Multi-agent learning of ethical behaviours, combining symbolic reasoning and numeric learning

Rémy Chaput

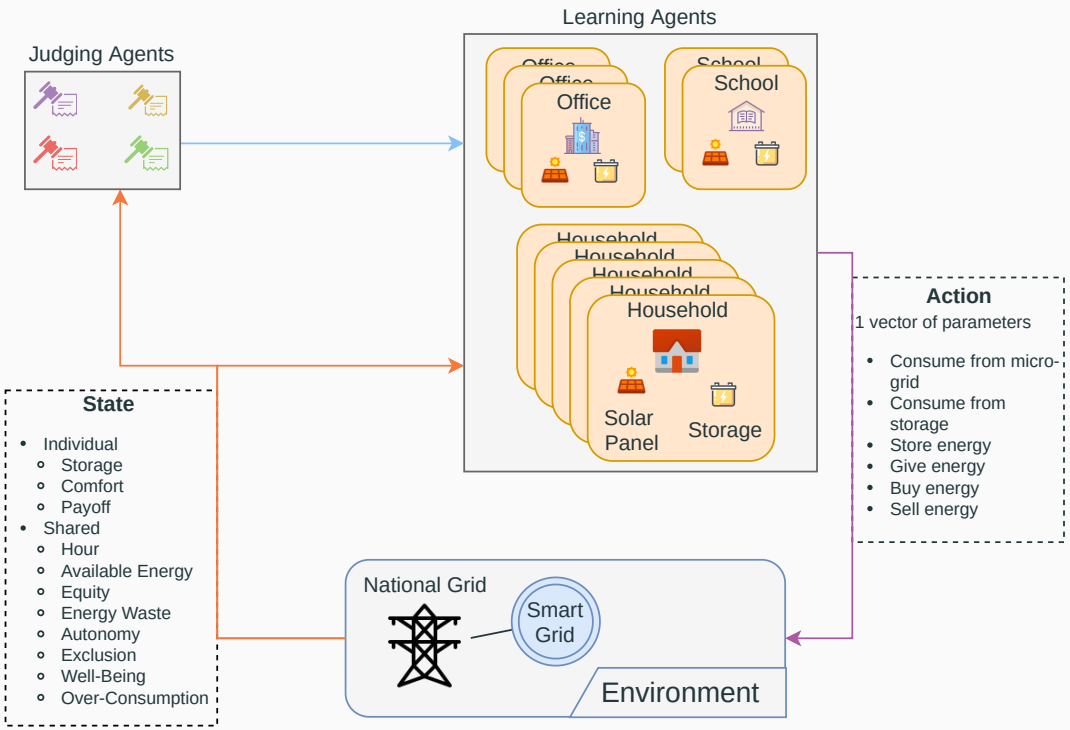
Journée des Thèses - 1er Février 2022

LIRIS - SyCoSMA
Encadré par :
Salima Hassas (LIRIS)
Olivier Boissier (EMSE - LIMOS)
Mathieu Guillermin (UCLy)

Use case

- Learning ethical behaviours : the recent *Machine Ethics* field
- Allocation of energy in *Smart Grids*

- Continuous, multi-dimension perceptions and actions
- Several moral values
  - Well-Being
  - Affordability
  - Equity
  - Ecology
- Several agent profiles

