C. Tessier
(Onera)
F. Dehais
(Isae)

E-mail: catherine.tessier@onera.fr

Authority Management and Conflict Solving in Human-Machine Systems

This paper focuses on vehicle-embedded decision autonomy and the human operator's role in so-called autonomous systems. Autonomy control and authority sharing are discussed, and the possible effects of authority conflicts on the human operator's cognition and situation awareness are highlighted. As an illustration, an experiment conducted at ISAE (the French Aeronautical and Space Institute) shows that the occurrence of a conflict leads to a perseveration behavior and attentional tunneling of the operator. Formal methods are discussed to infer such attentional impairment from the monitoring of physiological and behavioral measures and some results are given.

Introduction

There is a growing interest in unmanned vehicles for civilian or military applications, since they prevent the exposure of human operators to hazardous situations. In these domains, autonomy is crucial because the human operator is not embedded within the system [50] and hazardous events may interfere with the human-robot interactions (e.g. communication breakdowns and latencies). The design of authority sharing is therefore critical [29], because conflicts between the robot and the human operator are likely to compromise the mission [38, 52]. Interestingly, these findings are consistent with research in aviation psychology: crew-automation conflicts known as "automation surprises" [40, 41] occur when the autopilot does not behave as expected by the crew (e.g. the autopilot has disconnected and the crew, who is not flying, is not aware of that [35]). These situations can lead to accidents with an airworthy airplane if, despite the presence of auditory warnings [3], the crew persists in solving a minor conflict [4] «instead of switching to another means or a more direct means to accomplish their flight path management goals» [56]. Flight simulator experiments show that in the case of a cognitive conflict with the mission management systems, the human operators' attentional resources are almost exclusively engaged in solving the conflict [17]

to the extent that critical information such as visual or auditory alarms are neglected [18] - a phenomenon known as attentional tunneling [54].

Conflicts in a human-machine system stem from the fact that either the plan for the human operator or the machine is not being followed anymore, or the operator has a faulty awareness of the situation [53], or both. In order to prevent mission degradation, the agents' plans and, if need be, the authority allocation must be adapted, either to fit in with the authority change, or to go against it. This is a real challenge, since in human-machine systems the human agent is hardly controllable and no "model" of the human's decision processes is available. In this paper we will focus on autonomy and the human operator's role in autonomous systems. Then autonomy control will be discussed, before highlighting authority sharing and authority conflicts and discussing the possible effects of such authority conflicts on the human operator's cognition and situation awareness. As an illustration, we will highlight an experiment conducted at ISAE (the French Aeronautical and Space Institute) to show that the occurrence of a conflict leads to a perseveration behavior and attentional tunneling and that such an attentional impairment can be inferred thanks to the monitoring of physiological and behavioral measures.

Autonomy and the human operator's role

In this paper, autonomy stands for decision autonomy, i.e. an "autonomous" agent has the ability to make decisions on its own with embedded situation assessment and decision and planning functions. In ALFUS, methodology autonomy is defined as "a UMS's own ability of sensing, perceiving, analyzing, communicating, planning, decision-making, and acting/executing, to achieve its goals as assigned by its human operator(s) through designed HRI [30]".

While there is no universal definition of autonomy, this concept can be seen as a relational notion between some agents about an object [9, 6]: agent X is autonomous with respect to agent Y about goal G. In a social context, other agents or institutions may influence a given agent, thus affecting its decision making freedom and its behavior [8]. In the context of a robot or software agent in the real world, autonomy can be seen as the ability of the agent to minimize the need of human supervision and to act alone [43]: the primary focus is then the operational aspect of autonomy rather than the social one. In this context, pure autonomy is just a particular case of the artificial agent - human agent relationship, precisely consisting in not using this relationship. However, in practice, human supervision is needed since (1) algorithms can only do what they are designed for and cannot cope with unknown situations [7]; (2) some decisions must be taken by a human being (e.g. in military contexts); (3) the human operator must be able to take over from the algorithms. Moreover, it seems that human interventions significantly improve performance over time compared to a neglected agent [24, 23]: neglect corresponds to communication delays between the human operator and the artificial agent or moments when the human operator is absent or busy with other tasks. Therefore autonomy is still needed to make up for neglect, especially when the operators are far away and communication is not permanent (for security reasons or because of the physics of the system, e.g. the space domain).

In order to take advantage of the complementary skills of the human and artificial agents [31], autonomy variation has been widely considered in the literature:

• Evaluative approaches such as MAP (Mobility, Acquisition and Protection) [26, 11], ACL (Autonomous Control Level) [11] or ALFUS [28] a posteriori assess the autonomy of a robot agent used in a particular mission. Nevertheless, the assessments are either very general and qualitative (ACL), or quantitative with a lack of semantics (e.g. ALFUS final score is an aggregation of three different scores assessing mission complexity, environmental complexity and human independence [30]).

• Prescriptive approaches focus on the design of "autonomous" systems, including human control, with the issues of autonomy levels and how to switch levels. These approaches a several features, i.e. the roles and tasks of each agent within the system, initiative modes, the criteria for autonomy evolution and how the human operator is perceived by the artificial agent. They are dealt with in the next section.

Autonomy variation and autonomy control

Roles and tasks

A role is designed as a set of tasks to be achieved by a given agent [25]. Autonomy-level based prescriptive approaches then specify how the roles should be shared out between the human and the artificial agent. As early as in 1978 Sheridan [47] published a ten-level automation scale for a robotic system: nevertheless, it is an abstract model that does not take into account the environment complexity, or the mission of the robot. Since then, several other scales have been proposed, e.g. [19] for which an autonomy level is characterized by the complexity of the processed controls, [24] where a level represents the capacity of the robot to work independently from the human operator or [5] which claim that the agents' roles vary according to the tasks they must do, they are allowed to do or can do, and the initiative they have to perform them.

