Reinforcement Learning for Sizing and Operation

Adrien Bolland (adrien.bolland@uliege.be)

June 22, 2023

Limitations of mathematical Programming

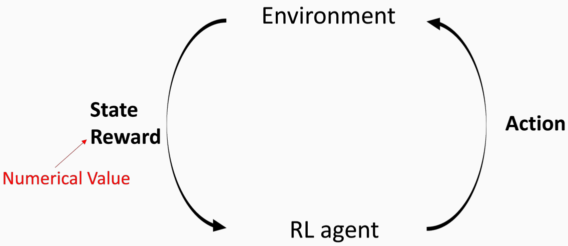
Mathematical programming suffers from some limitations :

- Requires an analytical model of the problem
- Time-consuming to solve nonconvex problems
- Hard to account for uncertainty

In practice, it may not be suited for online control and complex design problems.

Reinforcement Learning

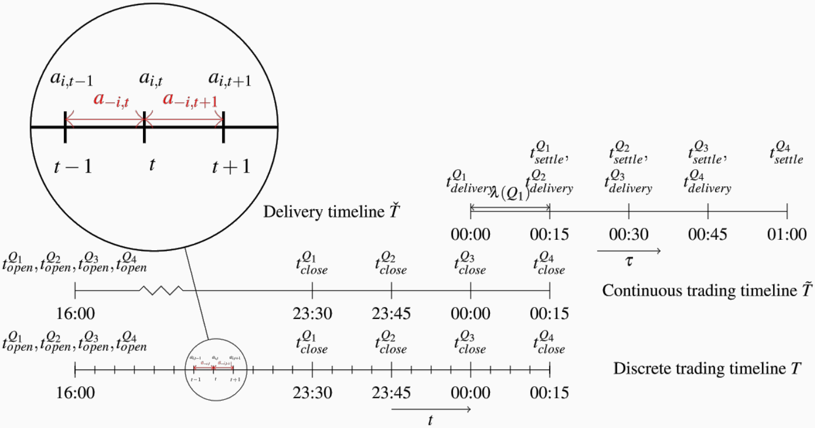
Reinforcement learning agents make decisions in a system based on the observed states in order to maximize the reward gathered.



Reinforcement learning

- Requires an oracle model
- Differentiates between optimization and execution time
- Solves offline a nonconvex stochastic optimization problem

An agent can refine its position by buying or selling electricity on the market.



CID and RL for Controlling a Battery

We assume having a battery that we can charge and discharge, and we want to use this asset for trading on the CID.

1. State space : physical state of charge of the battery and features of the current order book
2. Action space : choice between being greedy, i.e., refining the position with the currently available orders, or waiting for future orders
3. Reward : revenue at each market period

CID and RL for Controlling a Battery – In practice

We have built an environment from historical market data.

When taking an action, a linear program computes the optimal greedy position. The choice between the two actions is determined by a recurrent neural network.

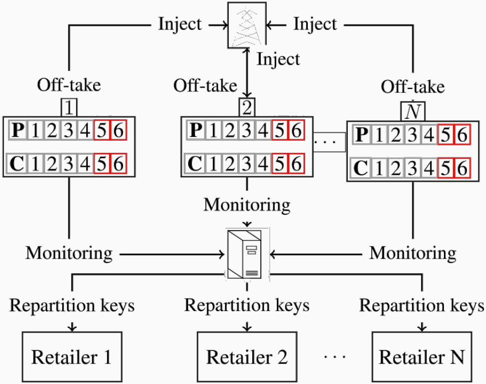
Two RL policies are computed, with fitted Q iteration and deep Q network, and compared to the rolling intrinsic policy that greedily refines it position at each time step.