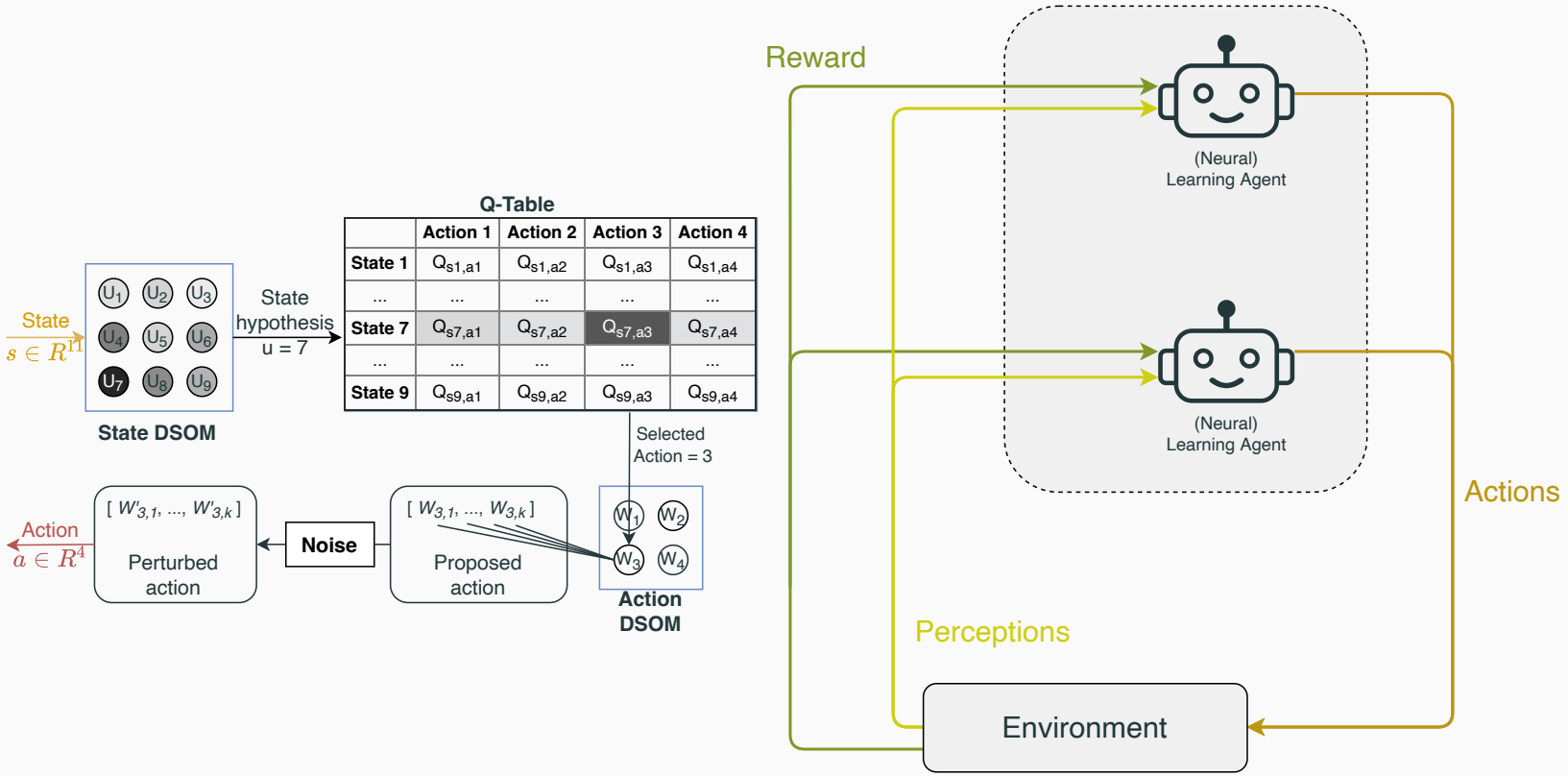


Key points and Contributions

- Multi-Agent
  - Not much regarded in the literature
  - But many “ethical situations” are assessed by interactions and conflicts between agents
  - $\Longrightarrow$ Use-case with multiple agents, 1st algorithm[1]
- Hybrid
  - Combines advantages of learning and symbolic reasoning
  - $\Longrightarrow$ 2nd algorithm[2,3]
- Several moral values
  - Can conflict between themselves
  - $\Longrightarrow$ With aggregation[2,3]
  - $\Longrightarrow$ Explicitly learning dilemmas (current work in progress)
- In interaction with human users
  - Ethics come from humans