The main limits of these approaches are the following:
• at a given autonomy level, the agents' roles cannot evolve;

• there is a limited number of levels, therefore a rigid framework is set and the variety of the situations encountered during a mission cannot be taken into account;
• no rules are given to switch levels;
• the scales are either very general and abstract, or dedicated to a particular system.

Autonomy may also be considered at the task level [46, 21]: an agent is autonomous to achieve a task whenever this task is allocated to it. This approach is more appropriate to deal with the features of a particular mission, nevertheless defining the "best" agent to achieve a given task is an issue in itself.

Initiative modes

Initiative modes are related to the dynamics of the artificial agents' autonomy: which of the artificial agent and the human agent can change the autonomy level of the artificial agent? Three initiative modes are highlighted in the literature: adaptive autonomy, giving the artificial agent exclusive control; adjustable autonomy, giving the human agent exclusive control; and mixed initiative, where the human and artificial agents collaborate to maintain the best perceived level of autonomy [25].

Adaptive autonomy mainly implements the capacity of the artificial agent to ask for the human operator's help, or to self-control. For instance [42] endow robot agents with learning capabilities allowing them to better manage the need for human intervention. Fong's collaborative control [22] is an approach aimed at creating dialogs between the operator and the robot: the robot sends requests to the human operator when problems occur so that these are able to provide the needed support. [10] design agents that can diagnose their own states and self-adapt thanks to predefined behaviors.

The main advantage of adaptive autonomy is that it allows the behaviors of the artificial agent to be well defined for well-identified tasks and situations. Moreover, reactions may be triggered faster than under human control. However, human operators cannot take over from the artificial agent whenever they want, especially when they believe the artificial agent behaves wrongly: their interaction with the artificial agent is restricted to what is expected from them [25].

On the contrary, adjustable autonomy is when only the human operator can control the artificial agent autonomy: the operator may choose the interaction level [19] or "advise" the artificial agent through behavioral or enabling rules [36]. Then, the human operator can analyze the situation, anticipate disruptive events and take over from the artificial agent. The main drawback is that performance may decrease when the operator reacts too slowly or wrongly: the human operator's actions on the artificial agent are not always beneficial [45].

The underlying idea of mixed initiative is to take advantage of the skills of both agents. [46] base task allocation between the robot and the operator on statistics to determine which agent will be the most efficient. This does not guarantee success, because statistics summarize very different situations. However, autonomy tuning at the task level is an interesting idea, since it provides the most adaptive solution to the mission. On a similar principle, [44] build a model allowing artificial agents and human operators to transfer decision making to each other and to compare their decisions. Inconsistencies in the team can be detected so that they can be solved. While the idea of inconsisten-

cies seems to be really relevant in the context of a team of agents, the authors do not say how they should be solved: who should have the priority if the artificial agent and the human operator disagree?

Mixed initiative seems to be the best approach, however it must be tuned properly to show its benefits in practice [25].

Criteria for autonomy evolution

As far as adaptive autonomy and mixed initiative are concerned several operational criteria have been proposed for the artificial agent to change its autonomy. Markov Decision Processes are proposed by [42] and [44], in order for the artificial agent to decide changes, especially to give the authority to the human agent. Furthermore, [44] deal with the possible inconsistencies among the agents' decisions. Such an approach needs to be able to compute the utility of each strategy. The criteria of [25] are explicit, since autonomy changes are triggered by predefined events (i.e. some human operator's actions, some mission events). Those criteria are objective, however they are very mission- and task-dependent. As far as the work of [36] is concerned, the behavior of the robot agent changes according to predefined rules; however, the potential conflicts between rules are not discussed.

Various criteria and metrics have been proposed in the literature to trigger autonomy changes in the artificial agent. Though they are grounded on objective mission features, they are generally very mission-dependent and therefore not re-usable in other contexts.

How is the human operator perceived by the artificial agent?

Generally speaking and even if it is likely to be erroneous, the human operator has knowledge of the capacities of the artificial agent and of its current and possible future states (situation awareness [20]). Conversely, when adaptive autonomy or mixed initiative are considered, the artificial agent should have a model of the human operator's "capacities" and "state". This is hardly the case in the literature, since the operator is often considered as an infallible resort. Some examples however can be found [31, 21] where the robot has models of the tasks the operator can perform: therefore, it can plan for itself and for the operator and track the operator's task execution.

More recent research [1, 12] considers some data from the operator (such as physiological data, eye tracking, expertise, workload) to take part in the reasoning process of the artificial agent, thus allowing it to adapt its autonomy when an "impaired state" of the operator is diagnosed. We will focus on that approach in the rest of the paper.

From autonomy to authority

Joining human and machine abilities aims at increasing the range of actions of "autonomous" systems. However, the relationship between both agents is dissymmetric, since the human operator's "failures" are often neglected when designing the system. Moreover, simultaneous decisions and actions of the artificial and the human agents are likely to create conflicts [16]: unexpected or misunderstood authority changes may lead to inefficient, dangerous or catastrophic situations. Therefore, in order to consider the human agent and the artificial agent in the same way [27] and the human-machine system as a whole [55], it seems more relevant to work on authority and authority control than on autonomy, which concerns the artificial agent exclusively.

Authority sharing and authority conflicts

One of the main issues in human-machine systems is to prevent the whole system from deteriorating and reaching undesired and possible dangerous states; this includes on-line failure detection and recovery, the maintenance of the operator's situation awareness and correct interaction with the artificial agent, as well as authority conflict detection and solving.

A change in authority allocation can be planned in the procedures or in the mission plan, or can be unexpected: this happens when the human operator takes over a task controlled by the artificial agent (software or robot) because they detect a failure, or for any reason of their own; or when the artificial agent takes over a task controlled by the operator because the operator's action violates some constraints (e.g. a potential excursion out of the flight domain), or because the communication with the operator is impaired; or when no agent has the authority anymore [35]. Therefore authority has to be formalized in order to identify those situations so as potential authority conflicts.

