

GYM4REAL: A BENCHMARK SUITE FOR EVALUATING REINFORCEMENT LEARNING IN REALISTIC DOMAINS

D. SALAORNI, V. DE PAOLA, S. DELPERO, G. DISPOTO,
P. BONETTI, A. RUSSO, G. CALCAGNO, F. TROVÒ,
M. PAPINI, A. M. METELLI, M. MUSSI AND M. RESTELLI
davide.salaorni@polimi.it

MOTIVATION

Can we **really** trust the widely used RL benchmarks when it comes to solving control problems in **real-world safety-critical scenarios**?

DAMENV

Goal: plan the water release to satisfy the daily water demand while preventing the reservoir from exceeding its maximum capacity and causing overflows.

$$\min \sum_{t=1}^T \left(\sum_{f=1}^F w_{f,t} + c_t \right)$$

Simulator: mathematical model of a water reservoir system influenced by a configurable lake model.

Data: historical daily records of water level, demand, and inflow in the period from 1946 to 2010.



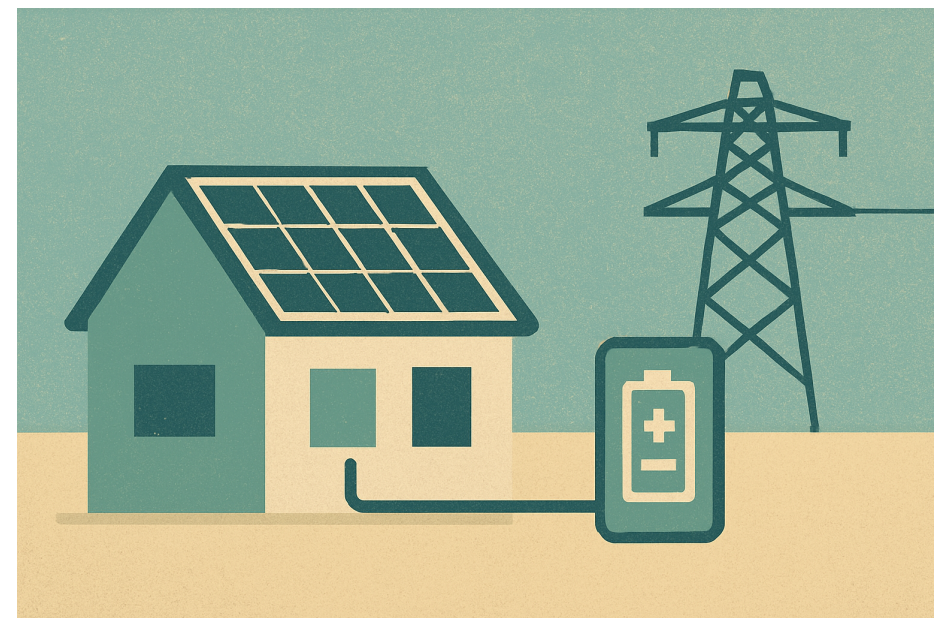
MICROGRIDENV

Goal: optimize the energy management by minimizing trading and battery costs.

$$\max \sum_{t=1}^T [r_{\text{trad}}(a_t) + r_{\text{deg}}(a_t)]$$

Simulator: digital twin of a battery integrated with exogenous realistic signals.

Data: real-world data of renewable energy generation, consumption and market prices.



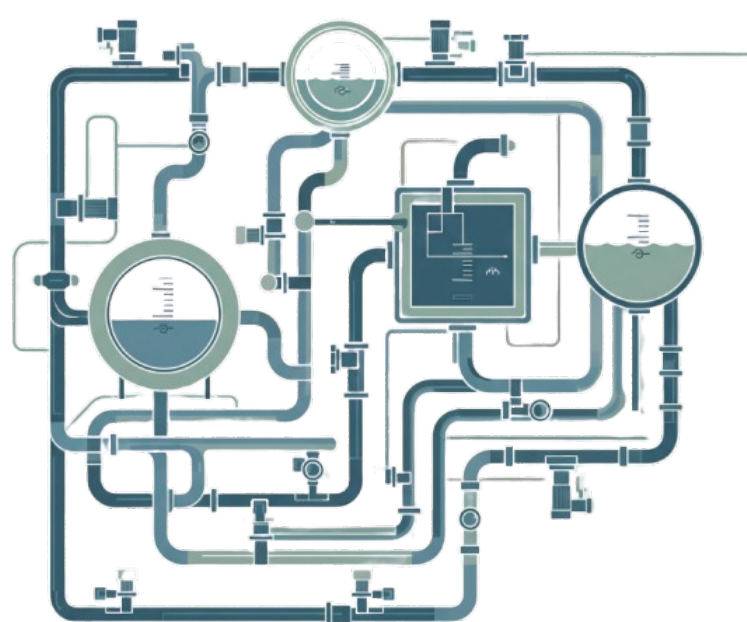
WATERDISTRIBUTIONSYSTEMENV

Goal: foster network resilience by maximizing the DSR while minimizing the overflow risk.

$$\max \sum_{t=1}^T [r_{\text{DSR}}(a_t) + r_{\text{of}}(a_t)]$$

Simulator: hydraulic analysis framework Epanet providing mathematical solvers.

