Action Masking for Safer Model-Free Building Energy Management

Sharath Ram Kumar, Rémy Rigo-Mariani, Benoit Delinchant, Arvind Easwaran

Action Masking for Safer Model-Free Building Energy Management

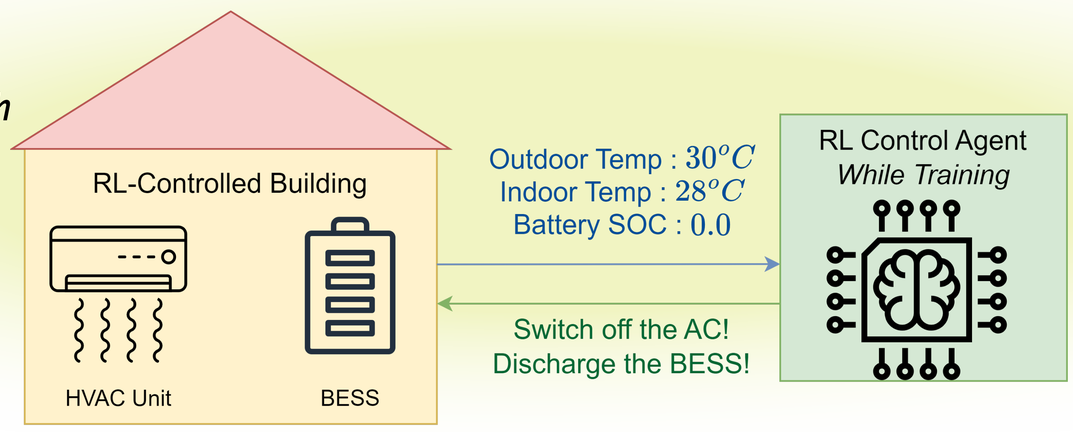
Sharath Ram Kumar[1,2], Remy Rigo-Mariani[1], Benoit Delinchant[1], Arvind Easwaran[2]

[1] *Univ. Grenoble Alpes, CNRS, Grenoble INP, G2Elab, Grenoble, France*
[2] *Nanyang Technological University, Singapore*

REINFORCEMENT LEARNING METHODS ARE NOT PRACTICAL FOR CRITICAL APPLICATIONS

Most RL Algorithms require agents to explore the environment through a trial-and-error approach

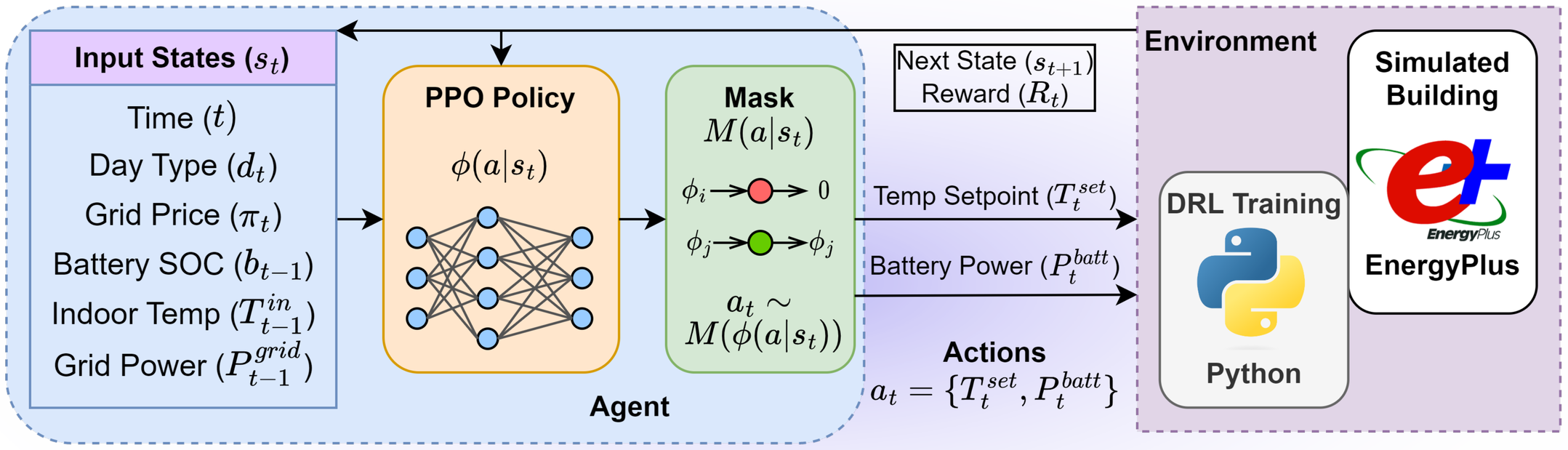
This is not feasible in many real-world control tasks, as unsafe actions may be taken in the learning phase