Authority: some definitions

An agent X has authority over a resource R of a system with respect to another agent Y [34] if X can control R to the detriment of Y. The control of X on R can be more or less strong against Y according to the following modes:
- access: agent X can use resource R in order to achieve a goal;
- pre-emptability: agent X can use resource R as soon as needed, taking it from agent Y if Y is already controlling R;
- control guarantee: once agent X controls R, agent Y will not be able to take R away from X through pre-emption.

Consequently authority is characterized by the following properties:
- a gradation of the agent's authority: agent X's control on resource R gets stronger as it is granted access, pre-emptability and control guarantee, in this order;
- authority, as autonomy [9] is a relative concept: agent X may have pre-emptability on R over agent Y, but not over agent Z. Consequently, there are as many authority relationships as there are couples of agents that may control R;
- authority is shared between the agents: for a couple of agents <X,Y> that may control resource R, the authority gain of agent X on resource R corresponds to an authority loss for agent Y. For instance, if agent X obtains the control guarantee on R, this means agent Y loses pre-emptability. Consequently, agent Y will not have access to R anymore: agent X prevents agent Y from accessing resource R, even if it does not use it.

The Petri net in figure 1 represents the authority relationship between two agents, X et Y, for a given resource R: each place corresponds to the State of agent X / State of agent Y regarding resource R. The state changes modify the status of R, i.e. they determine whether R can be allocated to X or Y, or not.

There are two intermediate states for which the agents' authority is equivalent, namely (Access / Access) and (Pre-emptability / Pre-emptability). As far as the first one is concerned, the agents cannot take the resource control from one another, each must wait for the other one to release the resource. This is a cooperation context. As far as the second one is concerned, the agents can take the resource control from one another indefinitely, which makes the behavior of the system inefficient or even dangerous. This is a competition context.
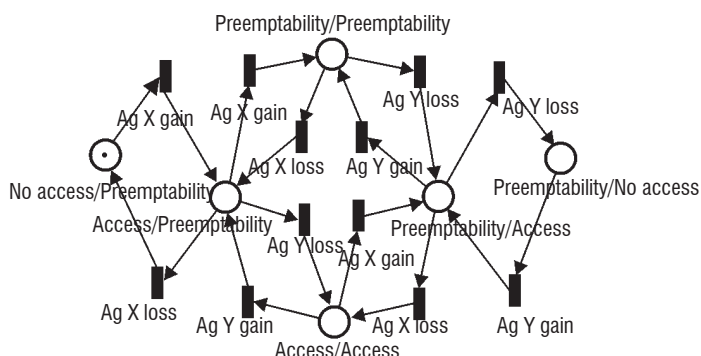

Figure 1 - Authority relationship between two agents X and Y on a given resource R

Authority conflicts

A conflict is a state of the world where one or several agents cannot achieve their goals: agent X is in conflict with agent Y if one of Y's goals prevents X from achieving one of its goals. Except for state (Access / Access), which corresponds to both agents having the lowest authority over resource R, all other states of the authority relationship are potential conflicts, since one agent has more authority than the other over R, or both agents have Pre-emptability (competition for R). Authority conflicts between the human operator and "autonomous" systems are often linked to "automation surprise" [41]: either the plan for both the human and robot is not followed anymore, or the operator has a faulty awareness of the situation [53], or both.

Experiments conducted in flight simulators reveal that the occurrence of such conflicts in mission management systems [17] leads to summoning up most of the human operator's capacities toward conflict solving. As a consequence, the operator's cognitive abilities are impaired with a strong tendency to attentional tunneling [54], where critical information, such as visual or audio alarms [18], is neglected. Because this critical information is not perceived, the human operator's situation awareness is degraded, which may lead to a dangerous vicious circle (see figure 2).
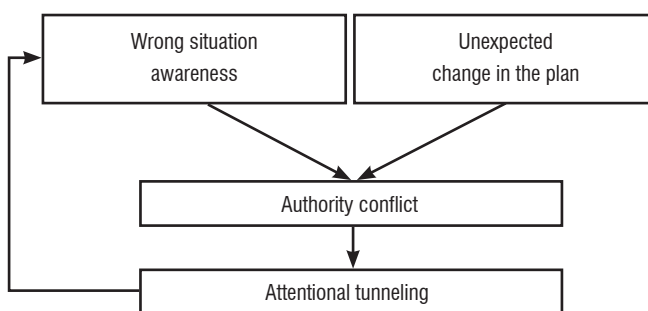

Figure 2 - The vicious circle of authority conflicts

Conflict detection and solving

Conflict detection and solving in human-machine systems involve (figure 3):
- an estimation of the state of the whole human-robot system, i.e. an estimation of the state of the robot, of the «state» of the human operator and of the state of the interaction between the two; conflicts

Box 3 - Cognitive countermeasures

Attentional tunneling is a paradox for interface designers: how can one expect to "cure" human operators from attentional tunneling if the alarms or systems designed to warn them are neglected? Experiments conducted in flight simulators have shown that the absence of response to either auditory or visual alarms may be explained by an inability to disengage attention: the warning systems are based on providing the operator with additional information, but this is of little use if the warning system is not also efficient in disengaging attention from the current task [17]. By contrast, cognitive countermeasures are based on the temporary removal of information on which the human operator is focusing, for it to be replaced by an explicit visual stimulus in order to change the operator's attentional focus. The user interface acts as a cognitive prosthesis as it performs the attentional disengagement and attentional shifting.

