Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is an author's version published in: https://oatao.univ-toulouse.fr/26588

Official URL: https://doi.org/10.1007/978-3-030-49183-3_10

To cite this version :

Ponzoni Carvalho Chanel, Caroline and Roy, Raphaëlle N. and Dehais, Frédéric and Drougard, Nicolas Mixed-Initiative Human-Automated Agents Teaming: Towards a Flexible Cooperation Framework. (2020) In: 22nd International Conference on Human-Computer Interaction - HCI INTERNATIONAL 2020, 19 July 2020 - 24 July 2020 (Copenhagen, Denmark).

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

Mixed-Initiative Human-Automated Agents Teaming: Towards a Flexible Cooperation Framework

Caroline P. C. Chanel, Raphaëlle N. Roy,
Nicolas Drougard, and Frédéric Dehais

ISAE-SUPAERO, Université de Toulouse, France
name.lastname@isae-supaero.fr

Abstract

The recent progress in robotics and artificial intelligence raises the question of the efficient artificial agents interaction with humans. For instance, artificial intelligence has achieved technical advances in perception and decision making in several domains ranging from games to a variety of operational situations, (e.g. face recognition [51] and firefighting missions [23]). Such advanced automated systems still depend on human operators as far as complex tactical, legal or ethical decisions are concerned. Usually the human is considered as an ideal agent, that is able to take control in case of automated (artificial) agent's limit range of action or even failure (e.g embedded sensor failures or low confidence in identification tasks). However, this approach needs to be revised as revealed by several critical industrial events (e.g. aviation and nuclear powerplant) that were due to conflicts between humans and complex automated system [13]. In this context, this paper reviews some of our previous works related to human-automated agents interaction driving systems. More specifically, a mixed-initiative cooperation framework that considers agents' non-deterministic actions effects and inaccuracies about the human operator state estimation. This framework has demonstrated convincing results being a promising venue for enhancing human-automated agent(s) teaming.

Keywords: Mixed-initiative interaction, Shared-autonomy, Human operator monitoring, Sequential decision making under uncertainty and partial observability, POMDP.

1 Introduction

Recent advances in coupling robotics and artificial intelligence have generated a large number of applications of multi-UAV or multi-robot systems to deal with 3D (Dirty, Dull, Dangerous) missions in industry and in research. Surveillance [40, 39], search and rescue missions [32, 48], exploration [5], or inspection missions [30] are a few examples of the potential power of those automated systems. Thereby, these automated systems still rely on human agents for the tricky tactical, ethical and moral decisions. Usually, the human is considered as an ideal agent in charge of taking over when the automated agent fails [44], or when the artificial agent is not able/suitable to make the decision. However, the capacity of human agents to take over when the automated system fails can be strongly affected, for instance, by a poor user interface design, the complexity of automated agents operation, a high operational pressure, or emotional commitment. Those factors could diminish the human operator performance and judgment during interaction leading to the need of careful design of authority sharing [12].

Indeed, the question of human-automated agent(s) interaction has mainly been studied in the literature through the concept of autonomy levels [24, 36]. Sliding autonomy design [4, 17], adaptive automation [43, 45] or the shared autonomy design [26] are examples of this concept. Related with the latter, the mixed-initiative (MI) framework is particularly interesting because it establishes a coordination strategy that defines the role of the humans and the automated agents according to their recognized skills [1, 2], considering them as teammates rather than master-slave agents. The main idea is that each agent might seize the initiative - it chooses to contribute to the task that it does best. That is to say if an agent (human or automated agent) controls the interaction, the other agents work to assist him [2, 27]. In a more general framework, tasks should not be determined in advance, but rather they should be negotiated between agents [6].

Following the classical MI concept [2], the recognized skills of the involved agents may define their role during mission execution. In our view, the role of a given agent (e.g. the tasks to perform) would not only be defined by its skills. A special attention should also be given to the current capabilities of such an agent. Artificial agents such as mobile robots, autonomous cars, or aircraft still have limitations concerning their embedded sensors in real-life environments (e.g. limited camera field of view, or bias in image processing algorithms, poor GPS signal confidence) that potentially limit their performance. It may bring, for instance, low confidence when they are performing identification tasks [20, 48, 30] or execution errors during autonomous navigation [16]. Similarly, human agents that interact with such automated agents may also be confronted to difficulties, such as physical and/or cognitive limitations. This crucial point is addressed in the following section.

1.1 Human skills *versus* human capabilities

In spite of the advantages of using autonomous (artificial) agents in several operational contexts, the human operators are still vital for the successful completion of a wide variety of missions. If it is not for assuming responsibility issues, the human operator is the agent that still produces tactical, moral, social and ethical decisions, while being flexible and creative, and able to handle complex and unknown situations [34]. These abilities are not (yet) embedded in artificial (automated) agents.

Interestingly, de Winter and Dodou (2014) [53] highlighted that the well-known Fitts' list (1951), *a list of 11 statements about whether a human or a machine performs a certain function better* still remains an interesting support - at least a starting point - for agent task allocation purposes. Yet this high-level concept does not take into account the capabilities of the agents during mission execution. Indeed, during mission execution, human operators may be subject to pressure, emotional commitment [50], task complexity [19] or long task duration [10], causing them to suffer from cognitive load or fatigue, and consequently, generating degraded mental states [13, 41]. These These various events can lead human operators to perform poorly by taking potentially poor decisions [13, 50].

For instance, a technical report from the U.S. Army, Navy, and Air Force [52] suggested the percentage of involvement of human factors in UAVs operations varied across aircraft models from 21% to 68%. It has also been shown in [31] that 50% of the terminal failures in disaster robotics are due to human error. To sum up, just like artificial agents, it must be accepted that human operators are not ideal agents either. Those examples show that, when mission efficiency is the goal, allocating tasks during mission execution to a non-performing agent, even if this agent is generally the most capable of executing the task, would not be the best solution. However, from the human operator's point of view, it is not always bearable or acceptable that artificial (automated) agents could seize the initiative, except if human cognitive capabilities or performance are degraded.

