The Firefighter Robot Game

NIcolas DROUGARD & Caroline CHANEL

DCAS / ISAE-SUPAERO

prenom.nom@isae-supaero.fr

Outline

Increasing use of automated systems

assembly lines, computers, autopilots in aircrafts, autonomous cars, unmanned vehicles: drones/ground robots for military operation or contaminated area . . .

Increasing use of automated systems

assembly lines, computers, autopilots in aircrafts, autonomous cars, unmanned vehicles: drones/ground robots for military operation or contaminated area ...

- **Increasingly decisional autonomy:**
  technical advances in AI, machine learning, vision, decision ...

Increasing use of automated systems

assembly lines, computers, autopilots in aircrafts, autonomous cars, unmanned vehicles: drones/ground robots for military operation or contaminated area ...

- **Increasingly decisional autonomy:**
  technical advances in AI, machine learning, vision, decision ...
- **Need to consider human behavior into decision loop:**

  $\rightarrow$ complementary skills (Fitts list [de Winter and Dodou, 2014]):
  - produces tactical, moral, social and ethical decisions;
  - flexible/creative, handles complex/unknown situations;

  $\rightarrow$ legal: - need for people for responsibility assessment issues.

For instance: the percentage of involvement of human factors in UAVs operations varied across aircraft from 21% to 68%. (U.S. Army, Navy, and Air Force report [Williams, 2004])

Constraints experienced by humans

- pressure (e.g. cause by a danger) → stress
- task complexity → mental workload
- task hardness and duration → fatigue, boredom

Consequences:

- disengagement, lower vigilance
- mind wandering
- over-engagement, attentional tunneling
- mental confusion... **affect human abilities!**

Increase in accident risk resulting in mission fails
or sub-optimal achievements

Mixed-Initivative framework

it defines the role of the human and artificial agents according to their recognized skills [Allen et al., 1999, Adams et al., 2004]

Mixed-initiative for HRI

[Jiang and Arkin, 2015]: *A collaboration strategy for human-robot teams where humans and robots opportunistically seize (relinquish) initiative from (to) each other as a mission is being executed*

Human operators are not providential agents

- not always reliable
- not always omniscient during the mission
- not always able to fix any occurring issue

Context

Improve mixed-initiative missions

A possible solution: Mixed-initiative strategy computation

it should define the role of the human and artificial agents according to their recognized skills and current capabilities [Chanel et al., 2020b].
→ need for **monitoring the operator**'s cognitive state (mental resource)
→ then **decide** when the automated system should or should not have the authority
→ **adaptation** of machine's behavior

Supervision strategy computation

taking into account the human operator behavior
→ task allocation between the human and the machine
→ sending appropriate information/alarms
→ adaption of machine's behavior

Need to consider the human for system dynamics modelling

→ human non-deterministic behavior
→ partial observability of human (mental state)
→ need for mission performance maximization

Outline

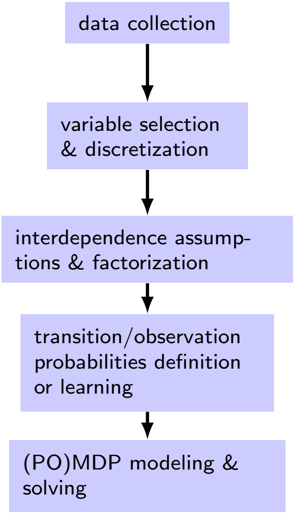
Human-Agent Interaction Model Learning

based on Crowdsourcing

human behavior & random dynamics
→ uncertain events
→ **Probabilistic Planning**
→ (PO)MDP

precise transition/observation probability values
→ need for a sufficiently large dataset
→ time consuming experiments
& numerous volunteers

(1) call on **remote volunteers** participating through a website
robot-isae.isae.fr
(2) call for volunteers in lab implementing an experimental protocol to acquire physiological data



Outline

Proof of concept mission

Mission and goal

Context:

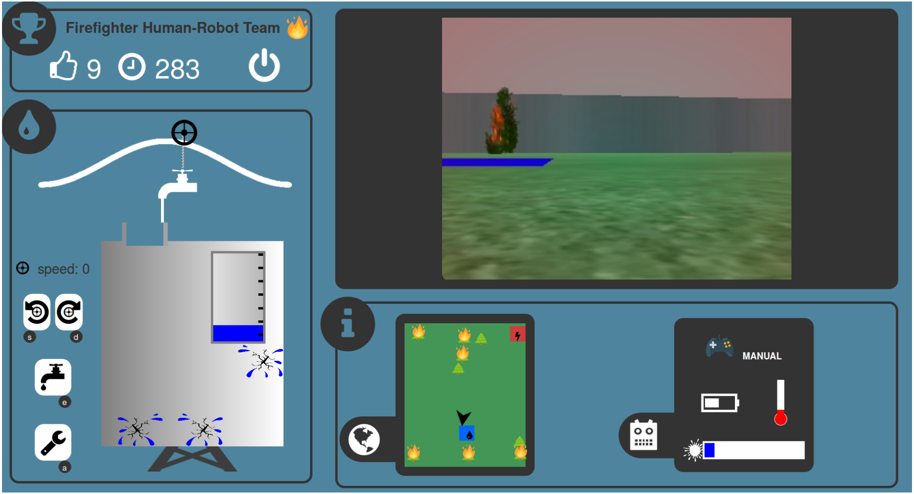
- A firefighter robot is present in a small area
- with few trees
- which have a weird tendency to self-ignite for some unknown reason...
- Through your graphical interface, you get the position of the robot,
- as well as the video from its camera.
- The battery charge level of the robot decreases with time: when the robot is on the red square, the battery recharges.
- The volume of water contained by the robot is not unlimited: to recharge its water tank, the robot has to be placed on the blue square and the ground tank has to be full enough. For that, you have to fill the ground tank using the buttons on the left-side of the interface.

Goal:

- With the help of this robot, your mission is to fight as much fires as you can.
- The robot can become autonomous at any time.
- Pay attention to the robot's temperature when it is too close of flames!

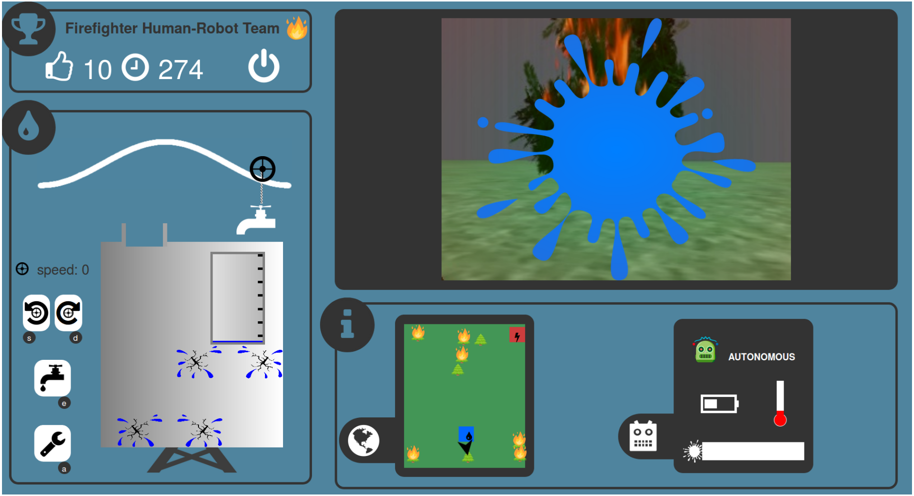
Proof of concept mission

Graphical user interface (GUI)



Proof of concept mission

Graphical user interface (GUI)