Data: synthetically generated residential water demand datasets using STREaM.

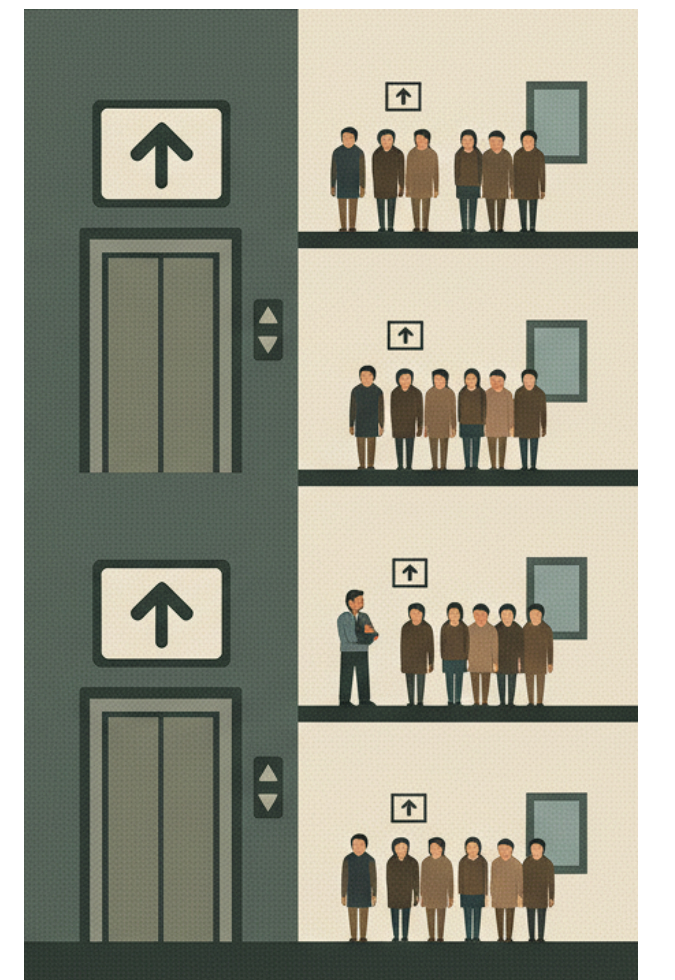


ELEVATORENV

Goal: minimize the cumulative *waiting time* of all transported passengers within a time period

$$\min \sum_{t=1}^T \left(\sum_{f=1}^F w_{f,t} + c_t \right),$$

Data: passenger arrival profiles are synthetically generated using a Poisson process, with the arrival rate uniformly sampled for each floor independently.

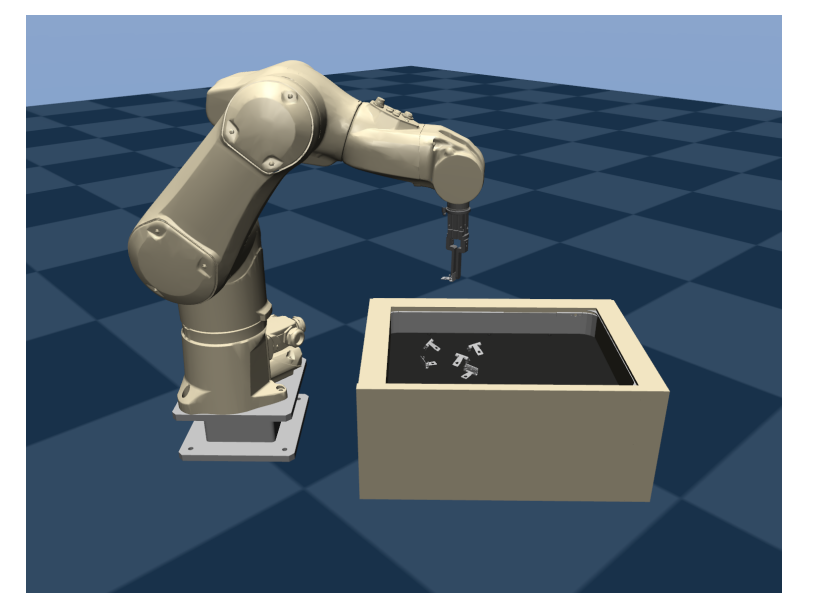


ROBOFEEDERENV

Goal: hierarchical task aimed to

1. determining the *picking order* of the objects
2. identifying the precise *grasping point* on each object for successful pickup and placement

Simulator: Mujoco robotic world + top-bottom robot viewpoint.



TRADINGENV

Goal: learn a profitable intra-day strategy dealing with non-stationary markets and low signal-to-noise ratio.

Data: historical foreign exchange time-series relative to the EUR/USD currency pair.



Paper



GitHub