The mixed-initiative framework then proposes a reasonable design to deal with non-performing agents as it enables to the other ones to seize the initiative during mission execution.

2 A new mixed-initiative interaction paradigm

In our view, in order to cope with non-ideal agents, including the human operator, the mixed-initiative framework should be redefined as: ***a cooperation strategy that defines the role of involved agents according to their recognized skills and current capabilities***.

Such a novel framework definition requires to monitor the capabilities of all involved agents (human and automated agents). Fortunately, monitoring the states and actions of automated agents is supported by a recognized field of research. For instance, works related to execution monitoring can be found in the spacecraft [21] or robotics [37] literature.

Besides, monitoring the state and actions of a human operator (e.g. mental resources and performance) when interacting with automated systems is not a trivial task. Research in the Human Factors and Neuroergonomics fields is globally focused on the study and evaluation of the conditions in which the human reaches her/his limits of engagement during task operation [38, 41, 14, 10]. Neuroergonomics contributes to the development of human agents monitoring tools in ecological settings [42, 15] to avoid such degraded states.

However, the evaluation of the generality and the robustness of such algorithms still remains to be proved, given that the human-related features are usually operator-dependent, task-dependent and session-dependent. Thus, it is necessary to consider such monitoring system inaccuracies. In addition to that, the involved agents may also present non-deterministic behavior, what could result in different outcome states during mission execution. As example, these uncertainties may relate to the mission in question or its progression, the automated agent and the cognitive states of the human agent. Thus, it is necessary to cope with those uncertainties while ensuring efficient mission completion through the use of an adequate decision-making framework.

Therefore, supposing that both automated agents and human monitoring tools are available, a Mixed-Initiative Interaction Driving System (MI-IDS) could explicitly define the tasks (role) of agents for maximizing mission performance (see Figure 1). In

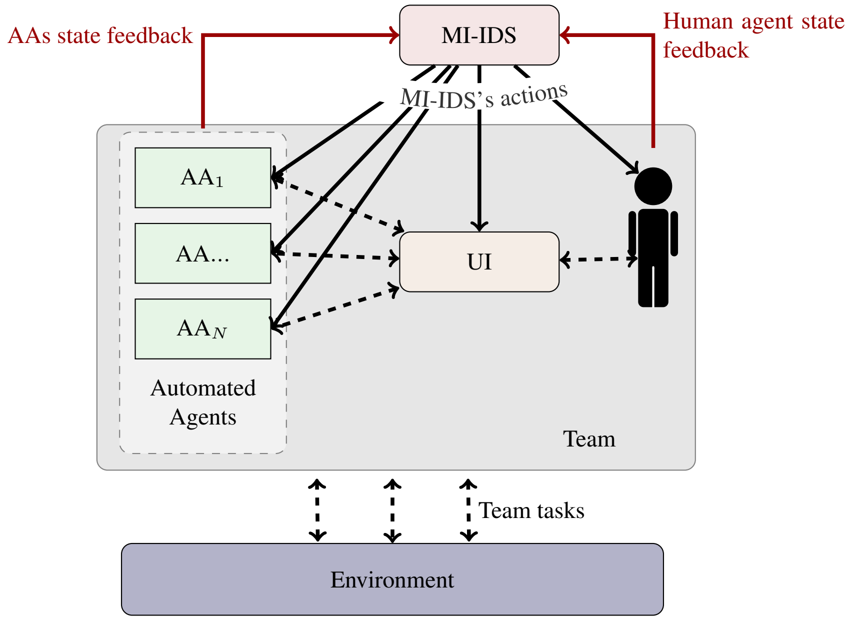


Fig. 1. Mixed-Initiative Interaction Driving System (MI-IDS) architecture. UI: User Interface.

other words, the MI-IDS estimates if an agent is able to seize the initiative from the other(s), while predicting if it is profitable for the long-term mission performance. Such a mixed-initiative concept has been applied to human-robot interaction [20, 49], and in assistive systems [22, 33], based on Partially Observable Markov Decision Processes (POMDPs) [28].

The POMDP framework takes into consideration uncertainties related to the partial observability of states (e.g. mental states) or the potentially non-deterministic behavior of the human operator [20, 9]. The POMDP solution can launch actions to mitigate the decline of the human agent's performance [49], or deliberately, can take the initiative to assign a given task to an automated agent [20] to ensure the proper accomplishment of the mission. The POMDP framework is detailed in the following part. Then, some of our previous works based on this framework are reviewed.

3 POMDP - theoretical background

The works presented in the following are built upon Partially Observable Markov Decision Processes (POMDPs), which offer a sound mathematical model for sequential decision-making under probabilistic uncertainty and partial observability.

A POMDP [47, 29] is a tuple $\langle \mathcal{S}, \mathcal{A}, \Omega, T, O, R, b_0 \rangle$, where: $\mathcal{S}$ is the set of states; $\mathcal{A}$ is the set of actions; Ω is the set of observations; $T : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ is the transition function, such that: $T(s, a, s') = p(s_{t+1} = s' | s_t = s, a_t = a)$ where $t \in \mathbb{N}$ is the time step of the process and s_t (resp. a_t) is the random variable representing the state (resp. action) at time step t; $O : \Omega \times \mathcal{A} \times \mathcal{S} \to [0,1]$ is the observation function such that: $O(o,a,s') = p(o_{t+1} = o | s_{t+1} = s', a_t = a)$ where o_{t+1} is the random variable representing the observation at time step $t+1$; $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the reward function, that defines the gain that a process going through a given state and executing a given action will add to its previous and future gains to define its total value; b_0 is the initial probability distribution over states. We denote by $\Delta_{\mathcal{S}} \subset [0,1]^{\mathcal{S}}$ the (continuous) set of probability distributions over states, named *belief state space*. Figure 2 depicts the dynamic influence diagram of a POMDP to drive human-automated agent interaction. The actions in this figure, the sequential choice of which must be optimised, are those of MI-IDS. Thus, these actions may relate to supervision actions, to launch alarms in order to mitigate the decline of the human agent's performance [49], or directly to assign a given task to an automated (artificial) agent [9, 20].