Proof of concept mission

Displayed alarms

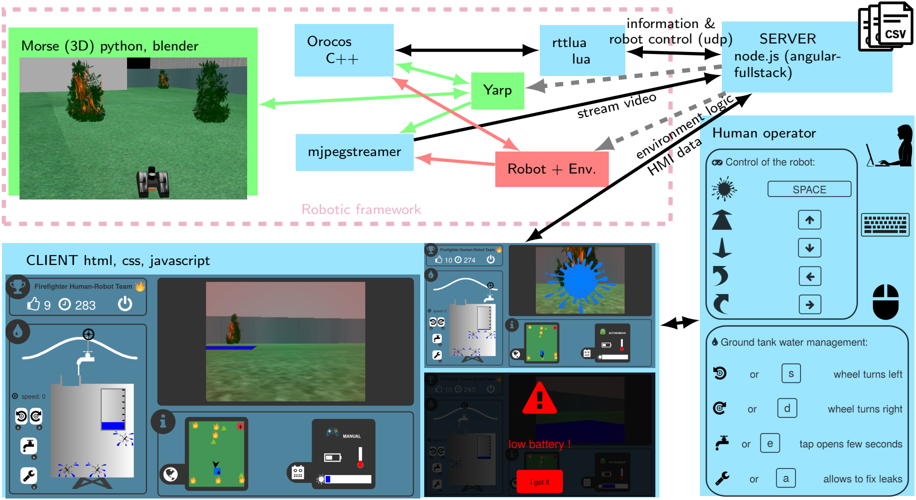
Proof of concept mission

Controlled dynamics

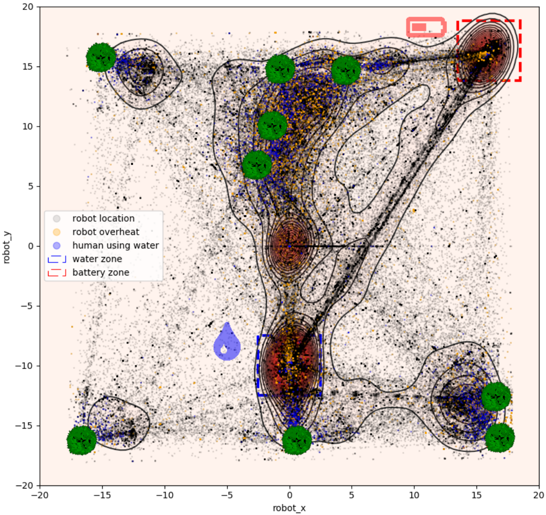
- **balanced dataset** (with & without autonomy/alarms)
  - → random robot's operating mode (auto/manual control)
  - → random display of information to the operator

- pre-testing for:
  - → **degraded human behaviors** emergence:
    - - water management + robot's control → demanding
    - - score/limited time → pressure
    - - battery empty or too hot robot → danger
  - → while **attractive** for people, in order to get data (challenging, amazing, feasible, *etc.*)

robot-isae.isae.fr



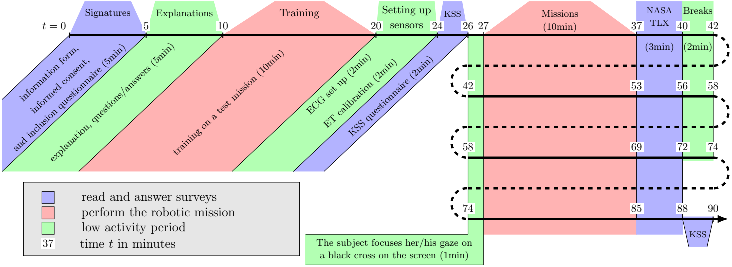
- clicks, keystrokes, state of the mission/robot, shared actions;
- total recorded time: 86h 25m 54s;
- with human actions: 48h 13m 26s.



