Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

Learning adjustable autonomy policies in Multi Human-AI agent Reinforcement Learning

Luca Iocchi
Fabio Patrizi
(and many others)

Sapienza University of Rome

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

MOTIVATION

- Multi-agent systems (Human-AI)
- Human-AI policies: joint state -> joint actions
- Multi-Agent Reinforcement Learning
- Policy repair

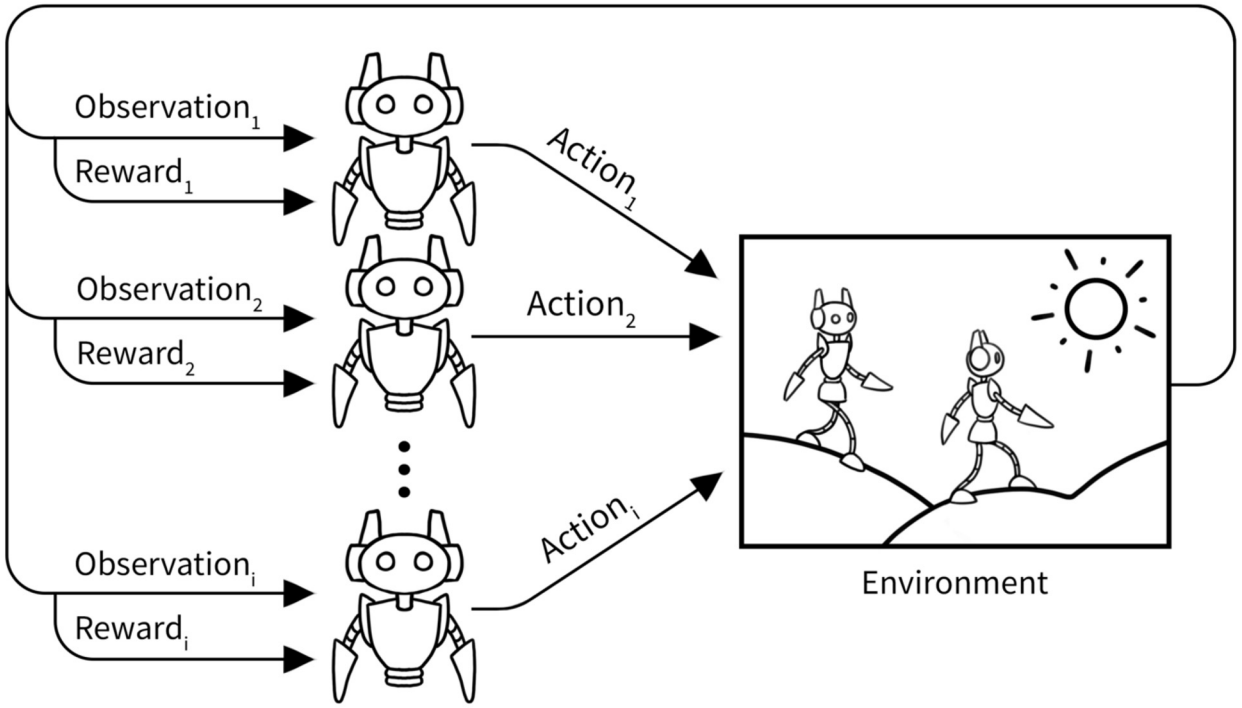
Adjustable autonomy in HAI = HAI policy repair switching responsibility between humans and AI

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

BUILDING BLOCKS

- Human-AI Decision Processes
- Hierarchical Reinforcement Learning
- Model-based Reinforcement Learning
- Trust Factors
- Strategy Repair

