Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-19 (2016) 195–200

Shared Mental Models in Human-Machine Systems

Clark Borst *

* *Control and Simulation, Faculty of Aerospace Engineering,*
Delft University of Technology, Kluyverweg 1 2629 HS, The Netherlands
(e-mail: c.borst@tudelft.nl).

Abstract: Imagine a future where humans and machines are able to share tasks, to monitor each other's performance, and to interchange (control) authority whenever required or desired. In aviation, this vision was conceptualized almost twenty-five years ago by the late Charles Billings and is formally known as Human-Centered Automation. Although the aviation community has embraced this perspective, it proves to be difficult to realize this envisioned level of human-machine collaboration, especially for cognitive tasks. To achieve a breakthrough, I argue that we first have to consider what seems to be missing from current forms of automation that is fundamental to effective inter-human collaboration: the possibility to share mental models (or representations) of the problem(s) to solve. When looking at human-human interaction, productive team thinking and problem-solving efforts are accomplished when teammates have a "common ground" or shared understanding of the work to be done and the various ways to do it. Similarly, when work would be distributed over human and automated agents, the constraints introduced by the other agents are properties of the work domain and must therefore be shared. But how can more information be shared while not overloading the human's capacity to learn and solve problems? In this paper I argue that the Ecological Interface Design paradigm can provide the means and guidelines to pursue shared human-automation mental models that will facilitate productive thinking. I will illustrate this by means of an example in the field of aircraft conflict detection and resolution.

Keywords: Man-Machine Systems, Automation, Supervisory Control, Ecological Interface Design.

1. INTRODUCTION

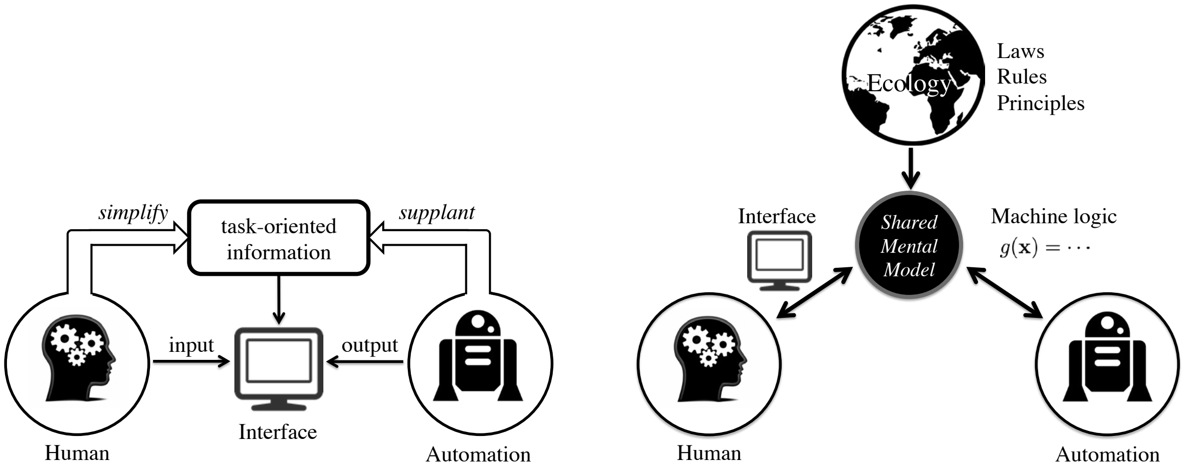
In the (near) future, humans will need to supervise highly automated planes, trains, and automobiles featuring more complex and more intelligent automation. This prospect has given rise to a "human-centered" perspective to automation design in response to broad objections and drawbacks to technology-driven automation (Billings, 1996). The perspective represents a fundamentally new approach to how we view and work with automated systems, namely that automation itself must become a member within the human-machine team. This view is built on the broader concept of human capabilities and limitations, and cautions that there should always remain a role for the human in the loop, to retain such abilities as inductive reasoning and complex pattern matching, which still escape computer design. But how do we create such human-centered automation?

Research communities are beginning to realize that conventional design approaches that are geared toward supplanting and simplifying human involvement are not the right ways to increase the level of automation. To achieve effective human-machine collaboration where both agents can share tasks, monitor each other's performance, and interchange (control) authority, several leading human factors researchers have agreed: more information needs to be shared between human and machine, not less (Norman, 1990; Christoffersen and Woods, 2002; Inagaki, 2008; Flach et al., 1998; Feigh and Pritchett, 2013; Jamieson and Vicente, 2005). But how can more information be shared while not overloading the human's capacity to learn and solve problems?