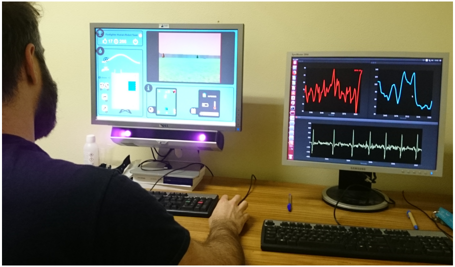
;

Proof of concept mission in lab conditions

Experimental protocol



18 participants in lab equipped with physiological sensors:
→ Human actions on the interface (HAI)
→ ECG
→ Eye-tracking (ET)



Proof of concept mission in lab conditions
Data collected

In lab experiments, the Lab Streaming Layer (LSL) was used to save all data throw LabRecorder.
$\rightarrow$ it provides all drivers to acquire data, and an unique data logger software.
$\rightarrow$ it enables data synchronization afterwards.

But, what we have ?

- **HAI:** at each second all keystrokes and mouse clicks were pushed to LSL
- **ECG:** Heart-rate (inter-beat interval), heart-rate variability (windowed), and raw-ECG data were pushed to LSL
- **ET:** Fixations events (80ms), and gaze position (250Hz) were pushed to LSL
- **Others:** robot position, external tank level, battery level, embedded water tank level, time to mission end, trees states, etc.

Outline

Previous work

Probabilistic planning

human behavior + random dynamics of environments
$\rightarrow$ uncertain events $\rightarrow$ **Probabilistic Planning**

Previous work

Probabilistic planning

human behavior + random dynamics of environments
$\rightarrow$ uncertain events $\rightarrow$ **Probabilistic Planning**

Markov Decision Process **MDP** [Puterman, 2014]

Formally, an **MDP** is: $\langle \mathcal{S}, \mathcal{A}, T, R, H \rangle$

- states $s \in \mathcal{S}$ describing the system
- actions $a \in \mathcal{A}$ to be chosen at each time step
- **transition** function $T(s, a, s') = \mathbf{p}\left(s' \mid s, a \right)$
- goal of a mission $\sim$ **reward** function $R(s, a) \in \mathbb{R}$

Previous work
Probabilistic planning

human behavior + random dynamics of environments
$\rightarrow$ uncertain events $\rightarrow$ **Probabilistic Planning**

Markov Decision Process **MDP** [Puterman, 2014]

Formally, an **MDP** is: $\langle \mathcal{S}, \mathcal{A}, T, R, H \rangle$

- states $s \in \mathcal{S}$ describing the system
- actions $a \in \mathcal{A}$ to be chosen at each time step
- **transition** function $T(s, a, s') = \mathbf{p}\left(s' \mid s, a\right)$
- goal of a mission $\sim$ **reward** function $R(s, a) \in \mathbb{R}$

strategy $\pi_t : s_t \mapsto a_t \in \mathcal{A}$

$$\text{maximizing } \mathbb{E}\left[\sum_{t=0}^{H} R\Big(s_t, \pi_t(s_t)\Big)\right], \; H > 0$$

Previous work

Markov Decision Process framework

- $s \in \mathcal{S}$: **system states**
- $a \in \mathcal{A}$: **actions**
- $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: **rewards**

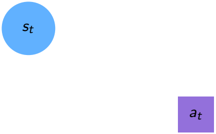
Previous work
Markov Decision Process framework

- $s \in \mathcal{S}$: **system states**
- $a \in \mathcal{A}$: **actions**
- $r : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$: **rewards**

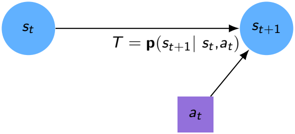
Previous work

Markov Decision Process framework

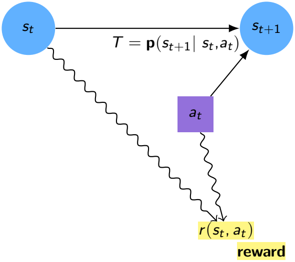
- $s \in \mathcal{S}$: **system states**
- $a \in \mathcal{A}$: **actions**
- $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: **rewards**