At each time step, the MI-IDS updates its current *belief state* b about the complete state of the system including human, artificial agents, their environment, mission, etc. This belief update is implemented by applying the Bayes' rule thanks to the information given by the action performed $a \in \mathcal{A}$ and the observation received $o \in \mathcal{O}$:

$$b_a^o(s') = \frac{p\,(\,o, s' \mid\; a, b\,)}{p\,(\,o \mid\; a, b\,)} \tag{1}$$

where

$$p\,(\,o, s' \mid\; a, b\,) = O(o,a,s') \sum_{s \in S} T(s,a,s') b(s)$$

and

$$p\,(\,o \mid\; a, b\,) = \sum_{s' \in \mathcal{S}} p\,(\,o, s' \mid\; a, b\,)\,. \tag{2}$$

Solving a POMDP consists in finding a policy function $\pi : \Delta_{\mathcal{S}} \to \mathcal{A}$ that maximizes a performance criterion. The expected discounted reward from any initial belief state

$$V^{\pi}(b) = E_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t r\big(b_t, \pi(b_t)\big) \;\middle|\; b_0 = b \right] \tag{3}$$

is usually the performance criterion optimized for an infinite horizon. In this criterion, $0 < \gamma < 1$ is called the *discount factor* and represents the probability that the process will stop at each time step. The optimal policy π^* is the policy that maximizes the value function (Equation 3). The resulting value function, called the *optimal value function* and denoted by V^*, satisfies the Bellman equation:

$$V^*(b) = \max_{a \in A} \left[r(b,a) + \gamma \sum_{o \in \Omega} p\,(\,o \mid\; a, b\,)\, V^*(b_a^o) \right] \tag{4}$$

where $r(b,a) = \sum_{s \in \mathcal{S}} R(s,a) b(s)$ and $p\,(\,o \mid\; a, b\,)$ is defined in Equation 2.

It has been proven that iterating the equation, starting with a value function which is piecewise linear and convex (PWLC), creates a sequence of PWLC value functions that converge to the optimal value function V^* [47]. Therefore, in the n^{th} stage of the optimization, the value function V_n can be parameterized as a set of hyperplanes over

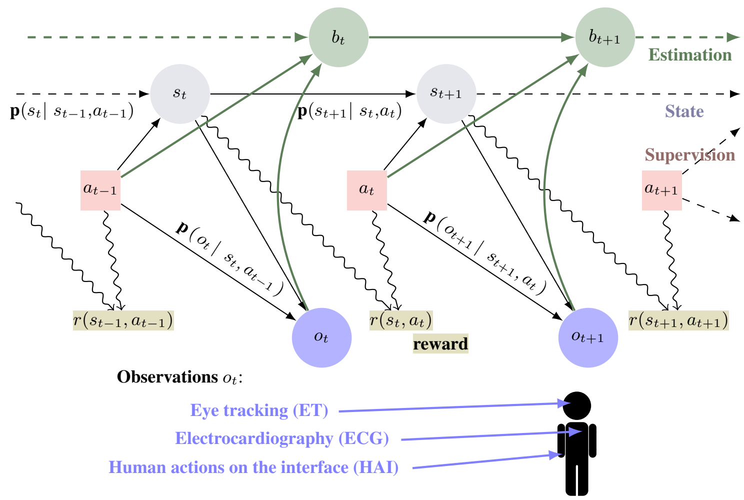


Fig. 2. Graphical representation of a POMDP for driving the human-automated agent interaction.

$\Delta_{\mathcal{S}}$ named α-*vectors* denoted by $\alpha_n \in \mathcal{V}_n \subset \mathbb{R}^{\mathcal{S}}$. Indeed, the value of a belief b can be defined as:

$$V_n(b) = \max_{\alpha_n \in \mathcal{V}_n} \sum_{s \in \mathcal{S}} \alpha_n(s) b(s).$$

Each α-vector is associated with an action $a(\alpha_n)$, e.g. $\forall a \in \mathcal{A}$, $R(.,a) \in \mathbb{R}^{\mathcal{S}}$ is an α-vector of V_0 such that:

$$V_0 = \max_{a \in \mathcal{A}} r(b,a) = \max_{a \in \mathcal{A}} \sum_{s \in \mathcal{S}} R(s,a) b(s) = \max_{\alpha_0 \in \mathcal{V}_0} \sum_{s \in \mathcal{S}} \alpha_0(s) b(s).$$

In the region of the belief space where α_n maximizes V_n, $a(\alpha_n)$ is the optimal action (at the n^{th} optimization stage). The optimal policy at this step is then: $\pi_n(b) = a^*(\alpha_n^*)$ such that $\alpha_n^* = \arg\max_{\alpha_n \in \mathcal{V}_n} \sum_{s \in \mathcal{S}} \alpha_n(s) b(s)$.

Recently, researchers proposed a structured POMDP model, named Mixed Observability Markov Decision Processes (MOMDPs, see [35, 3]), which factorizes the state space $\mathcal{S}$ in fully and partially observable parts: $\mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$. MOMDPs exploit the specific structure of the state set to reduce the dimension of the belief state space, resulting in a significant gain in policy computation time [35]. The principle of MOMDP solving is the same as POMDPs, it consists in finding a set of policies $\pi_{s_v} : \Delta_{\mathcal{S}_h} \rightarrow \mathcal{A}$, $\forall s_v \in \mathcal{S}_v$, which maximize the criterion:

$$\pi_{s_v}^* \in \underset{\pi_{s_v} : \Delta_{\mathcal{S}_h} \rightarrow \mathcal{A}}{\arg\max} \; E_{\pi_{s_v}} \left[\sum_{t=0}^{\infty} \gamma^t r_{s_v^t} \left(b_h^t, \pi_{s_v^t}(b_h^t) \right) \middle| \; b_0 = (s_v^0, b_h^0) \right] \tag{5}$$

where $b_h \in \Delta_{\mathcal{S}_h}$ and $r_{s_v}\big(b_h, \pi_{s_v}(b_h)\big) = \sum_{s_h \in \mathcal{S}_h} r(s_v, s_h, a) b_h(s_h)$.

