditions, including automation—leaving the pilots to figure out the mitigation. The point is not that there are shortcomings in pilots, pilot training, procedures, or designs; the point is that off-nominal conditions in the context of complex systems, complex procedures, and complex operations make it fundamentally challenging for pilots to understand the situation. Rather than designing automated systems to be more complex autonomous team agents, as in HAT, perhaps the focus should be on simplification.

Example 2: Air Traffic Control

Current research by the FAA on unmanned aircraft systems (UAS) illustrate solutions centered on simplification. UAS have great potential for intelligent automation to fly the aircraft—aviate, navigate, communicate—within the national airspace system (NAS). But as great as their automation potential is, we are seeing the importance of human understanding of UAS behavior through standardization. As an example, during off-nominal "lost link" situations, in which the pilot-in-command loses the control link, the UAS must fly autonomously for some period of time. During this time, what is best for the UAS can involve a complex decision that depends on its goals, system health, altitude, weather, communications infrastructure, available airfields, etc. But from the ATC perspective, the primary need is simple: predictability. This has been documented in ATC interviews, knowledge elicitation, and cognitive walkthroughs during research on contingency operations in the Terminal and Enroute domains. Although lost link standards have not yet been defined, it is clear that controllers primarily need to be able to understand exactly when the UAS will maneuver (with sufficient time to issue clearances to traffic), and what exactly the maneuver will be (from a small set of operationally suitable options). Therefore, the sophistication of UAS automation to autonomously maneuver based on its needs might be irrelevant because the operational constraints from the human operators (ATC), to understand what the UAS will do, largely determines UAS behavior in that situation. In this case, human understanding during off-nominal situations has so far been a primary driver for standardizing lost link procedures and technologies.

These aviation examples illustrate that there are practical limits to consider in automated system design from the perspective of human operators, especially in off-nominal conditions. Pilots, air traffic controllers, and other human actors in the NAS are responsible for safety, and held accountable for safety. They might benefit at times from a computer-based team member as viewed in the HAT metaphor, but history suggests the need something simpler, and current research suggests solutions moving forward could be centered on human understanding versus more complex autonomous agents.

Discussion

This paper represents the first part within a series of papers associated with a Panel Session that examines the nature of HAT. Within this paper we have discussed the need for defining the role of the team members and draw on examples in human teaming. We suggest that the same philosophy should be applied to teams that are composed of both human and autonomous members. The dynamic aspects of defining team roles and, perhaps more importantly, who has authority within the team is worth considering. There may very well be instances where the synthetic team member may be in a position to give commands to a human team member. In order for a team to be effective and to respond to changing goals then the acceptance of such authority within the team is a significant factor. It is likely however, that any composition of HAT would still ultimately require a human to supervise and monitor the team. Of course this is dependent on the context of the operation.

The autonomous team member must also be viewed as possesing a degree of intelligence that may not be as tightly *bounded* to rationality as the human team members. When we consider applications in aviation, where HAT could afford measurable benefits, care must be taken. The automation literature in aviation human factors provides us a with a wealth of knowledge as to the benefits and pitfalls of applying complex systems alongside the human. When we look at the potential use of HAT across the flight deck and within ATC there are limitations that need to be considered. On days when the operation is routine and mundane a HAT would function in a manner that would free up resource of the human, but in off-nominal conditions the human will be less likely to understand (and trust) the system as to the rationale behind decisions benig presented. This has implications to safety, acceptance and shaping the future regulatory requirement.

References

Baxter, J., & Richards, D. (2010). Whose Goal is it anyway? User Interaction in an Autonomous System. Twenty-Fourth AAAI Conference on Artificial Intelligence, Atlanta, Georgia, USA, July 11–15, 2010.

Belbin, R.M. (1981) *Management Teams: Why They Succeed or Fail*, Butterworth-Heinemann, Oxford.

Belbin, R.M. (1993) *Team Roles at Work*, Butterworth-Heinemann, Oxford.

Brown, J.P. (2016) The Effect of Automation on Human Factors in Aviation. The Journal of Instrumentation, Automation and Systems, 3(2), 31-46.

Cooke, N.J., Kiekel, P.A. & Helm, E.E. (2001) “Comparing and validating measures of team knowledge”, *Proceedings of the Human Factors and Ergonomics Society*, pp. 361-365.

Cox, M. T. (2013) Goal-driven autonomy and question-based problem recognition. Retrieved from http://www.cogsys.org/papers/2013poster9.pdf

Dekker, S. (2009). *Report of the flight crew human factors investigation conducted for the Dutch safety board into the accident of TK1951, Boeing 737-800 near Amsterdam Schiphol Airport, February 25, 2009* (p. 127). Lund: Lund University, School of Aviation.

Endsley, M.R. (2016) From here to Autonomy: Lessons learned from Human-Automation research. Human Factors: The Journal of the Human Factors and Ergonomics Society, 59(1), 5-27.

Hoffman, R.R. & Woods, D.D. (2011) Simon’s Slice: Five Fundamental Tradeoffs that Boundthe Performance of Human Work Systems, in: 10th International Conference on Naturalistic Decision Making, Orlando FL. pp. 5–31.

Lyons, J. B. (2013). Being transparent about transparency: A model for human-robot interaction. Proceedings of AAAI Spring Symposium on Trust and Autonomous Systems, Palo Alto , 48-53.

Lyons, J.B., & Stokes, C. (2012) Human-Human reliance in the context of automation. Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(1), 112-121.

Richards, D. (2017) Escape from the factory of the robot monsters: Agents of change. Team Performance Management: An International Journal, 23, 96-108.

Salas, E., Cooke, N.J. & Gorman, J.C. (2010) “The science of team performance: progress and the need for more…”, Human Factors, Vol. 52 No. 2, pp. 344-346.

Scholtz, J. (2003) Theory and evaluation of human robot interactions, *Hawaii International Conference on System Science 36 (HICSS 36)*, *Hawaii*, January 2003.

Selkowitz, A.R., Larios, C.N., Lakhmani, S.G. & Chen J.Y.C. (2017) Displaying Information toSupport Transparency for Autonomous Platforms. Advances in Human Factors in Robots andUnmanned Systems. Orlando, FL: Springer International Publishing, 161. doi: 10.1007/978-3-319-41959-6_14.

Sheridan, T. B. & Verplank,W. (1978) Human and Computer Control of Undersea Teleoperators. Cambridge, MA, Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT.

Woods, D.D. & Branlat, M. (2011) How human adaptive systems balance fundamental trade-offs: Implications for polycentric governance architectures, in: Proceedings of the Fourth Resilience Engineering Symposium, Sophia Antipolis, France

Zhou, J., & Chen F. (2015) Making Machine Learning Useable. International Journal of Computational Science and Engineering14(2): 91–109. doi: 10.1504/IJISTA.2015.074069.