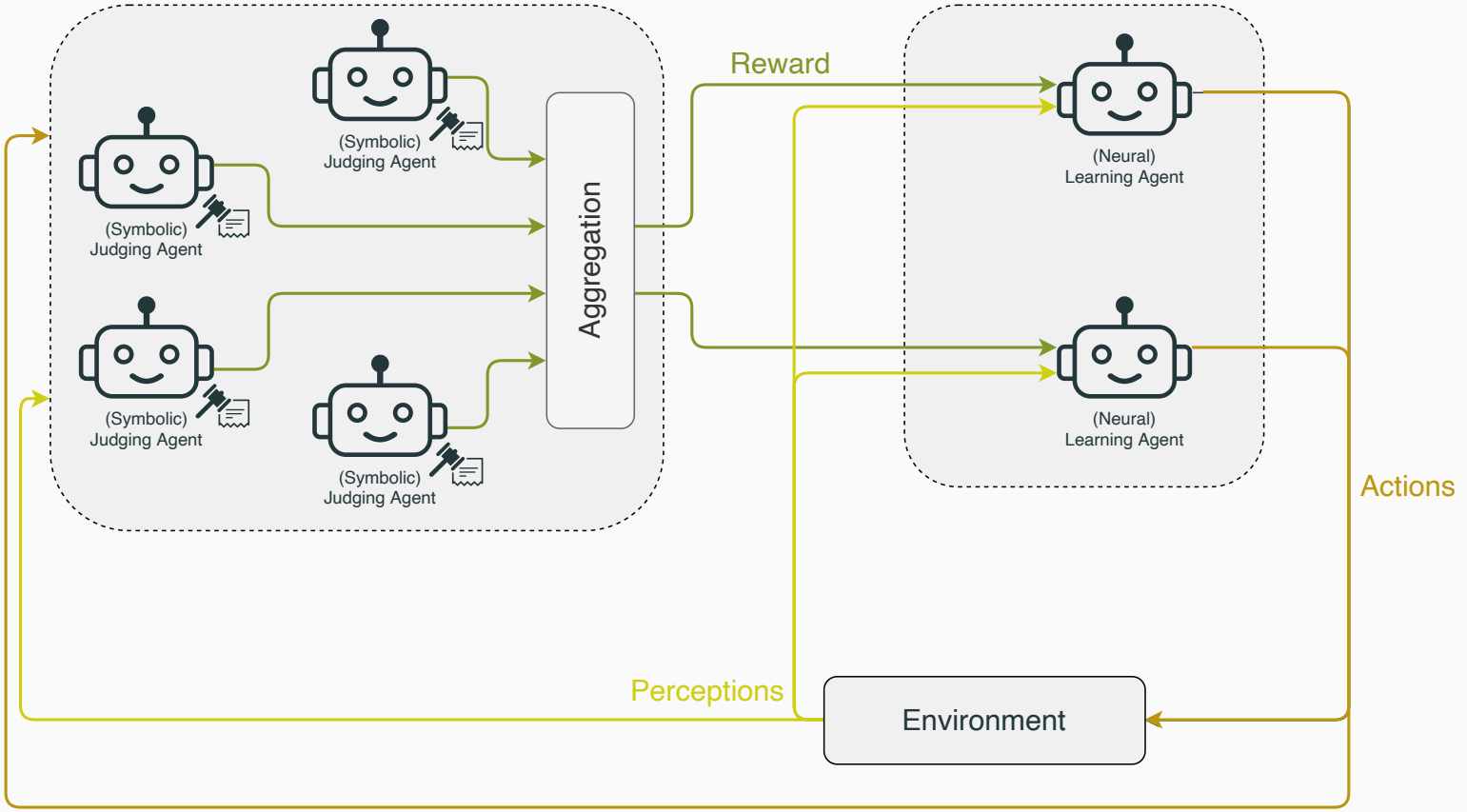
[1]Chaput et al., JFSMA 2020 ; [2]Chaput et al., AIES 2021 ; [3]Chaput et al., JFSMA 2021