MULTI-AGENT DECISION PROCESSES

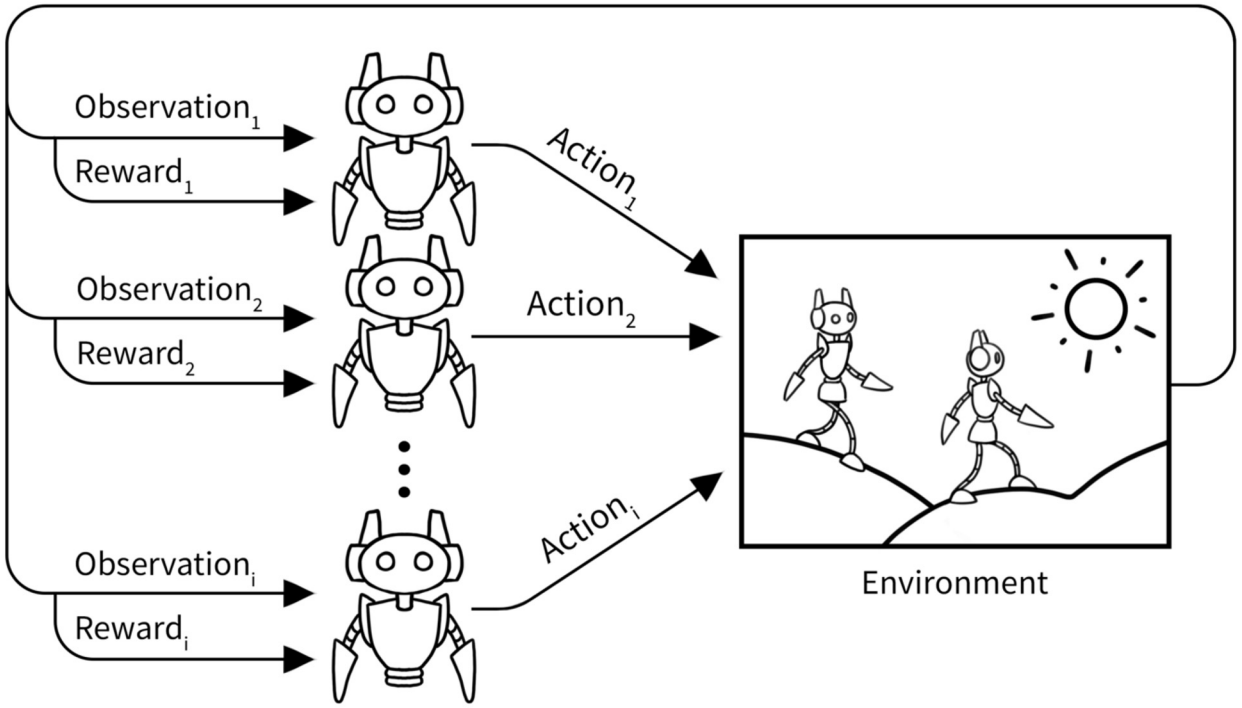


Source

- Markov Games: $(\mathcal{P},\mathcal{S},\{\mathcal{A}\}_i,\mathcal{R},\mathcal{T})$
  - $\mathcal{P}$ set of n agents
  - $\mathcal{S}$ set of state $(s_1,s_2,...,s_k)$
  - $\mathcal{A}$ set of joint actions $(\mathcal{A}_1 \times \mathcal{A}_2 \times ... \times \mathcal{A}_n)$
  - $\mathcal{R}$: $\mathcal{S}\times\mathcal{A}\times\mathcal{S}\rightarrow\mathbb{R}$ reward function
  - $\mathcal{T}$: $\mathcal{S}\times\mathcal{A}\times\mathcal{S}\rightarrow[0,1]$ transition function

- Dec-POMDPs: $(\mathcal{P},\mathcal{S},\{\mathcal{A}\}_i,\mathcal{R},\mathcal{T},\{\mathcal{O}\}_i)$
  - $\mathcal{O}_i$ observation of agent i