correspond to inconsistent or unwanted states of the human-robot system;
   • conflict solving thanks to an adaptation of the human-machine system, i.e. re-planning, changes in authority sharing [34, 33], cognitive countermeasure sending [18].
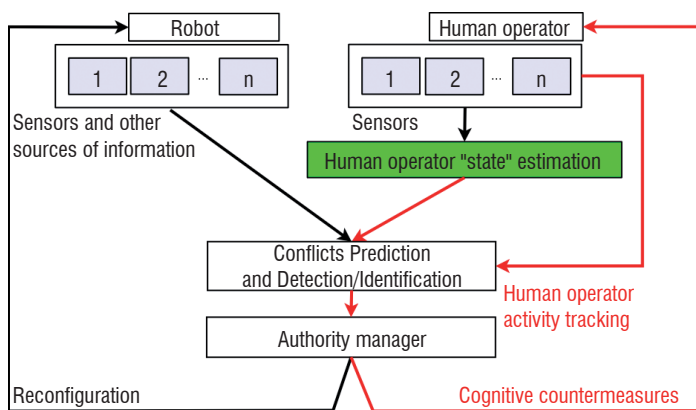

Figure 3 - Conflict solving and the role of the human operator "state" estimation

State estimation

State estimation results from the matching between measures and models of the expected behaviors of the human-machine system. Conflict markers are inconsistent states (e.g. a dead marking in a Petri net representing the operation of the system), and/or measures matching a model of an unwanted behavior (e.g. the operator's attentional tunneling).

Several kinds of measures are required to estimate the state of the human-machine system: some measures linked to the state of the physical system (the machine), e.g. the configuration, the speed, the alarms, etc.; some measures linked to the "state" of the human operator; and some measures linked to the interaction between the human and the machine, e.g. the operator's actions on the machine interfaces. The analysis of these different state vector estimations allows the current and future situations to be assessed in order to adapt the tasks and the authority of both the human and artificial agents.

The estimation of the «state» of the human operator is a challenge. Indeed both data and models for recognizing special «states» are needed. As far as attentional tunneling is concerned, the human operator's excessive focus is associated with a decreased saccadic activity and long concentrated eye fixations [13] and consequently less scanned areas of interest on the user interface [48]. The heart rate also confirms that the catabolic activity increases. Therefore models of those phenomena have to be built or learnt [32] to further characterize unwanted "states" of the operator.

Conflict solving

Conflict solving consists in adapting the behavior of the machine and the information sent to the human operator, at least for a while. This may involve:
   • action re-planning and/or resource reallocation within the system (e.g. in case of a failure), possibly with goal changes (e.g. land on the nearest emergency landing ground);
   • changes in some authority relationships on some resources of the human-machine system [33] (e.g. automatic protection of the flight domain: the autopilot can take over from the crew);
   • cognitive countermeasure sending to the human operator, through the HMI [17, 18, 14] (e.g. in case of attentional tunneling in the human operator).

An experiment and some results

This section is focused on an experiment and first results that we have obtained for conflict identification and solving. More details can be found in [39, 14 and 15].

Experimental framework

Experiments have been conducted at ISAE on a target search mission achieved by a ground robot and a remote human operator [33, 15]. The human operator is equipped with an eye-tracker and with an electrocardiogram device. The robot is equipped with decision functions that allow it to navigate while avoiding obstacles, to detect targets and to adapt its behavior when some disruptive events occur. Information is available on the HMI (see Figure 4) for the human operator to supervise the robot and take over if necessary. As soon as a target is detected by the robot, the operator must take over and operate the robot so as to identify the target precisely.

While the operator takes over the robot for target identification, a battery failure triggered by the experimenter activates a safety procedure that makes the robot go back to base autonomously. This event is presented on the operator's interface via three alarms: the battery icon (Area 7) switches from green to orange, the piloting mode (Area 3) blinks twice from «manual» to «supervised», and the display (Area 5) shows «Back to base» in green. However, since this unexpected event occurs at a critical time in the mission when the operator is particularly focused on the panoramic vision
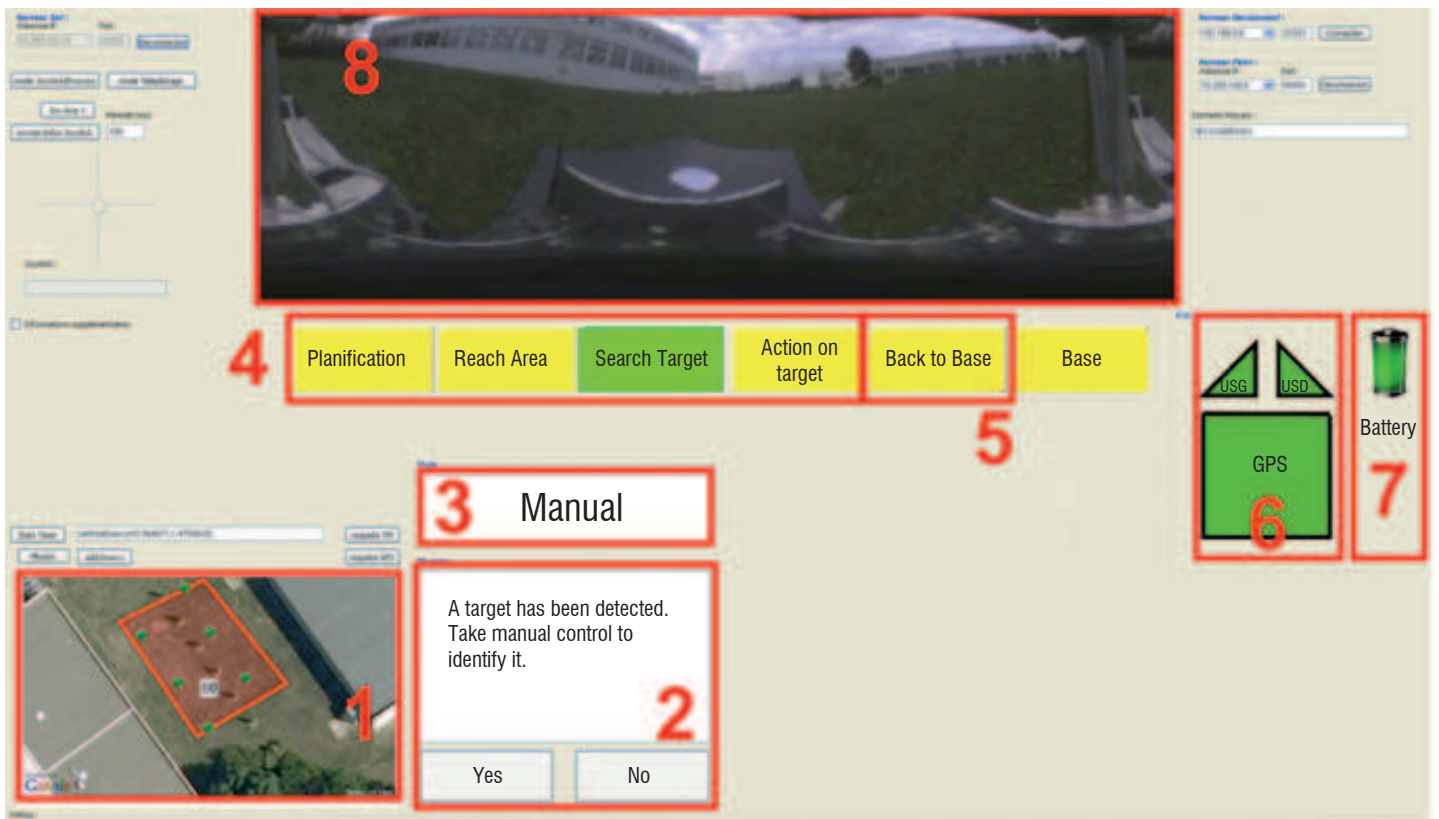
Figure 4 - The human-machine interface

