OIST

Distributed Reinforcement Learning for DC Open Energy Systems

2023/5/4

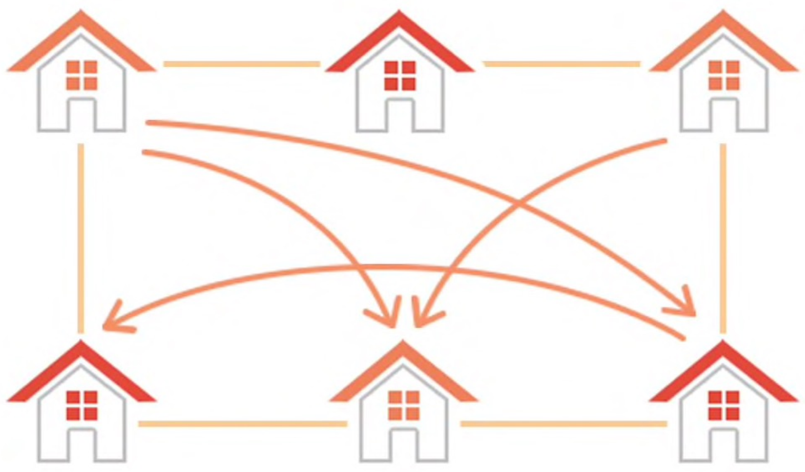
Qiong Huang, Kenji Doya

Neural **C**omputational **Unit**

Tackling Climate Change with Machine Learning, ICLR 2023

DC-Based Open Energy System (DCOES)

- It focuses on re-defining the conventional electricity grid systems.
- A **bottom-up** distributed electric power system which allows self-determined energy **exchanges** between residential nodes within a local DC microgrid **community**.



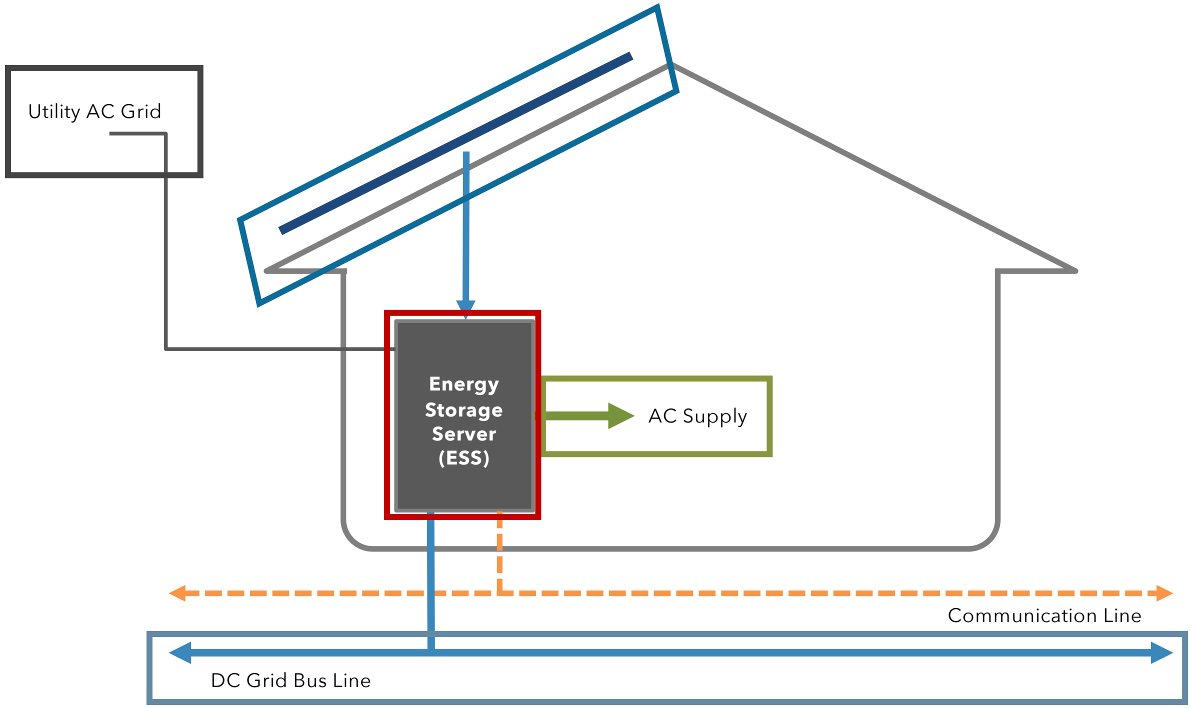
DCOES Power System

- **Sustainable**
  - Renewable energy sources
  - Local Storage (Battery)
- **Dependable**
  - Keep risk manageable
  - Local consumption
- **Affordable**
  - Accessible to everyone on earth
  - Start small yet expandable