As for the POMDP, the value function at a time step $n < \infty$ can also be represented by a set of α-vectors $\forall s_v \in \mathcal{S}_v$:

$$V_n(s_v, b_h) = \max_{\alpha \in \mathcal{V}_n(s_v)} \sum_{s_h \in \mathcal{S}_h} \alpha(s_h) b_h(s_h) \tag{6}$$

where, α is a hyperplan over the sub-space $\Delta_{\mathcal{S}_h}$. In this way, the value function over the complete state space is parametrized by the set $\mathcal{V}_n$ that in turn is composed by the sets $\mathcal{V}_n(s_v)$, i.e. $\mathcal{V}_n = \{\mathcal{V}_n(s_v), \forall s_v \in \mathcal{S}_v\}$. So, given a belief state (s_v, b_h) the optimal action $a^*(\alpha^*)$ is defined by the α^*-vector, such as:

$$\alpha^* \in \underset{\alpha \in \mathcal{V}_n(s_v)}{\arg\max} \sum_{s_h \in \mathcal{S}_h} \alpha(s_h) b_h(s_h). \quad (7)$$

For more details about MOMDP resolution and algorithms, please refer to [35, 3].

This efficient model framework has been used in our laboratory to drive the human-automated agents interaction with promising results as detailed in the following section.

4 Proof-of-concept systems

4.1 MOMDP implementations to drive the human-autonomous agents interaction

Search and rescue scenario. Humans are generally confronted with concurrent tasks while performing an automated system supervision. Gateau et al. (2016) [20] have addressed a dual-task paradigm in a search and rescue scenario. In the primary-task the human operator had to collaborate with autonomous artificial agents to perform target identifications by means of a user interface. In his/her secondary-task, the human operator had to memorize a series of digits (Short-Term Item Memorization task) and report it via the user interface. It was shown the volunteers had a lower performance when achieving both tasks simultaneously than when achieving each of them separately. Thus, special attention could be given to when a request should be launched to the human operator to perform target identification.

The main point in this work was to implement an integrated MOMDP-based system that would consider the *availability* of the human operator and her/his non-deterministic behavior based on the average time-to-answer a given request [20]. The *availability* of the human operator was measured by means of an eye-tracking (ET) device. The online processing of the gaze data acquired with such a device made possible to associate it to a region of the screen (or area of interest) where the human operator might be paying attention at a given time step t (visual attention estimation). The figure 3, taken from [20], illustrates the proposed architecture.

The use of such information in the MOMDP model allowed to compute a policy that performs requests to the human operator respecting his/her supposed *availability*, while allocating tasks to the automated (artificial) agents such as *go to target i, send request, proceed identification*. It was shown that applying a MOMDP policy that integrates the *availability* estimation (with gaze monitoring condition) improved the human operator's performance on the secondary-task, compared to a driving policy that does not consider it (without gaze monitoring condition). Moreover, it was also shown the performance of the automated agents in the search and rescue mission had not been penalized. The average expected rewards had the same order of magnitude in both studied conditions (with and without gaze monitoring). It means that even when the human

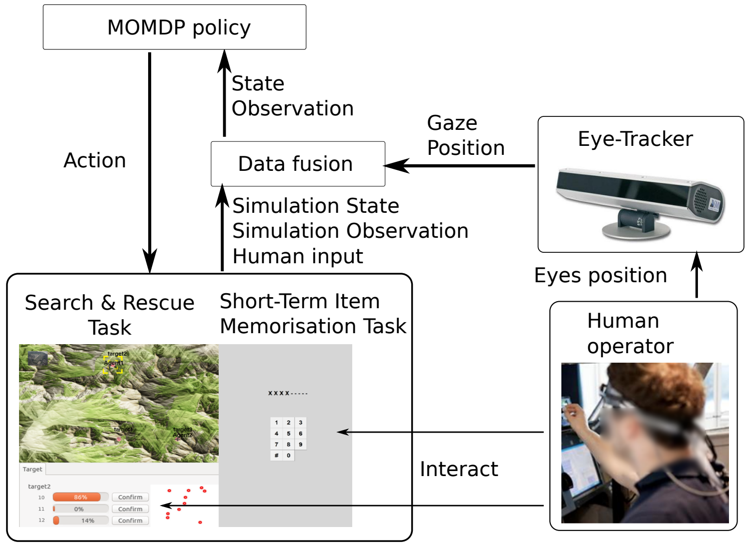


Fig. 3. MOMDP-based supervision architecture for the search and rescue scenario (from [20]).

operator was not available to help the autonomous agents in identifying targets (by answering a request), the system could achieve the target recognition task alone (but with more risk of error).

However, several mission parameters used to define the MOMDP model transition and observation functions were roughly estimated. For instance, the average time-to-answer a request given the stimuli used (yellow bounding box visual effect) was taken from [25], and fake values concerning the accuracy performance of the human operator or of the automated agents in the target identification task were used. And finally, beside the good results obtained with this dual-task paradigm, the mission performance metric (the one optimized by the policy) did not include the scores obtained by the human operator in the short-term memorization task (secondary-task). This score was not directly considered by the model. It means the policy was not taking it into account for the maximization of the expected sum of rewards problem. The performance improvement of the operator in this secondary-task was only a consequence of the fact that the policy chose to achieve the recognition task without help of the human operator, that, at a given moment was estimated as *not available* to the automated agents. So, the policy indirectly favored the secondary-task performance improvement.

Cognitive state estimation during a target search mission. As highlighted in section 1.1, the human operator should not be considered as a providential agent always capable to take over when the automated (artificial) agent fails. In this sense, the originality of the work proposed in [49] relies on the fact that a MOMDP model, approached by experimental data [41], was defined to supervise the interaction between a mobile teleoperated robot and a human operator, this last, supposed to be monitored thanks to eye-tracking (ET) and electrocardiogram (ECG) devices.

In the proposed MOMDP model, the robot operation mode (autonomous or manual mode) and mission states (e.g. going to zone, searching the target, handling the target, returning to base, on base and failed) could be considered as fully observable state variables, while the operator's cognitive ability, considered as *Cognitive Availability*, was modeled as a partially observable state variable. The *Cognitive Availability* was defined as the opposite of *Attentional Tunneling* [11], a degraded mental state in which the human operator fails in detecting changes in the environment because his/her attention is focused on another specific task.