Video 1: despite the alarms, the operator's gaze (red cross) is fixed on the panoramic vision display. http://www.aerospacelab-journal.org/al4/authority-management-and-conflict-solving

Video 2: the countermeasure disengages the operator's visual focus (red cross) from the panoramic vision and leads him to glance at the battery icon within 1 second. http://www.aerospacelab-journal.org/al4/authority-management-and-conflict-solving

display (Area 8) for the identification task in manual mode, the results show that 9 participants out of 13 faced attentional tunneling and were not aware of the automation logic: the eye-tracking data showed that the relevant information (Areas 3, 5 and 7) had not been seen (see video 1). Consequently, each agent (the robot and the operator) persisted in fulfilling their goals, i.e. going back to base and identifying targets respectively. This is an authority conflict concerning the control of resource "robot navigation" with a (Pre-emptability / Pre-emptability) competitive state of both agents (figure 1).

Conflict solving through countermeasures

A second experiment was conducted with 11 subjects. The goal was to test empirically one of the possible actions to solve a conflict,

i.e. a planned interaction with the operator aimed at modifying the operator's behavior: this cognitive countermeasure consists in removing the panoramic vision display (Area 8) where the operator is focused, and replacing it during 4 seconds with the message "Battery failure, robot returning to base" (see video 2). What is expected is that the operator will drop their target identification task and switch their attention towards the relevant areas of the interface.

The results show that 9 participants out of 11 immediately gave the control back to the robot to let it return to base autonomously. The remaining two participants claimed that they had understood the conflict with the countermeasure, but thought that they had enough time to operate the robot before the total discharge of the battery.

Characterization of attentional tunneling

Characterizing attentional tunneling is necessary to design models (figure 5 left) so as to be able to automatically detect this phenomenon on-line (figure 5 right) and trigger relevant actions to compensate for it.
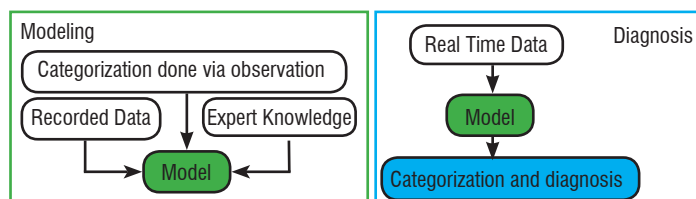

Figure 5: The model is designed (left) to be used on-line (right)

Thanks to signal processing of the data collected during the experiments (i.e. the operator's gaze position and heart rate) and fuzzy aggregation rules, first characterizations of attention tunneling have been obtained. The measures - percentage of time spent on the video (Area 8), number of areas of interest (AOIs) scanned in a defined time interval, number of changes of AOIs in a defined time interval; heart rate and heart rate standard deviation - are aggregated thanks to expert rules so as to derive the Focus and the Cardiac Stress which are in turn aggregated to derive the Attentional Tunneling - see [39] for more details.

As an example, figures 6 and 7 show the results for two subjects previously labeled by the experimenters as "Attentional tunneling" and "OK, conflict perceived" respectively.

Time references on the x-axis (graduated in seconds) are: P1: start of phase "research area"; P2: start of phase "search target"; P3: start of phase "identify target" (i.e. manual piloting); P4: failing battery alarms, piloting mode «supervised», start of the conflict; P5 (if present): observed end of the conflict. On the time axis the alert level is represented by a three-color code (red, yellow and green).