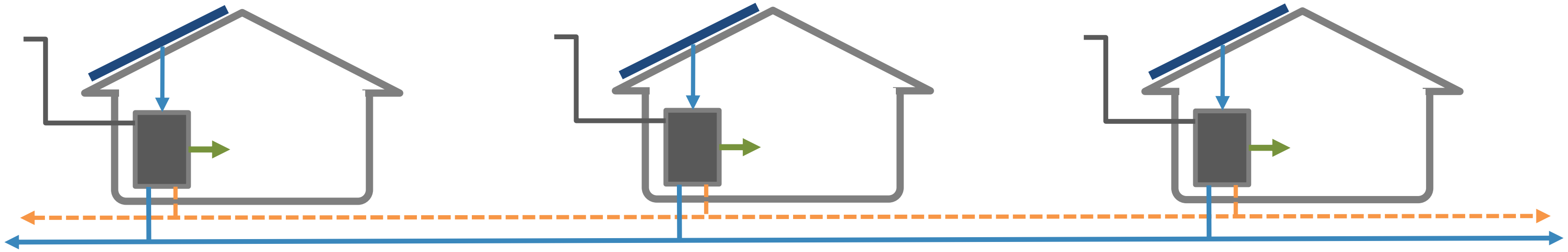
- Adapt from: T. Sakagami et al. Performance of a DC-based microgrid system in Okinawa. 2015 ICRERA

DC Open Energy System on OIST Campus

- Configuration of each house



OIST Faculty House Area
19 houses in total

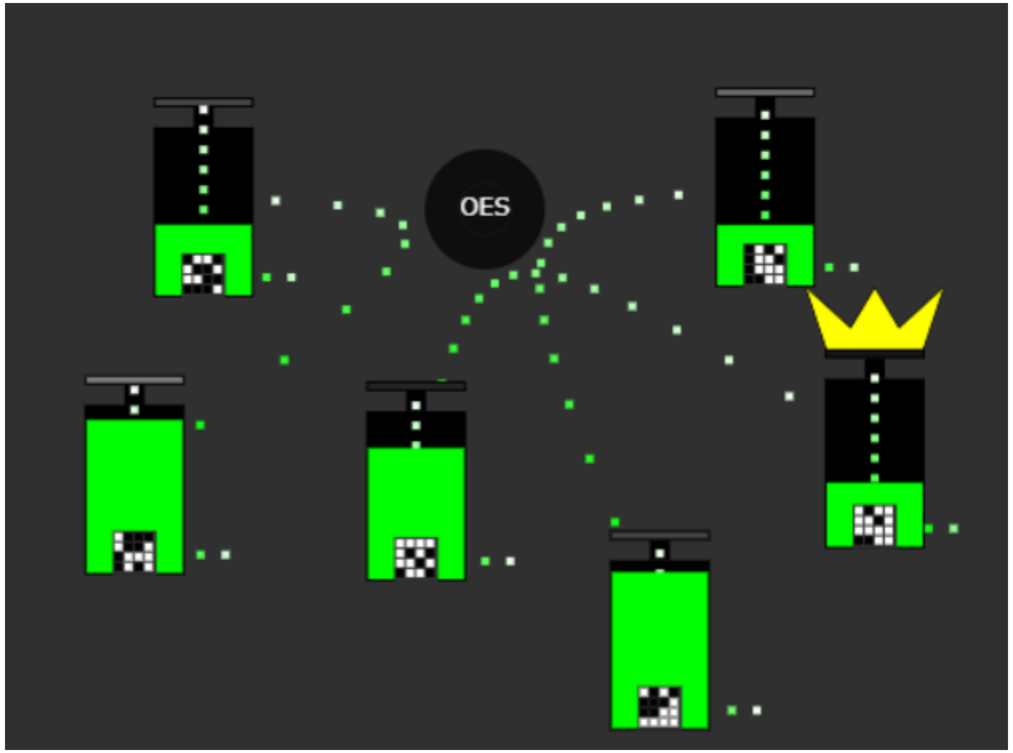


Autonomous Power Interchange System (APIS)

- Key software for simulating energy exchange

APIS System

- apis-main
  - scenarios
  - config files
- apis-main-controller
- apis-emulator (can run alone)
  - data
  - core
- apis-web
- . . .