HUMAN-AI DECISION PROCESSES



Source

- Human-AI Decision Process: $(\boldsymbol{\Theta}_H, \boldsymbol{\Theta}_A, \boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{R}}, \boldsymbol{\mathcal{T}}, [\boldsymbol{\mathcal{C}}], [\boldsymbol{\mathcal{b}}])$
  - $\boldsymbol{\Theta}_H$ human agents: $(\boldsymbol{\mathcal{S}}_H, \boldsymbol{\mathcal{A}}_H, \boldsymbol{\mathcal{O}}_H)$
    - $\boldsymbol{\mathcal{S}}_H$ set of states (human)
    - $\boldsymbol{\mathcal{A}}_H$ set of actions (human)
    - $\boldsymbol{\mathcal{O}}_H$ observation (human)
  - $\boldsymbol{\Theta}_A$ artificial agents: $(\boldsymbol{\mathcal{S}}_A, \boldsymbol{\mathcal{A}}_A, \boldsymbol{\mathcal{O}}_A)$
    - $\boldsymbol{\mathcal{S}}_A$ set of states (AI)
    - $\boldsymbol{\mathcal{A}}_A$ set of actions (AI)
    - $\boldsymbol{\mathcal{O}}_A$ observation (AI)
  - $\boldsymbol{\mathcal{S}}$ set of state $(s_1, s_2, \ldots, s_k)$ (environment)
  - $\boldsymbol{\mathcal{R}}$: $\boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{A}} \times \boldsymbol{\mathcal{S}} \rightarrow \mathbb{R}$ reward function
  - $\boldsymbol{\mathcal{T}}$: $\boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{A}} \times \boldsymbol{\mathcal{S}} \rightarrow [0,1]$ transition function
  - $\boldsymbol{\mathcal{C}}$: (optional) communication channel
  - $\boldsymbol{\mathcal{b}}$: (optional) belief function

HUMAN-AI POLICY

- Human-AI Policy

  Π: $\mathcal{O}_H$ x $\mathcal{O}_A$ x [$\mathcal{C}$ x $\mathcal{b}$] $\rightarrow$ $\mathcal{A}_H$ x $\mathcal{A}_A$

- Autonomy: balance among actions in $\mathcal{A}_H$ x $\mathcal{A}_A$
- Policy synthesis through MARL
- Policy repair $\Pi_1 \rightarrow \Pi_2$
- Adjustable autonomy: policy repair $\Pi_1 \rightarrow \Pi_2$ with different balance of autonomy
- **Learning adjustable autonomy policies** through
  Model-based MARL on HAIDP for policy repair

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

ON-GOING WORK

Community involvement

- Hierarchical RL
  - Exploiting Multiple Abstractions in Episodic RL via Reward Shaping. **AAAI 2023**
  - Realizable Abstractions: Near-Optimal Hierarchical Reinforcement Learning. Submitted to NeurIPS 2024
  - Deep Abstractions for Tabular Reinforcement Learning. Work in progress

- Model-based RL
  - Model-Based Reinforcement Learning in Discrete Non-Markovian Reward Decision Processes. Submitted to AAAI 2025
  - Multi-Agent Model-Based Reinforcement Learning in Discrete Non-Markovian Reward Decision Processes. 1st **International Workshop on Adjustable Autonomy and Physical Embodied Intelligence (AAPEI 24)** ECAI Workshop.

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

ON-GOING WORK

- HAI MARL
  - Towards computational models for reinforcement learning in human-AI teams. **2nd International Workshop on Multidisciplinary Perspectives on Human-AI Team Trust** (HAI 2023)
  - Modeling a Trust Factor in Composite Tasks for Multi-Agent Reinforcement Learning. **HAI 2024**
  - Developing Targeted Communication through a Trust Factor in Multi-Agent Reinforcement Learning. **3nd International Workshop on Multidisciplinary Perspectives on Human-AI Team Trust** (HHAI 2024)