As we can see for case A, the alert level goes from low (green) to high (red) during manual piloting. The alert level is stable on high for the rest of the mission. This is in accordance with the observed behavior: the subject faced attentional tunneling and did not understand the conflict.
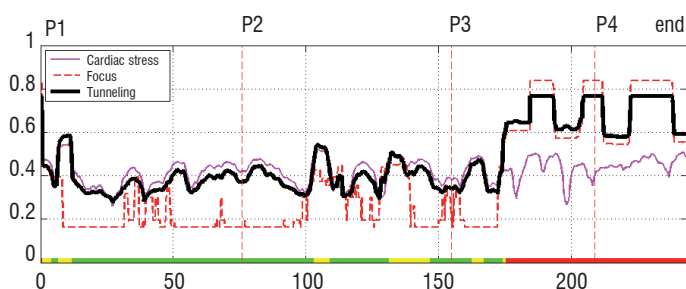

Figure 6: Case A - Attentional tunneling

As for case B, the alert level goes from low to high during manual piloting. After the start of the conflict the alert level is stable on high. After the end of the conflict the alert level goes from high to medium (yellow) in about 5 seconds and from medium to low in about 5 more seconds. In this case also the calculated behavior is in accordance with the observed behavior.
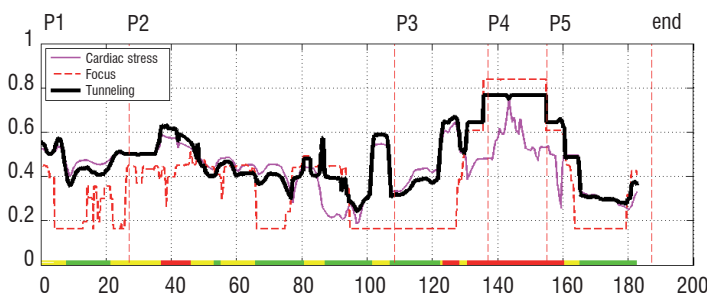

Figure 7: Case B - OK, conflict perceived

Conclusions and further work

The main drawback of the concept of variable autonomy – though widely studied in the literature – is that the human operator is not placed on the same plane as the machine (a robot or a software agent): the human operator is often considered as an infallible resort within the human-machine system. On the other hand, the concept of authority allows symmetric roles to be considered: the authority on a given resource can be transferred from one agent to the other according to the context. Furthermore, the concept of conflict allows degraded situations within the whole human-machine system to be detected, provided measures that are relevant to identify or predict unwanted behaviors are available. Therefore, some models of the operator's specific behaviors must be designed. We have shown that model building from experimental data gives promising results as far as attentional tunneling is concerned.

The main challenge in human-machine mixed initiative systems, such as robots or aircraft, is to avoid conflicting situations, i.e. situations where the operator and the decision algorithms "do not understand each other" and attempt to keep their authority on some resources of the system. Further work must focus on the closed loop involving on-line conflict detection – thanks to further investigation of the "human" metrics and of the correlation of the "human" and "machine" metrics, and the design of robust models of degraded human behaviors; and on on-line conflict solving through authority dynamic management, so as to allocate authority to the most capable agent in the current context. This involves further issues, such as agents' cohesion, the maintenance of the human operator's situation awareness and the operator's acceptance. Solutions allowing the operator's actions to be influenced without disturbing them (e.g. "subliminal" guidance, actions on the operator's situation awareness using countermeasures, etc.) must be further investigated ■

References

[1] D. BARBER, L. DAVIS, D. NICHOLSON - *The Mixed Initiative Experimental (MIX) Test-Bed for Human Robot Interactions with Varied Levels of Automation*. 26th Army Science Conference, 2008.

[2] J. BEAUVOIS, R. JOULE - *A Radical Point of View on Dissonance Theory, in Cognitive Dissonance : Progress on a Pivotal Theory in Social Psychology*. E. Harmond-Jones and J. Mills, eds., Washington : American Psychological Association, 1999, pp. 43–70.

[3] D. BERINGER, H. HARRIS JR - *Automation in General Aviation: Two Studies of Pilot Responses to Autopilot Malfunctions.* The International Journal of Aviation Psychology, 9 (1999), pp. 155–174.

[4] E. BILLINGS - *Aviation Automation : the Search for a Human-Centered Approach*. Lawrence Erlbaum associates, Inc., Mahwah, NJ, USA, 1996.

[5] J. BRADSHAW, M. SIERHUIS, A. ACQUISTI, R. FELTOVICH, R. HOFFMAN, R. JEFFERS, D. PRESCOTT, N. SURI, A. USZOK, R. VAN HOOF - *Adjustable Autonomy and Human-Agent Teamwork in Practice: An Interim Report on Space Application.* Agent Autonomy. H. Hexmoor, C. Castelfranchi, and R. Falcone, eds., Kluwer Academic Publishers, 2003, pp. 243–280.

[6] S. BRAINOV, H. HEXMOOR - *Quantifying Relative Autonomy in Multiagent Interaction.* Agent Autonomy, H. Hexmoor, C. Castelfranchi, and R. Falcone, eds., Kluwer Academic Publishers, 2003, pp. 55–73.

[7] J. BROOKSHIRE, S. SINGH, R. SIMMONS - *Preliminary Results in Sliding Autonomy for Coordinated Teams*. AAAI04 Spring Symposium Series, Stanford, CA, USA, 2004.

[8] C. CARABELEA, O. BOISSIER - *Coordinating Agents in Organizations Using Social Commitments*. Electronic Notes in Theoretical Computer Science, 150 (2006).

[9] C. CASTELFRANCHI, R. FALCONE - *From Automaticity to Autonomy: the Frontier of Artificial Agents*. Agent Autonomy, H. Hexmoor, C. Castelfranchi, and R. Falcone, eds., Kluwer Academic Publishers, 2003, pp. 103–136.

[10] C. CHOPINAUD, A. EL FALLAH-SEGHROUCHNI, P. TAILLIBERT - *Prevention of Harmful Behaviors within Cognitive and Autonomous Agents*. ECAI 2006: 17th European Conference on Artificial Intelligence, Riva del Garda, Italy, 2006.

[11] B. T. CLOUGH - *Metrics, Schmetrics! How the Heck do you Determine a UAV's Autonomy anyway?* Performance Metrics for Intelligent Systems Workshop, Gaithersburg, MD, USA, 2002.

[12] G. COPPIN, F. LEGRAS, S. SAGET - *Supervision of Autonomous Vehicles: Mutual Modelling and Interaction Management.* HCI International 2009, San Diego, CA, USA, 2009.

[13] L. COWEN, L. BALL, J. DELIN - *An Eyemovement Analysis of Web-Page Usability, in Computers XVI: Memorable Yet Invisible*. Proceedings of HCI'02, London, UK, 2002.