Few research communities are exploring ways to accomplish this. Most recently, the automotive domain is investigating information sharing between drivers and automatic lane keeping systems on an intuitive and skill-based level by utilizing haptic feedback (Flemisch et al., 2014; Abbink and Mulder, 2009). Similar tactile solutions are explored in aviation for flight envelope protection systems. Haptic-shared control may indeed be an effective way to support humans in skill-based, manual control tasks, such as driving a car or flying an aircraft. But automation is increasingly more capable of taking control over strategic decision-making – the type of knowledge-based work associated with well-trained professionals. However, a framework for sharing information required for such cognitive work does not yet exist. Therefore, we need a way to productively share the principles underlying the control problem and the automation's rationality to keep humans involved and smart on cognitive (knowledge-based) levels. To achieve this, a paradigm shift is needed in the way in which both automation and the human-automation interface are designed in order to create shared mental models.

In this paper I argue that both the design of automation and the design of the interface (i.e., the communication medium) should first and foremost be guided by an analysis of the deep structure (semantics) of the problem domain in order to realize shared mental models (Borst et al., 2015). Thereby, the goal is to ensure that the automation's rationality becomes observable through the interface in ways that facilitate understanding and learning. To formalize this view, I argue that the Cognitive Systems Engineering (CSE) and Ecological Interface Design (EID) frameworks (Vicente and Rasmussen, 1992) can provide the means and guidelines to pursue the development of shared representations and how to communicate them in meaningful ways. I will illustrate this idea by an example shared represen-

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.517



(a) dyadic approach

(b) triadic approach

Fig. 1. Perspectives on human-machine systems. Whereas the dyadic approach puts the emphasis om either the human or the technology, the triadic approach gives priority to capture and visualize the lawful nature of the work domain.

tation for aircraft Conflict Detection and Resolution (CD&R), useful for both pilots and air traffic controllers.

2. FINDING THE RIGHT REPRESENTATION

2.1 Dyadic Approach

The crux in designing a successful human-machine system is to find the *right* representation for the control problem that needs to be solved (Flach et al., 1998). For a long time, determining what right is has been approached from a dyadic perspective, that is, by looking at human and/or machine capabilities (Fig. 1(a)). In this perspective the human, or rather his mind, is viewed as a limited capacity information channel that is a separate entity within the system; an information processor, with inputs, outputs, while the specific rules or logic that describe the information processor vary. When the focus is put on the human user, it is important that the system representation does not exceed the working memory capacity of the operator and that it conforms with the user's expectations or mental model. Of course the danger arises that such a representation may not correspond with the way the system really works and/or it may trivialize a complex control problem by leaving out crucial details that may be important when dealing with off-normal events.

When the emphasis of a system representation is placed on the capabilities of the technology rather than those of the human, it is often assumed that the human operator is the weakest link and should therefore be replaced by a computer of which its capabilities far exceed human capabilities. In this perspective, engineers develop the automation based on a representation of the world that can easily be automated. After that, human factors specialists are typically tasked to design an interface that fits the automation's logic and provide corresponding instructions that help operators use the machinery on rule-based levels so as to support their nominal tasks at acceptable workload levels. The idea is that the human's spare cognitive capacity can be used to complement the system on knowledge-based behavioral levels, required for the type of adaptive and creative decision-making that still escapes computer design. The issue with this approach, however, is that such decision-making requires insight into the complexity (i.e., constraints and relationships) underlying the control problem and the rationality guiding the automation in order to decide when and how to intervene. The task-oriented rules propagated by such interfaces generally provide insufficient information for such behavior. Basic system insights must therefore be gained through extensive training and real-life experiences. This, however, does not guarantee the *right* mental model to be developed required for successful intervention in case the machine reaches its boundaries.

2.2 Triadic Approach

Different from technology- and user-centered approaches, a triadic approach focuses on defining a system relative to its function(s) within a larger work space or ecology (Fig. 1(b)) (Bennett and Flach, 2011). By including and giving priority to this third component in the human-machine system, all cooperating agents need to first and foremost comply with the physical and intentional structure (e.g., laws of physics, safety margins, etc.) of the work domain. This structure is therefore neither dependent on specific technological implementations nor on the capabilities and preferences of the human operator. As such, it would be a good starting point to share the underlying *constraints* of work between humans and automation, before thinking about how to distribute work between agents.

Design frameworks that embrace the triadic perspective are the Cognitive Systems Engineering (CSE) and Ecological Interface Design (EID) paradigms. Whereas CSE puts the emphasis on analysing the deep structure underpinning the work domain through a Cognitive Work Analysis (CWA) (Rasmussen et al., 1994; Vicente, 1999), the goal of the EID framework is to create interface representations that reveal the deep structure of a control problem in meaningful ways for operators to "chunk" information, reducing the demands on memory and supporting productive thinking (e.g., through direct manipulation, metaphorical design, clever geometrical shapes, etc.) (Vicente and Rasmussen, 1992). In its most succinct form, an ecological interface portrays the physical and intentional boundaries for safe control (e.g., go and no-go regions) rather than single optimized solutions that may only work well when