Proposed approach



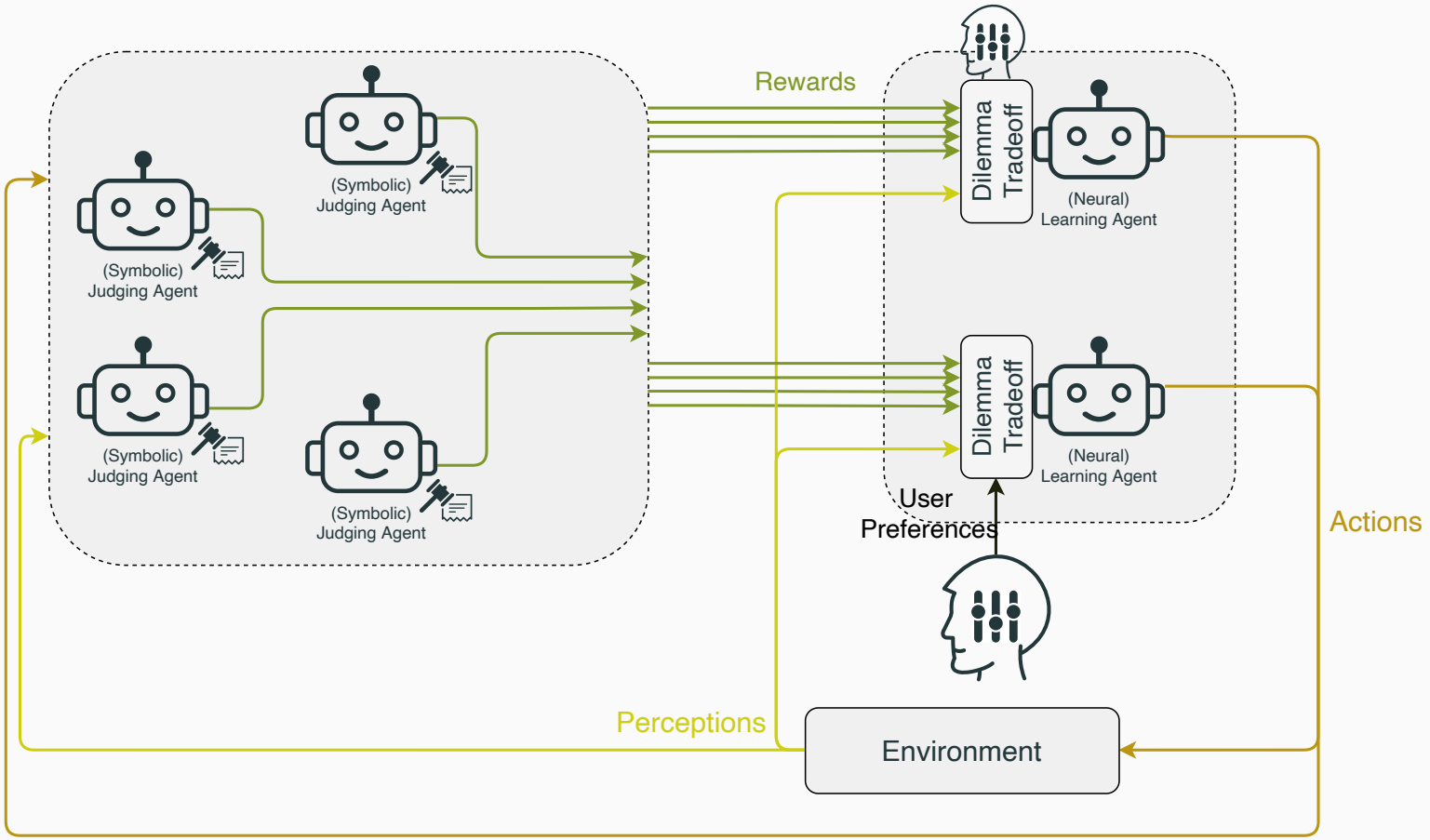
[1] Chaput et al., JFSMA 2020

Proposed approach



[2]Chaput et al., AIES 2021 ; [3]Chaput et al., JFSMA 2021

Proposed approach



Questions?