[14] F. DEHAIS, M. CAUSSE, S. TREMBLAY - *Mitigation of Conflicts with Automation*. Human Factors, 5 (2011), pp. 448–460.

[15] F. DEHAIS, M. CAUSSE, F. VACHON, S. TREMBLAY - *Cognitive Conflict in Human-Automation Interactions: A Psychophysiological Study.* Applied Ergonomics, 43 (2012), pp. 588–595.

[16] F. DEHAIS, A. GOUDOU, C. LESIRE, C. TESSIER - *Towards an Anticipatory Agent to Help Pilots*. AAAI'05 Faill Symposium «From reactive to anticipatory cognitive embodied systems», Arlington, VA, USA, 2005.

[17 F. DEHAIS, C. TESSIER, L. CHAUDRON - *Ghost: Experimenting Conflicts Countermeasures in the Pilot's Activity*. IJCAI'03, 18th International Joint Conference on Artificial Intelligence, Acapulco, Mexico, 2003.

[18] F. DEHAIS, C. TESSIER, L. CHRISTOPHE, F. REUZEAU - *The Perseveration Syndrome in the Pilot's Activity: Guidelines and Cognitive Countermeasures*. 7th International Working Conference on Human Error, Safety, and System Development (HESSD), Brussels, Belgium, 2009.

[19] G. DORAIS, P. BONASSO, D. KORTENKAMP, B. PELL, D. SCHRECKENGHOST - *Adjustable Autonomy for Human-Centered Autonomous Systems*. IJCAI'99 Workshop on Adjustable Autonomy Systems, Stockholm, Sweden, 1999.

[20] M. ENDSLEY - *Toward a Theory of Situation Awareness in Dynamic Systems*. Human Factors, 37 (1995), pp. 32–64.

[21] A. FINZI, A. ORLANDINI - *Human-Robot Interaction Through Mixed-Initiative Planning for Rescue and Search Rovers*. Lectures Notes in Computer Science - AIIA'05, Milan, Italy, 2005, pp. 483–494.

[22] T. FONG, C. THORPE, C. BAUR - *Collaboration, Dialogue and Human-Robot Interaction*. 10th International Symposium on Robotics Research, Lorne, Victoria, Australia, 2002.

[23] M. GOODRICH, T. MCLAIN, J. CRANDALL, J. ANDERSON, J. SUN - *Managing Autonomy in Robot Teams: Observations from Four Experiments.* ACM/IEEE International Conference on Human-robot interaction, Washington, DC, USA, 2007.

[24] M. GOODRICH, D. OLSEN, J. CRANDALL, T. PALMER - *Experiments in Adjustable Autonomy.* Ijcai'01 Workshop on Autonomy, Delegation and Control: Interacting with autonomous agents, Seattle, WA, USA, 2001.

[25] B. HARDIN, M. GOODRICH - *On Using Mixed-Initiative Control: a Perspective for Managing Large-Scale Robotic Teams*. HRI '09 - Proceedings of the 4th ACM/IEEE international conference on Human Robot Interaction, San Diego, CA, USA, 2009.

[26] B. HASSLACHER, M. W. TILDEN - *Living Machines.* Robotics and Autonomous Systems, 15 (1995), pp. 143–169.

[27] E. HOLLNAGEL, ed. - Handbook of Cognitive Task Design. Mahwah, NJ: Erlbaum, 2003.

[28] H. HUANG, K. PAVEK, B. NOVAK, J. ALBUS, E. MESSIN - *A framework for Autonomy Levels for Unmanned Systems ALFUS*. AUVSI's Unmanned Systems North America 2005, Baltimore, MD, USA, 2005.

[29] T. INAGAKI - *Automation and the Cost of Authority*. International Journal of Industrial Ergonomics, 31 (2003), pp. 169–174.

[30] J. JONES - *ALFUS - Autonomy Levels for Unmanned Systems.* Framework. Slides: http://sstc-online.org/2008/pdfs/JBJ2079.pdf, 2008.

[31] D. KORTENKAMP, P. BONASSO, D. RYAN, D. SCHRECKENGHOST - *Traded Control with Autonomous Robots as Mixed Initiative Interaction*. AAAI-97 Spring Symposium on Mixed Initiative Interaction, Stanford University, CA, USA, 1997.

[32] R. MANDRYK, M. ATKINS - *A fuzzy Physiological Approach for Continuously Modelling Emotion During Interaction with Play Environments.* International Journal of Human-Computer Studies, 6 (2007), pp. 329–347.

[33] S. MERCIER, C. TESSIER, F. DEHAIS - *Authority Management in Human-Robot Systems*. IFAC/IFIP/IFORS/IE Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Valenciennes, France, 2010.

[34] S. MERCIER, C. TESSIER, F. DEHAIS - *Détection et résolution de conflits d'autorité dans un système homme-robot*. Revue d'Intelligence Artificielle, numéro spécial «Droits et Devoirs d'Agents Autonomes», 24 (2010), pp. 325–356.

[35] R. MUMAW, N. SARTER, C. WICKENS - *Analysis of Pilots' Monitoring and Performance on an Automated Flight Deck*. Proceedings of the 11th International Symposium for Aviation Psychology, Columbus, OH, USA, 2001.

[36] K. MYERS, D. MORLEY - *Human Directability of Agents*. K-CAP 2001, 1st International Conference on Knowledge Capture, Victoria, Canada, 2001.

[37] J. ORASANU, L. MARTIN, J. DAVISON - *Cognitive and Contextual Factors in Aviation Accidents: Decision Errors*. Linking expertise and naturalistic decision making, E. Salas and G. Klein, eds., Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2001, ch. 12, pp. 209–225.

[38] R. PARASURAMAN, C. WICKENS - *Humans: Still Vital After all These Years of Automation.* Human factors, 50 (2008), pp. 511–520.