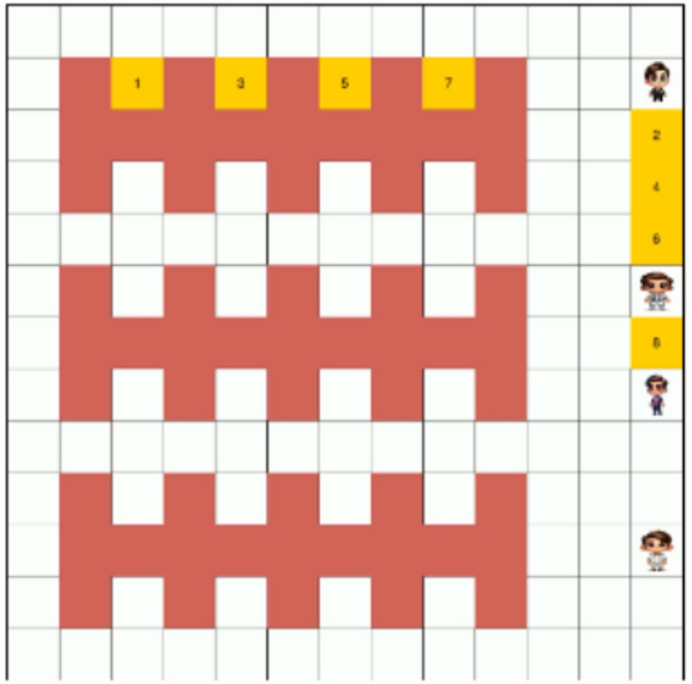
- Policy repair
  - Strategy Repair in Reachability Games. **ECAI 2023**
  - Towards Strategy Repair for Adjustable Autonomy. 1st **International Workshop on Adjustable Autonomy and Physical Embodied Intelligence (AAPEI 24)** ECAI Workshop.
  - Strategy Repair for Probabilistic Games. Submitted to AAAI 2025

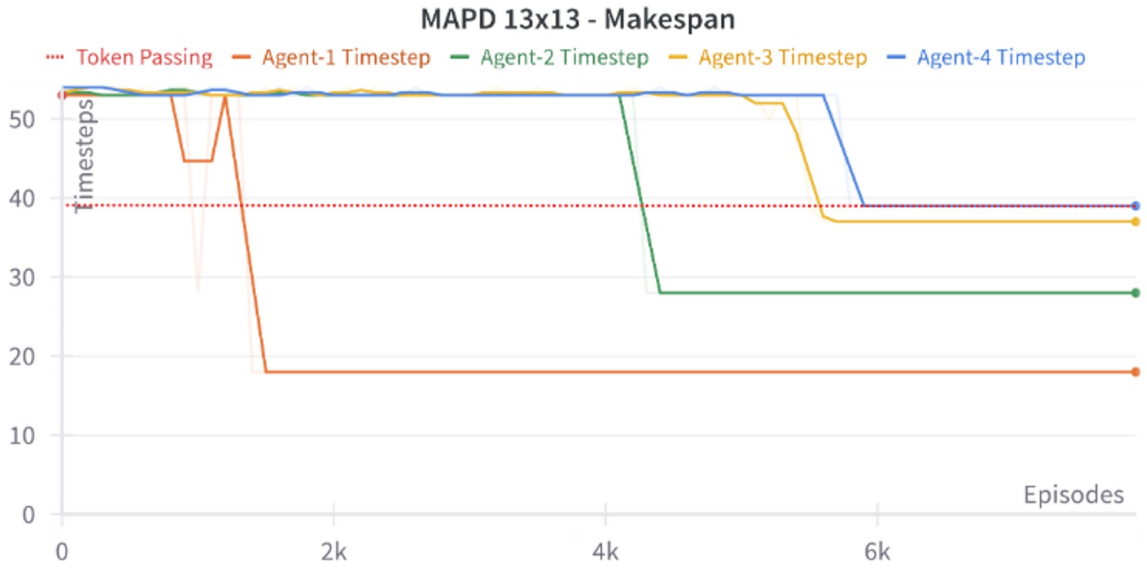
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

RESULTS SO FAR

- Theoretical analysis and experimental evaluation of abstractions in HRL
- Novel algorithms for model-based (MA)RL in NMRDP
- Novel models for representing trust factors in HAI teams
- Novel algorithms for strategy repair