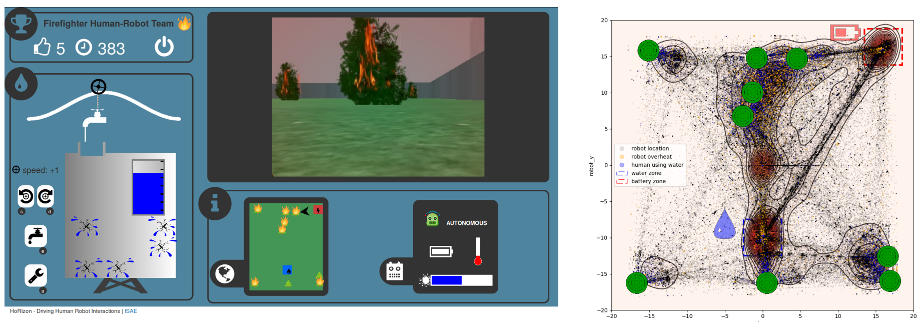
The estimation of this partially observable state variable, ensured by the *belief state update*, was based on the output of the ANFIS classifier proposed in [41]. This classifier demonstrated a good performance, being able to detect the *Attentional Tunneling* state based on eye-tracking (ET) and electrocardiogram (ECG) features. Technically, the confusion matrix of the ANFIS classifier was used to approximate the $p(o_{t+1}|s_{t+1}, a_t)$ probability function used in the MOMDP model.

The actions considered in the model were related with high-level instructions such as: *go to zone*, *search target*, *handle target*, *return to base*, *get attention* and *countermeasure launching*. The first four actions could be performed by the policy only if the robot was in autonomous mode. For instance, it was assumed the mobile robot was able to autonomously navigate and avoid obstacles, but if the policy chooses *go to zone* and the ultrasound embedded sensor fails, the robot mode turns to manual mode. Interestingly, the *get attention* action, supposed to be implemented in the user interface, should be used when the robot needed help and the operator's Cognitive Availability was estimated as *not available*. The *countermeasure launching* action, also supposed to be implemented in the user interface, should be executed when an important event arrives during a manual operation and the operator was considered as *not available* (e.g his attention was focused on handling the robot and he would not notice the alerts on the user interface).

In spite of the interesting methodology described in [49] to approach the MOMDP model based on experimental data and on the classifier accuracy [41], the obtained policy was only evaluated in simulation. Statistical analysis about mission states visitation, and *get attention* or *countermeasure* actions launching were discussed. Moreover, the transition function of several state variables where roughly approximated given that the previous experiments [41] had not provided enough data concerning all mission states transitions.

The firefighter robot game. A more recent project, based on a firefighter robotic mission [18, 9, 7] has been addressed in our lab in order to mitigate the drawbacks of our previous approaches: (i) the need for sufficient data to feed (MO)MDP models; (ii) the definition of a global common score reflecting mission performance, and being dependent of the efficiency of all involved agents; (iii) the implementation of a monitoring system to estimate the cognitive state of the human-operator during the interaction; and finally (iv) the evaluation in situ of the resulting (MO)MDP policy to drive such an interaction system.

The Firefight Robot game is a human-robot mission [18] that immerses the human-operator in a scenario where she/he plays a fireman who must cooperate with a robot that is present in a small area with few trees. The goal of the mission is to fight as many fires as possible in a limited amount of time (ten minutes). Through the user interface (UI), shown in 4(a), the human operator can supervise all robot parameters (position, temperature, battery, embedded water tank level, operation mode) and can receive the video streaming from its camera.



(a) The user interface (UI).

(b) Crowdsourcing data plot example.

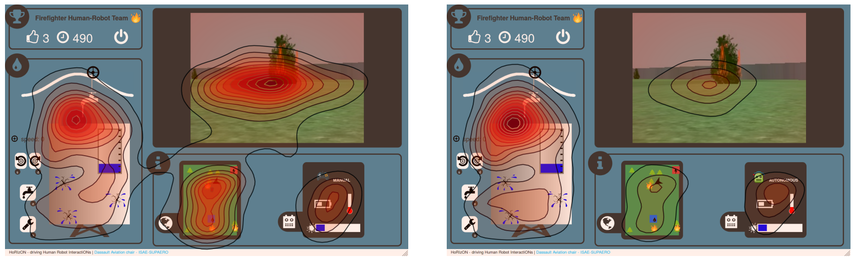
Fig. 4. The Firefighter Robot game. Figure (a) shows the human operator control interface. On the left of this interface is shown the current score and remaining mission time. The ground tank sub-task is bellow it. On the right of this interface the robot video stream is displayed, and below it, the map locating the robot in the arena and robot parameters. Figure (b) plots several robot trajectories illustrating the amount of data acquired thanks to the crowdsourcing platform.

In this mission scenario, the battery charge level of the robot decreases with time, then needing probably to be recharged several times during the mission. If the battery is empty the mission fails and is finished. The volume of water contained by the robot is not unlimited: to recharge in water, the robot has to reach the a water tank that in turn should dispose of enough water. For that, the human operator has to continuously fill this ground tank using dedicated buttons in the UI. Unfortunately, leaks may appear on the ground tank during the mission, and the human-operator needs to fix them using the dedicated buttons for that. The temperature of the robot can increase when it is too close to flames and the mission terminates when it is too hot. The robot has two modes of operation: a manual mode, in which it is directed piloted by the human operator; and an mode, in which the robot drives itself with a hard-coded strategy, including shooting water and the recharge of water or battery when necessary.