Problem Statement

To control the cooling HVAC and battery storage systems of a grid-connected building in Singapore, with a trade-off between the electricity bill and the thermal comfort.

ACTION MASKING TO ENFORCE RULES ON THE AGENT



* Cannot charge (discharge) a full (empty) battery
* Cooling system switched off from 10 PM to 5 AM
* Cooling system must stay ON if $T_{indoor} > 26.5^oC$

The agent is trained using PPO, a popular DRL algorithm

The action mask constrains the exploration space by dynamically limiting the actions the agent can take.

MASKED AGENTS CAN OUTPERFORM DIRECT RL AGENTS

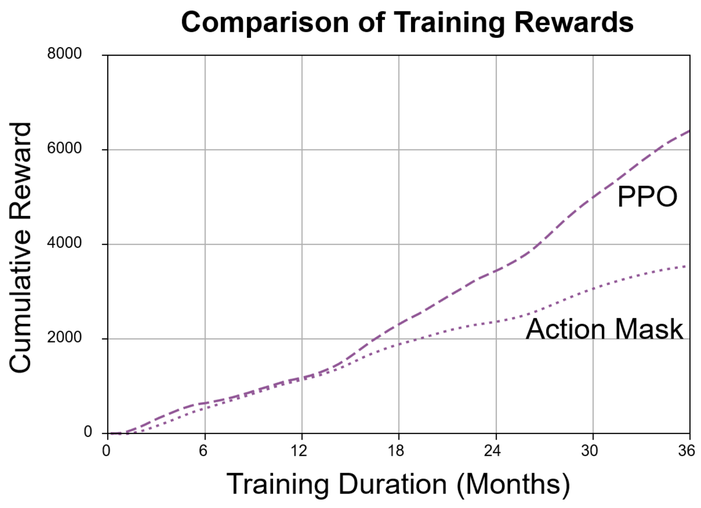


Table 1: Summary of Results

Agent	Cost ($)	Comfort Score
RBC	468.43	0.43
Direct RL	428.89	0.09
Masked RL	439.27	0.39

Key Results

1. Both DRL controllers achieved a lower cost compared to the baseline RBC.

2. The direct RL controller led to a significantly worse comfort score.

3. Action masking achieved a similar comfort score to the baseline while reducing costs.

Conclusions

1. The Direct RL controller prioritized a lower energy bill over thermal comfort (local optima) due to the lack of constraints.

2. The use of Action Masking resulted in a policy that reduced the energy bill while respecting thermal comfort rules, without any modifications to the reward function or hyperparameters.

References and Acknowledgements

[1] Sharath Ram Kumar, Remy Rigo-Mariani, Arvind Easwaran, Benoit Delinchant. 2023. Towards Safe Model-Free Building Energy Manament using Masked Reinforcement Learning. Proceedings of ISGT Europe 2023 (2023)
[2] Ján Drgoňa, Javier Arroyo, Iago Cupeiro Figueroa, David Blum, Krzysztof Arendt, Donghun Kim, Enric Perarnau Ollé, Juraj Oravec, Michael Wetter, Draguna L. Vrabie, and Lieve Helsen. 2020. All you need to know about model predictive control for buildings. Annual Reviews in Control 50 (2020), 190–232.
[3] Shengyi Huang and Santiago Ontañón. 2020. A Closer Look at Invalid Action Masking in Policy Gradient Algorithms. (2020).
[4] Karl Mason and Santiago Grijalva. 2019. A Review of Reinforcement Learning for Autonomous Building Energy Management.
[5] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal Policy Optimization Algorithms. (07 2017).

This work was supported by the National Research Foundation, Prime Minister's Office, Singapore, under its CREATE programme.