# Mixed-initiative planning considering human operator state estimation based on physiological sensors

Nicolas DROUGARD
Caroline CHANEL
Raphaëlle N. ROY
Frédéric DEHAIS

## Context: Human-Machine Interaction (HMI)

Increasing use of automated systems: aircrafts, cars, unmanned vehicles, for military operation, contaminated area, ...

- **Increasingly autonomous robots:** - technical advances in AI, in particular in Machine Learning, Computer Vision & Decision Marking ...

The *Fitt's list* is a compilation of the strenghts of machines (higher computational performance, produce repeatable results, can handle multiple simultaneous tasks, etc.), and humans ones:

- **Human operator still vital:** - capable of complex social interaction and moral judgement, produces tactical and ethical decisions,
  - highly flexible, creative and operate well in dynamic/complex/unknown situations.

⚠ **Human factors involved in** 80% of Autonomous Air Vehicles (AAVs) **accidents** (*Williams 2004*).

*How to combine efficiently human and robot strengths? How to avoid fails due to human factors?*

## Consequences and Idea

**Potential effects of missions on human operators:** stress (danger, pressure), workload (hard, multi tasks), fatigue, boredom (duration), etc.

**Consequences:** confusion, attentional tunneling, mind wandering, lower vigilance, etc. → increase in accidents/mission fails risk

**IDEA: refine supervision of human-robot team** by using data **from physiological sensors** on human operators

**Previous and current works** (DCAS – "Aerial & Spacial vehicle Conception Department"):

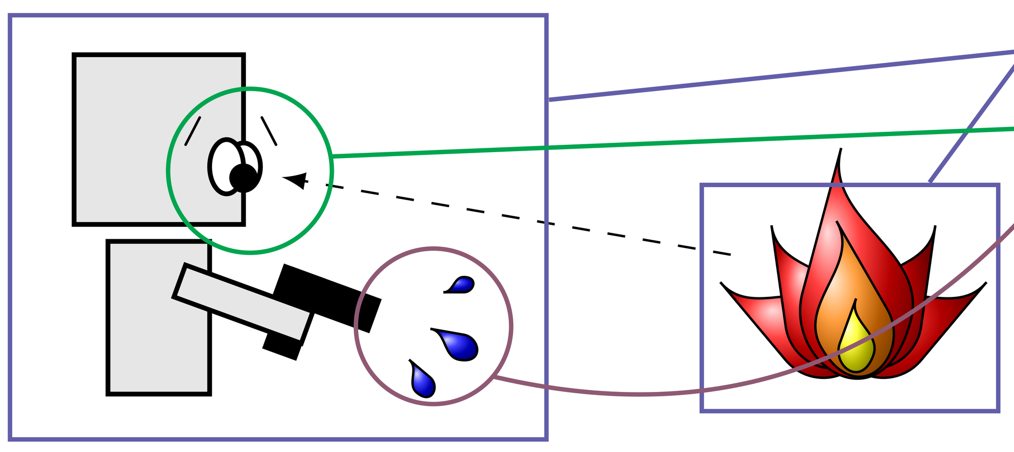
- *Target identification task with a ground robot* (*Souza et al. 2015*)
  → use of Eye Tracking (ET) and ElectroCardioGraphy (ECG)
- *Pilots performing a flight simulator mission* (*Drougard, PhD Thesis 2015*)
  → use of Human Actions on the User Interface (HAUI)
- *Search and rescue task with AAVs* (*Gateau et al. 2016*) → use of ET
- *Current work:* ***firefighter team*** → use of ET + ECG + HAUI

 - Human state estimation: cognitive availability, current human task, type of human behavior (*Nikolaidis et al. 2016*), or any relevant feature: awareness, involvement, etc.

 - Supervision optimization using POMDP/Reinforcement Learning (RL) frameworks.

## POMDP/RL for Robotics

Partially Observable Markov Decision Processes (*Smallwood et al. 1973*)