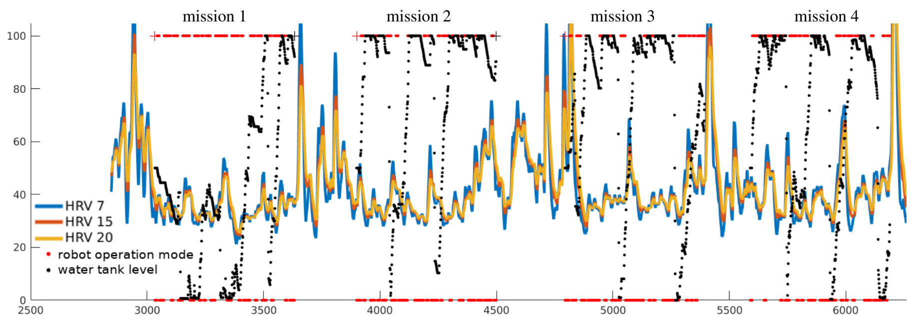
In this mission the presence of fires is supposed to be felt as a danger by the operator. The limited time for mission accomplishment would induce pressure. The temperature and battery risk would require human operator to monitor robot parameters all the time. Moreover, the ground tank filling sub-task, only performed by the human operator is demanding and requires a constant attention. All these elements are suppose to lead to a high engagement from the human-operator during the mission, and may favor the appearance of degraded cognitive states (e.g. performance decline).

This robotic mission is still available on an opened crowdsourcing platform[1]. Since the website launch, advertising was done in the authors' (professional and social) net-

[1] http://robot-isae.isae.fr/welcome



(a) Eye gaze fixations density with respect to operation modes: manual (in the left), and autonomous (in the right).



(b) Heart rate variability (HRV) of one of the participants during mission realisations. One may note that as the ground tank level increases HRV decreases. It may relates to an increase of the cognitive load [8, 7] of the human operator.

Fig. 5. Examples of physiological data acquired during Firefighter Robot mission realizations. The figure (a) illustrates the eye gaze density of participants with respect to the robot operation mode. The figure (b) plots the heart rate variability (HRV) for one of the participants among missions. The HRV feature have been online computed [8] and than recorded for further analysis.

works to encourage Internet users to carry out the mission in order to collect as much anonymous data as possible. During mission realization, the robot operation mode can change randomly, as well as, the appearance of alarms related to the robot parameters. These random mode changes and alarm launches were implemented with the aim of obtaining a representative and balanced dataset. Currently, we have collected more than a thousand mission realizations (see Figure 4(b) where those data is illustrated).

In a first work [9] based on those data, we proposed the application of machine learning techniques, namely classification and Markov Chain learning, to define the parameters of a Markov Decision Process (MDP) model - it consists in the same principle as the MOMDP with only fully observable states. The suggested MDP model has explored behavioral data (mouse clicks and keystrokes) from users. Until today, the evaluation of the obtained MDP policy was only performed in simulation [9]. In any case, the proposed methodology for the model learning and the obtained simulation results suggested the approach is promising.

In complement of the data acquired thanks to the crowdsourcing platform, experiments have been done in ISAE-SUPAERO facilities. During these experiments several volunteers equipped with eye-tracking (ET) and electrocardiogram (ECG) have played the Firefighter Robot game following an experimental protocol including 4 mission realizations per participant. The acquired physiological and behavioral data were studied and allowed us to reveal interesting results concerning performance prediction [7].

The participants performed differently so that we could identify high and low score mission groups that also exhibited different behavioral, cardiac (e.g heart rate, heart rate variability) and ocular patterns (e.g. number of fixations in the areas of interest). The Figure 5 shows some ET and ECG features collected during experiments. More specifically, our results, recently published in [7], indicated that the higher level of automation could be beneficial to low-scoring participants but detrimental to high-scoring ones and vice versa. In addition, inter-subject single-trial classification results showed that the studied behavioral and physiological features were relevant to predict mission performance. The highest average balanced accuracy (74%) was reached using the features extracted from all input devices (ET, ECG, mouse clicks and keystrokes). We believe these features, computed on sliding 10-second time windows, and the achieved classification accuracy will allow us to approach an observation function (i.e. O). This observation function will be explored to maintaining a *belief state* on the human-operator engagement during mission.

The combination of the methods suggested in [9] and in [7] is under study. The aim will be to proceed to a MOMDP model learning (with all experimental and crowdsourcing data acquired) followed by the policy computation. We expect the achieved policy will choose the operation mode of the robot, or to launch alarms, depending on the current human-operator engagement (or performance), in order to maximize the overall mission performance.

5 Conclusion and Perspectives

The research concerning human-autonomous agents interaction still is an actual field. Among authority sharing approaches, the mixed-initiative interaction design has been shown to be a promising framework to drive such interactions, as it considers all involved agents (human and artificial) as teammates rather than master-slave agents. In this paper we suggested a redefinition of the mixed-initiative interaction concept in order to integrate the notion of *agents' current capabilities*, assuming that performance of the involved agents - including the human agent - may not be forthcoming. This new definition implicitly relies on the development of agent monitoring systems.

Our current contributions on this topic propose to rely on the POMDP model, or on the (MO)MDP model following state variables observability assumptions. The POMDP framework allows to model non-deterministic (probabilistic) action effects of the human-automated agents system and observation uncertainties (inaccuracies) related to the human operator monitoring systems. This decision-making framework enables, on the one hand, to maintain a *belief state* on human operator cognitive state, and on the other hand, to compute an interaction driving policy able to adapt the automated agents' behavior in function of such an estimation.

The presented proof-of-concept missions have shown promising results, as those systems have demonstrated: to increase the human performance during mission accomplishment [20] and to be able to launch countermeasures [49] when the interaction driving system estimates it necessary. However, as discussed, in order to perform well the

POMDP framework needs an accurate modelling of the human-automated agent interaction, which in turn depends on huge data acquisition and confident human monitoring means based on human performance-related features.

In this vein, a crowdsourcing platform was developed [18], and the collected data have been explored [9] to approximate the state variables transition function - it allows to model non-deterministic (probabilistic) action effects. Lab experiments with volunteer participants have also been done, based on the same platform [7]. The aim of this study was to identify behavioral (mouse clicks, keystrokes, eye gaze fixations) and physiological (cardiac activity) features useful to predict performance using classifiers. Thus, a possible solution to approximate an observation function of the cognitive state of the human operator could be based on such classifier outputs.