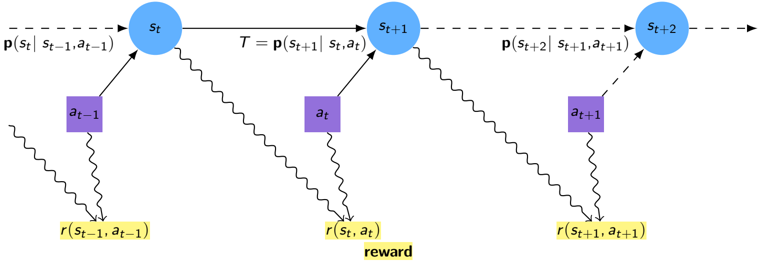
- $s \in \mathcal{S}$: **system states**
- $a \in \mathcal{A}$: **actions**
- $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: **rewards**



- $s \in \mathcal{S}$: **system states**
- $a \in \mathcal{A}$: **actions**
- $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: **rewards**



- $s \in \mathcal{S}$: **system states**
- $a \in \mathcal{A}$: **actions**
- $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: **rewards**



Previous work

Variable selection & discretization

- pos = robot's position ($\#pos = 9$ zones)
- bat = robot's battery level ($\#bat = 2$ levels)
- wat = embedded water tank ($\#wat = 2$ levels)
- fire = number of burning trees ($\#fires = 10$ forest states)
- space = use of the space key (boolean variable)
- end = game over (boolean variable)
- int = human intention ($\#int = 9$ movements)

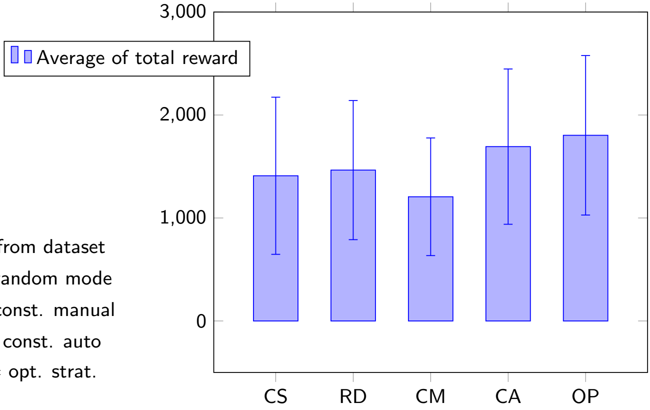
Full state space size $\#\mathcal{S} = 9 \times 2 \times 2 \times 10 \times 2 \times 2 \times 9 = 12960$.

- **Policy computation**

$\rightarrow$ PROST (on-line MDP planner [Keller and Eyerich, 2012])

- **Model description**

$\rightarrow$ RDDL (Relational Dynamic influence Diagram Language [Sanner, 2010])



CS = from dataset

RD = random mode

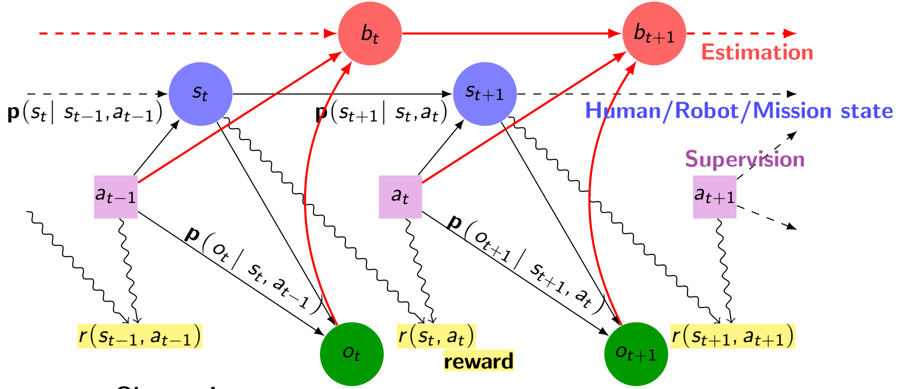
CM = const. manual

CA = const. auto

OP = opt. strat.