$s \in \mathcal{S}$: **system states**

$o \in \mathcal{O}$: **system observations**

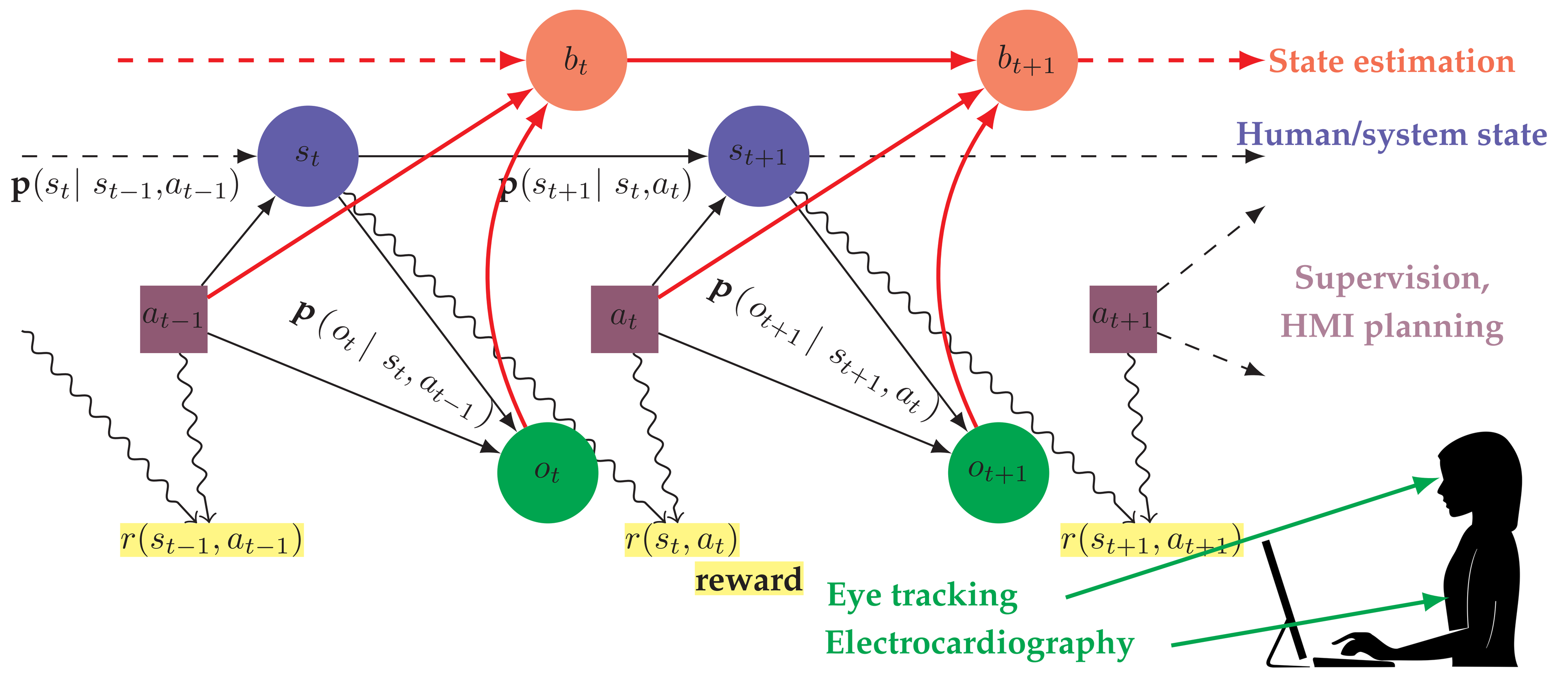
$a \in \mathcal{A}$: **agent's actions**

**transition** probs $\mathbf{p}\,(\, s' \,|\, s, a\,)$

**observation** probs $\mathbf{p}\,(\, o' \,|\, s', a\,)$

**rewards** $r(s,a) \in \mathbb{R}$

## Planning in HMI context (POMDP/RL)




**Belief state** $b_t(s) = \mathbb{P}\,(\, s_t = s \,|\, a_0, o_1, \ldots, a_{t-1}, o_t\,)$

with $s_t \in \mathcal{S}$ describing human operator state

**Strategy** $d_t : b_t \mapsto a_t \in \mathcal{A}$ maximizing $\mathbb{E}\left[\sum_{t\geqslant 0} \gamma^t \cdot r(s_t,a_t)\right]$, $\gamma \in ]0,1[$.

## Mission and User Interface

**Firefighter team mission:** ground robot assisted by a human operator

- **Mission goal: fight fires – some trees are catching fire randomly**

→ reward: $r_t = \#\{\text{safe trees at time } t\}$

→ simple score: use of POMDP/RL & improvement quantification

- **Shared & human-specific actions**: the robot can move, and fight fires autonomously or manually, water source managed only by human


- The robot and its environment are real (ISAE – DCAS) or simulated →
- Multi-task, stressing (fires), and demanding mixed-initiative mission

 → need for a fine and extensive HMI supervision highlighted.

## Design, Optimization & Validation

- **Supervision actions** $a \in \mathcal{A}$ are the mode (autonomous or manual), the countermeasures (visual stimulus) display and robot's actions
- **Model** from an **online video game** (many volunteers without sensors):
  - RL based on rewards → optimization of robot/supervision actions
  - demonstrations dataset → human behavior design
- **Statistical analysis and resulting POMDP**
  - features identification: look for sub-optimal human action sequences, human behavior types identification (by clustering) (*Nikolaidis et al. 2016*)
  - observation probs computation: volunteers equipped with sensors → relation between sensors data $o \in \mathcal{O}$ and human hidden states $s \in \mathcal{S}$
- **Final strategy**: POMDP solving on actions selected by RL
- **Missions comparison using score/rewards:** quantify the improvement using physiological data in planning Human-Robot missions.