[39] S. PIZZIOL, F. DEHAIS, C. TESSIER - *Towards Human Operator «State» Assessment.* ATACCS'2011 - 1st International Conference on Application and Theory of Automation in Command and Control Systems, Barcelona, Spain, 2011.

[40] N. SARTER, D. WOODS - *How in the World did we ever get into that Mode? Mode Error and Awareness in Supervisory Control.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 37 (1995), pp. 5–19.

[41] N. SARTER, D. WOODS, C. BILLINGS - *Automation Surprises.* Handbook of Human Factors and Ergonomics (2nd edition), G. Salvendy, ed., New York, NY: Wiley, 1997, pp. 1926–1943.

[42] P. SCERRI, D. PYNADATH, M. TAMBE - *Adjustable Autonomy for the Real World*. Agent Autonomy, H. Hexmoor, C. Castelfranchi, and R. Falcone, eds., Kluwer Academic Publishers, 2003, pp. 211–241.

[43] D. SCHRECKENGHOST, D. RYAN, C. THRONESBERY, P. BONASSO, D. POIROT - *Intelligent Control of Life Support Systems for Space Habitat*. Proceedings of the AAAI-IAAI Conference, Madison, WI, USA, 1998.

[44] N. SCHURR, J. MARECKI, M. TAMBE - *Improving Adjustable Autonomy Strategies for Time-Critical Domains.* AAMAS'09, Budapest, Hungary, 2009.

[45] N. SCHURR, P. PATIL, F. PIGHIN, M. TAMBE - *Using Multiagent Teams to Improve the Training of Incident Commanders.* AAMAS'06 Industry Track, Hakodate, Japan, 2006.

[46] B. SELLNER, F. HEGER, L. HIATT, R. SIMMONS, S. SINGH - *Coordinated Multi-Agent Teams and Sliding Autonomy for Large-Scale Assembly*. Proceedings of the IEEE - Special Issue on Multi-Robot Systems, 94 (2006), pp. 1425 – 1444.

[47] T. SHERIDAN, W. VERPLANK - *Human and Computer Control of Undersea Teleoperators.* tech. rep., MIT Man-Machine Systems Laboratory, Cambridge, MA, USA, 1978.

[48] L. THOMAS, C. WICKENS - *Eye-Tracking and Individual Differences in off-Normal Event Detection when Flying with a Synthetic Vision System Display.* 48th Conference of the Human Factors and Ergonomics Society, Santa Monica, CA, USA, 2004.

[49] W. TRUSZKOWSKI, H. HALLOCK, C. ROUFF, J. KARLIN, J. RASH, M. HINCHEY, R. STERRITT - *Autonomous and Autonomic Systems: With Applications to NASA Intelligent Spacecraft Operations and Exploration Systems.* NASA Monographs in Systems and Software Engineering, Springer, 2010.

[50] A. TVARYANAS - *Visual Scan Patterns During Simulated Control of an Uninhabited Aerial Vehicle (UAV)*. Aviation, space, and environmental medicine, 75 (2004), pp. 531–538.

[51] B. VAN DER KOLK - *The Body Keeps the Score: Memory and the Evolving Psychobiology of Post Traumatic Stress*. Harvard Review of Psychiatry, 1 (1994), pp. 253–265.

[52] H. VAN GINKEL, M. DE VRIES, J. KOENERS, E. THEUNISSEN - *Flexible Authority Allocation in Unmanned Aerial Vehicles*. Conference Proceedings of the Human Factors and Ergonomics Society, San Francisco, CA, USA, 2006.

[53] C. WICKENS - *Situation Awareness: Review of Mica Endsley's 1995 Articles on Situation Awareness Theory and Measurement.* Human factors, 50 (2008), pp. 397–403.

[54] C. WICKENS, A. ALEXANDER - *Attentional Tunneling and Task Management in Synthetic Vision Displays*. International Journal of Aviation Psychology, 19 (2009), pp. 182–199.

[55] D. WOODS, E. ROTH, K. BENNETT - *Explorations in Joint Human-Machine Cognitive Systems*. Cognition, computing, and cooperation, S. Robertson, W. Zachary, and J. Black, eds., Ablex Publishing Corp. Norwood, NJ, USA, 1990, pp. 123–158.

[56] D. WOODS, N. SARTER - *Learning from Automation Surprises and Going Sour Accidents.* Cognitive engineering in the aviation domain, N. Sarter and R. Amalberti, eds., Lawrence Erlbaum, New York, 2000, pp. 327–353.

Acronyms:

ALFUS (Autonomy Levels For Unmanned Systems Framework)
AOI (Area Of Interest)
BEA (the French national institute for air accident analysis)
ISAE (the French Aeronautical and Space Institute)
UAV (Unmanned Aerial Vehicle)

UGV (Unmanned Ground Vehicle)
UMS (UnManned System)
HMI (Human Machine Interface)
HRI (Human Robot Interaction)

AUTHORS

Catherine Tessier is a researcher at Onera, the French Aerospace Lab, in Toulouse, France. She received her PhD in 1988 and her HDR (Habilitation à Diriger les Recherches) in 1999. She is also a part-time Professor at ISAE. Her research areas include cooperating agents, authority sharing, adaptive autonomy, situation assessment and tracking and ethical issues, mainly in projects involving uninhabited vehicles. She is a member of the boards of the EDSYS Doctoral School, of the French National Research Group (GDR) for Robotics and of the French Association for Artificial Intelligence (AFIA).

Frédéric Dehais is an associate professor at Isae (Institut Supérieur de l'Aéronautique et de l'Espace). He defended his PhD in 2004 at Onera (the French Aerospace Lab) on the topic of modelling cognitive conflict in pilot's activity. He then held a 2-year post-doctoral position funded by Airbus to apply the research developed during his PhD. Since 2006, he has been leading the human factor team at Isae and working on different academic and industrial projects for providing real time assistance to human operators.