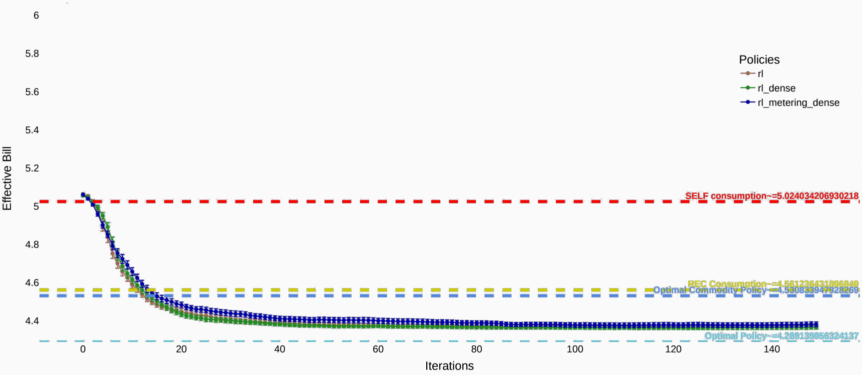
CID and RL for Controlling a Battery – Results

	π^{FQ}		$\pi^{APEXDQN}$	
	V (€)	r (%)	V (€)	r (%)
Mean	667.9	3.8	669.1	3.9
Min	153.7	– 26.7	187.9	– 9.4
25%	490.9	– 0.7	501.0	0.4
50%	649.9	4.0	632.3	3.3
75%	814.1	9.9	772.0	7.1
Max	1661	40.9	1471.4	19.9
Sum	102,937	–	101,708	–

The RL policies outperform by 4% the rolling intrinsic policy.

A group of prosumers are interconnected and can exchange electricity within a local market that is more profitable than the retailer. Each prosumer has to control when to produce and consume while accounting for physical and market constraints.





Preliminary results indicate that RL is a viable solution that challenges model predictive control techniques.

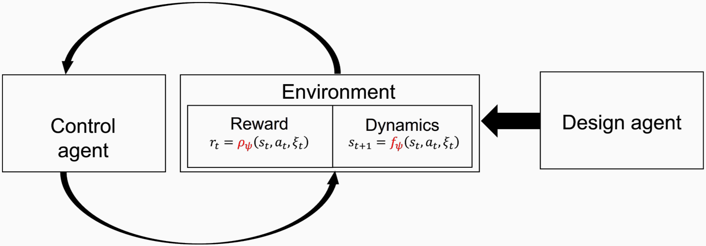
Joint Design and Control

In the current framework, the model is fixed in RL !

In many engineering problems, the environment that is controlled can also be designed.

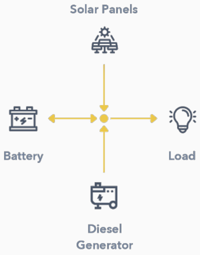
Joint Design and Control

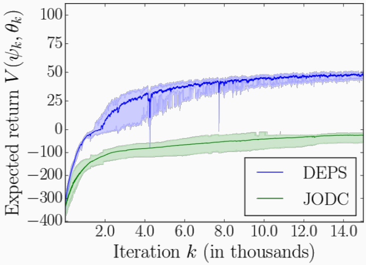
We introduce environment parameters that shall be jointly optimized with the policy to maximize the return.



Two agents are learned, one decides the design, the other the actions.

Joint Design and Control





- Design / Investments : capacity of the production units and battery
- Control : power output of the units
- Objective : minimize the total costs

We manage to extract optimal designs and optimal policies with RL.

Conclusion

Reinforcement learning is a promising and powerful tool for sizing and operation !

References

Ioannis Boukas, Damien Ernst, Thibaut Théate, Adrien Bolland, Alexandre Huynen, Martin Buchwald, Christelle Wynants, and Bertrand Cornélusse. A deep reinforcement learning framework for continuous intraday market bidding. *Machine Learning*, 110:2335–2387, 2021.

Adrien Bolland, Ioannis Boukas, Mathias Berger, and Damien Ernst. Jointly learning environments and control policies with projected stochastic gradient ascent. *Journal of Artificial Intelligence Research*, 73:117–171, 2022.

Marine Cauz, Adrien Bolland, Bardhyl Miftari, Lionel Perret, Christophe Ballif, and Nicolas Wyrsch. Reinforcement learning for joint design and control of battery-pv systems. *36th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental impact of energy systems*, 2023.