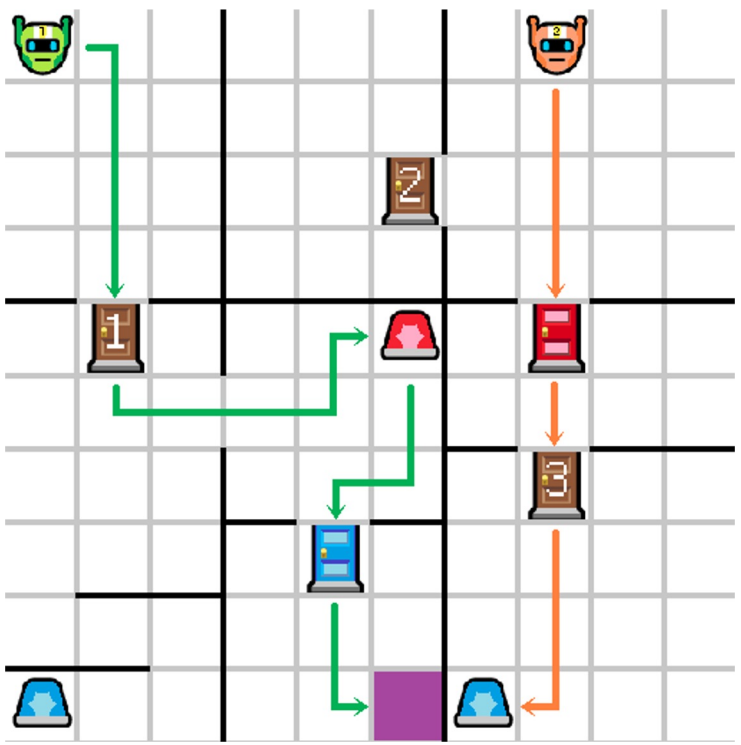
RESULTS (model based MARL)



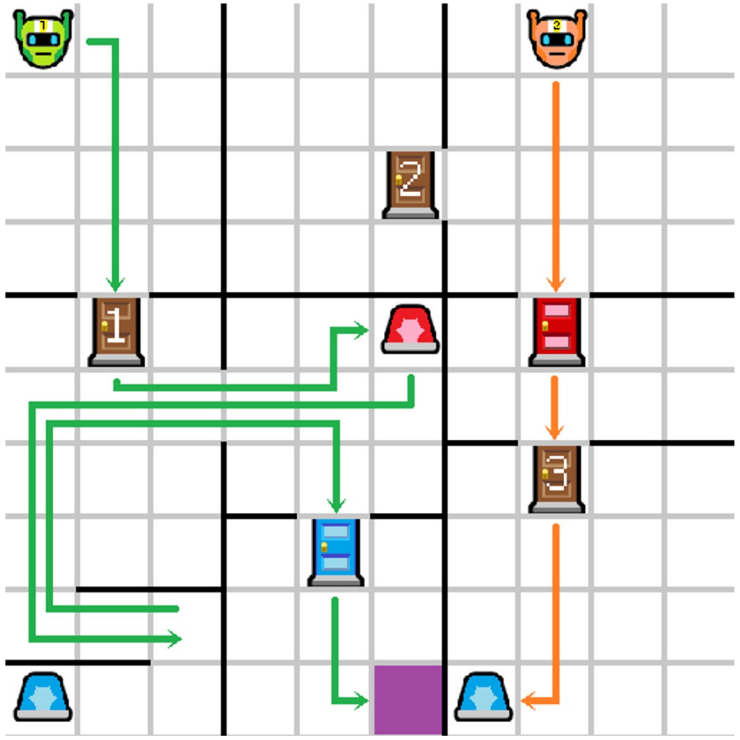


Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

RESULTS (modelling trust factors)



With CTF: green waits for orange to open the door



Without CTF: green tries to reach the blue button

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
FAIR Future Artificial Intelligence Research

CONCLUSIONS

Learning adjustable autonomy in HAI teams

- new perspective (human-centred AI vs. autonomous systems)
- long-term research combining many aspects of the problem
- there is a roadmap !
- community effort (involve more researchers)