Future work will address the problem of merging crowdsourcing and lab experiments datasets extending [9, 7] works. A special attention will be given to an automated way to learn which are the relevant state variables, as well as their granularity in order to attain a sufficient faithful model that can be processed and solved by existing POMDP (MOMDP) algorithms. It specially relates to the trade-off between a fair model and a solvable model. On another hand, one can question the confidence of the human operator monitoring system in detecting degraded mental states. Ongoing work is searching for online solutions that exploits, in addition of the ET and ECG features, electroencephalogram (EEG) features [46] in order to increase confidence of the monitoring system. And finally, the *in situ* evaluation with a real robot, of the resulting POMDP policy is planned to be done.

Acknowledgments

Many thanks to all undergraduate, graduate master and PhD students, technical employees, and research colleagues that have contributed to the development of these works. These works were financially supported by the Direction Générale de l'Armement (DGA) nowadays Agence de l'Innovation de Défense (AID), the Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brazil (CNPq), and the Dassault Aviation Chair[2] (CASAC 2016-2021).

References

1. Adams, J.A., Rani, P., Sarkar, N.: Mixed initiative interaction and robotic systems. In: AAAI Workshop on Supervisory Control of Learning and Adaptive Systems. pp. 6–13 (2004)
2. Allen, J., Guinn, C.I., Horvtz, E.: Mixed-initiative interaction. IEEE Intelligent Systems and their Applications **14**(5), 14–23 (1999)
3. Araya-Lopez, M., Thomas, V., Buffet, O., Charpillet, F.: A Closer Look at MOMDPs. In: Proc. of the 22nd IEEE International Conference on Tools with Artificial Intelligence - Volume 02. pp. 197–204. ICTAI '10, IEEE Computer Society, Washington, DC, USA (2010)
4. Beer, J.M., Fisk, A.D., Rogers, W.A.: Toward a framework for levels of robot autonomy in human-robot interaction. Journal of human-robot interaction **3**(2), 74–99 (2014)

[2] https://www.isae-supaero.fr/fr/isae-supaero/mecenat-relations-avec-la-fondation-isae-supaero/chaire-dassault-aviation-casac

5. Benavides, F., Ponzoni Carvalho Chanel, C., Monzón, P., Grampín, E.: An auto-adaptive multi-objective strategy for multi-robot exploration of constrained-communication environments. Applied Sciences **9**(3), 573 (2019)
6. Carneiro, J., Martinho, D., Marreiros, G., Novais, P.: Intelligent negotiation model for ubiquitous group decision scenarios. Frontiers of Information Technology & Electronic Engineering **17**(4), 296–308 (2016)
7. Chanel, C.P., Roy, R.N., Dehais, F., Drougard, N.: Towards mixed-initiative human–robot interaction: Assessment of discriminative physiological and behavioral features for performance prediction. Sensors **20**(1), 296 (2020)
8. Chanel, P.C., Wilson, M.D., Scannella, S.: Online ecg-based features for cognitive load assessment. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). pp. 3710–3717. IEEE (2019)
9. Charles, J.A., Chanel, C.P., Chauffaut, C., Chauvin, P., Drougard, N.: Human-agent interaction model learning based on crowdsourcing. In: Proceedings of the 6th International Conference on Human-Agent Interaction. pp. 20–28. ACM (2018)
10. Cummings, M.L., Mastracchio, C., Thornburg, K.M., Mkrtchyan, A.: Boredom and distraction in multiple unmanned vehicle supervisory control. Interacting with Computers **25**(1), 34–47 (2013)
11. Dehais, F., Causse, M., Tremblay, S.: Mitigation of conflicts with automation: use of cognitive countermeasures. Human Factors **53**(5), 448–460 (2011)
12. Dehais, F., Causse, M., Vachon, F., Tremblay, S.: Cognitive conflict in human–automation interactions: a psychophysiological study. Applied ergonomics **43**(3), 588–595 (2012)
13. Dehais, F., Hodgetts, H.M., Causse, M., Behrend, J., Durantin, G., Tremblay, S.: Momentary lapse of control: A cognitive continuum approach to understanding and mitigating perseveration in human error. Neuroscience & Biobehavioral Reviews (2019)
14. Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., Gateau, T.: Automation surprise in aviation: Real-time solutions. In: Proceedings of the 33rd annual ACM conference on Human Factors in Computing Systems. pp. 2525–2534. ACM (2015)
15. Dehais, F., Rida, I., Roy, R.N., Iversen, J., Mullen, T., Callan, D.E.: A pbci to predict attentional error before it happens in real flight conditions. In: International Conference on Systems, Man, and Cybernetics, 9th Workshop on Brain-Machine Interface Systems. IEEE (2019)
16. Delamer, J.A., Watanabe, Y., P. Carvalho Chanel, C.: Solving path planning problems in urban environments based on a priori sensors availabilities and execution error propagation. In: AIAA Scitech 2019 Forum. p. 2202 (2019)
17. Desai, M.: Sliding scale autonomy and trust in human-robot interaction. Ph.D. thesis, University of Massachusetts. Lowell. (2007)
18. Drougard, N., Chanel, C.P.C., Roy, R.N., Dehais, F.: Mixed-initiative mission planning considering human operator state estimation based on physiological sensors. In: IROS workshop on Human-Robot Interaction in Collaborative Manufacturing Environments (HRI-CME) (2017)
19. Durantin, G., Gagnon, J.F., Tremblay, S., Dehais, F.: Using near infrared spectroscopy and heart rate variability to detect mental overload. Behavioural brain research **259**, 16–23 (2014)
20. Gateau, T., Chanel, C.P.C., Le, M.H., Dehais, F.: Considering human's non-deterministic behavior and his availability state when designing a collaborative human-robots system. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 4391–4397. IEEE (2016)
21. Georgeff, M.P., Ingrand, F.F.: Real-time reasoning: The monitoring and control of spacecraft systems. In: Sixth Conference on Artificial Intelligence for Applications. pp. 198–204. IEEE (1990)