Outline

POMDP for driving HMI using human-related observations

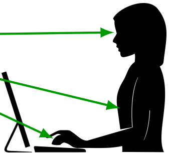


Observations o_t:

Eye tracking (ET)

Electrocardiography (ECG)

Human actions on the interface (HAI)



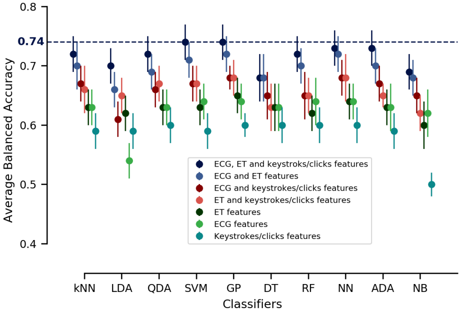
Would the physiological/behavioral data be useful?

A recent work [Chanel et al., 2020a] showed:

High-scoring *vs* low-scoring missions

→ significant differences on behavioral and physiological (ECG, ET) markers were observed
→ robot autonomous mode in high-scoring missions was detrimental to mission performance, while helpful in low-scoring missions

It allowed to train classifiers in order to approach the observation function $p(o|s)$ informing about the joint-performance given the current ET, ECG and HAI markers and robot operation mode.



Let's try it !

Practical work proposition (12h)

- Propose a (simple yet effective) POMDP model to drive the interaction for the Firefighter Robot Game study case.
- Using the data collected with the Firefighter Robot Game:
  - define state, action, and observation sets and transition function.
  - learn the observation and reward functions.
  - proceed with classifiers training to **observe** the human operator engagement/performance based on ET, ECG and HAI data.
  - suggestion : use the classifier's confusion matrix as observation function
- Note that, the resulting policy should decide at each decision time step (every 10s) the robot operation mode to be used.
- Evaluate the resulting policy with simulations of the POMDP model.

References I

Adams, J. A., Rani, P., and Sarkar, N. (2004).
Mixed initiative interaction and robotic systems.
In *AAAI Workshop on Supervisory Control of Learning and Adaptive Systems*, page 613.

Allen, J. E., Guinn, C. I., and Horvtz, E. (1999).
Mixed-initiative interaction.
IEEE Transactions on Intelligent Systems and their Applications, 14(5):14–23.

Chanel, C. P., Roy, R. N., Dehais, F., and Drougard, N. (2020a).
Towards mixed-initiative human–robot interaction: Assessment of discriminative physiological and behavioral features for performance prediction.
Sensors, 20(1):296.

Chanel, C. P., Roy, R. N., Drougard, N., and Dehais, F. (2020b).
Mixed-initiative human-automated agents teaming: Towards a flexible cooperation framework.
In *International Conference on Human-Computer Interaction*, pages 117–133. Springer.

de Winter, J. C. F. and Dodou, D. (2014).
Why the fitts list has persisted throughout the history of function allocation.
Cognition, Technology & Work, 16(1):1–11.

References II

Jiang, S. and Arkin, R. C. (2015).
Mixed-initiative human-robot interaction: definition, taxonomy, and survey.
In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, pages 954–961. IEEE.

Keller, T. and Eyerich, P. (2012).
Prost: Probabilistic planning based on uct.
In *Twenty-Second International Conference on Automated Planning and Scheduling*.

Puterman, M. L. (2014).
Markov decision processes: discrete stochastic dynamic programming.
John Wiley & Sons.

Sanner, S. (2010).
Relational dynamic influence diagram language (rddl): Language description.
Unpublished ms. Australian National University, 32:27.

Williams, K. W. (2004).
A summary of unmanned aircraft accident/incident data: Human factors implications.
U.S. Department of Transportation, Federal Aviation Administration, Civil Aerospace Medical Institute, 4.