22. Hoey, J., Von Bertoldi, A., Poupart, P., Mihailidis, A.: Assisting persons with dementia during handwashing using a partially observable markov decision process. In: International Conference on Computer Vision Systems: Proceedings (2007)
23. Hrabia, C.E., Hessler, A., Xu, Y., Seibert, J., Brehmer, J., Albayrak, S.: Efffeu project: Towards mission-guided application of drones in safety and security environments. Sensors **19**(4), 973 (2019)
24. Huang, H.M., Pavek, K., Novak, B., Albus, J., Messin, E.: A framework for autonomy levels for unmanned systems (alfus). Proceedings of the AUVSI's Unmanned Systems North America pp. 849–863 (2005)
25. Imbert, J.P., Hodgetts, H.M., Parise, R., Vachon, F., Dehais, F., Tremblay, S.: Attentional costs and failures in air traffic control notifications. Ergonomics **57**(12), 1817–1832 (2014)
26. Javdani, S., Admoni, H., Pellegrinelli, S., Srinivasa, S.S., Bagnell, J.A.: Shared autonomy via hindsight optimization for teleoperation and teaming. The International Journal of Robotics Research **37**(7), 717–742 (2018)
27. Jiang, S., Arkin, R.C.: Mixed-initiative human-robot interaction: definition, taxonomy, and survey. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics. pp. 954–961. IEEE (2015)
28. Kaelbling, L.P., Littman, M.L., Cassandra, A.R.: Planning and acting in partially observable stochastic domains. Artificial intelligence **101**(1-2), 99–134 (1998)
29. Kaelbling, L., Littman, M., Cassandra, A.: Planning and acting in partially observable stochastic domains. AIJ **101**(1-2) (1998)
30. Liu, C., Kroll, A.: A centralized multi-robot task allocation for industrial plant inspection by using a* and genetic algorithms. In: International Conference on Artificial Intelligence and Soft Computing. pp. 466–474. Springer (2012)
31. Murphy, R.R.: Disaster robotics. MIT press (2014)
32. Murphy, R.R., Tadokoro, S., Nardi, D., Jacoff, A., Fiorini, P., Choset, H., Erkmen, A.M.: Search and rescue robotics. In: Springer Handbook of Robotics, pp. 1151–1173. Springer (2008)
33. Nikolaidis, S., Ramakrishnan, R., Gu, K., Shah, J.: Efficient model learning from joint-action demonstrations for human-robot collaborative tasks. In: 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 189–196. IEEE (2015)
34. Nothwang, W.D., McCourt, M.J., Robinson, R.M., Burden, S.A., Curtis, J.W.: The human should be part of the control loop? In: 2016 Resilience Week (RWS). pp. 214–220. IEEE (2016)
35. Ong, S.C.W., Png, S.W., Hsu, D., Lee, W.S.: Planning under uncertainty for robotic tasks with mixed observability. Int. J. Rob. Res. **29**(8), 1053–1068 (Jul 2010)
36. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: 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 (2000)
37. Pettersson, O.: Execution monitoring in robotics: A survey. Robotics and Autonomous Systems **53**(2), 73–88 (2005)
38. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biological psychology **40**(1-2), 187–195 (1995)
39. Portugal, D., Rocha, R.P.: Distributed multi-robot patrol: A scalable and fault-tolerant framework. Robotics and Autonomous Systems **61**(12), 1572–1587 (2013)
40. Portugal, D., Rocha, R.P.: Cooperative multi-robot patrol with bayesian learning. Autonomous Robots **40**(5), 929–953 (2016)
41. Régis, N., Dehais, F., Rachelson, E., Thooris, C., Pizziol, S., Causse, M., Tessier, C.: Formal detection of attentional tunneling in human operator–automation interactions. IEEE Transactions on Human-Machine Systems **44**(3), 326–336 (2014)

42. Roy, R.N., Charbonnier, S., Campagne, A., Bonnet, S.: Efficient mental workload estimation using task-independent eeg features. Journal of neural engineering **13**(2), 026019 (2016)
43. Scerbo, M.W.: Adaptive automation. Neuroergonomics: The brain at work pp. 239–252 (2008)
44. Schurr, N., Marecki, J., Tambe, M.: Improving adjustable autonomy strategies for time-critical domains. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1. pp. 353–360. International Foundation for Autonomous Agents and Multiagent Systems (2009)
45. Sheridan, T.B.: Adaptive automation, level of automation, allocation authority, supervisory control, and adaptive control: distinctions and modes of adaptation. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans **41**(4), 662–667 (2011)
46. Singh, G., Roy, R.N., Ponzoni Carvalho Chanel, C.: Towards multi-uav and human interaction driving system exploiting human mental state estimation. In: 10th International Conference on Bioinformatics Models, Methods and Algorithms (2019)
47. Sondik, E., LABS, S.U.C.S.E.: The Optimal Control of Partially Observable Markov Processes. Stanford University, California (1971)
48. Ubaldino de Souza, P.E.: Towards mixed-initiative human-robot interaction: a cooperative human-droneteam framework. Ph.D. thesis, Institut Supérieur de l'Aéronautique et de l'Espace, Université de Toulouse (2017), https://depozit.isae.fr/theses/2017/2017_Ubaldino_de_Souza_Paulo-Eduardo_D.pdf
49. de Souza, P.E.U., Chanel, C.P.C., Dehais, F.: Momdp-based target search mission taking into account the human operator's cognitive state. In: 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI). pp. 729–736. IEEE (2015)
50. Souza, P.E., Chanel, C.P.C., Dehais, F., Givigi, S.: Towards human-robot interaction: a framing effect experiment. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). pp. 001929–001934. IEEE (2016)
51. Sun, X., Wu, P., Hoi, S.C.: Face detection using deep learning: An improved faster rcnn approach. Neurocomputing **299**, 42–50 (2018)
52. Williams, K.W.: A summary of unmanned aircraft accident/incident data: Human factors implications. Tech. rep., Federal Aviation Administration Oklahoma City, Civil Aeromedical Institute (2004)
53. de Winter, J.C., Dodou, D.: Why the fitts list has persisted throughout the history of function allocation. Cognition, Technology & Work **16**(1), 1–11 (2014)