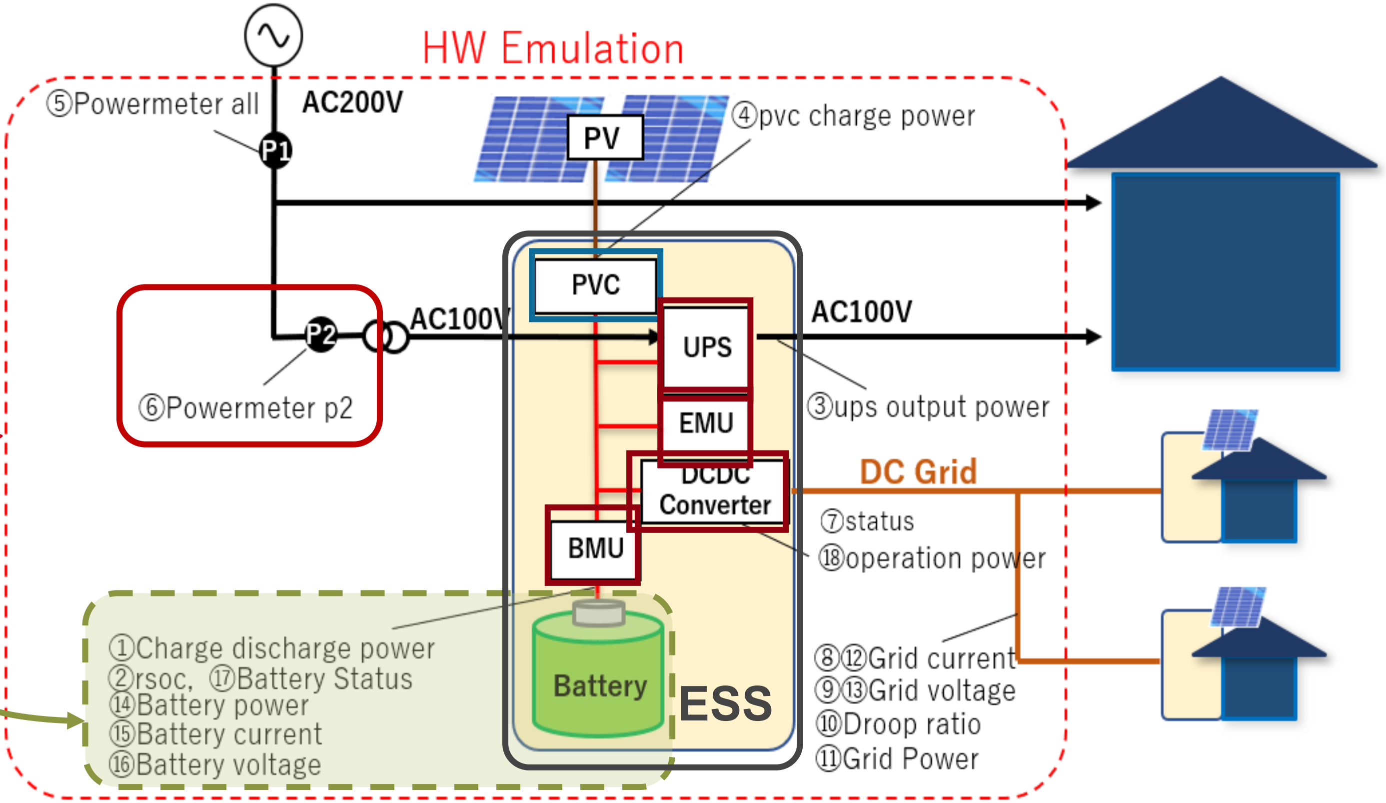
- Adapt from: Sony CSL, Inc. APIS https://github.com/SonyCSL/APIS, 2020, Dec.

Autonomous Power Interchange System (APIS)

- Key software for simulating energy exchange

APIS Simulator

- apis-main
  - scenarios
  - config files
- apis-main-controller
- apis-emulator (can run alone)
  - data (input, output)
  - core (power flow updates)
- apis-web
- . . .



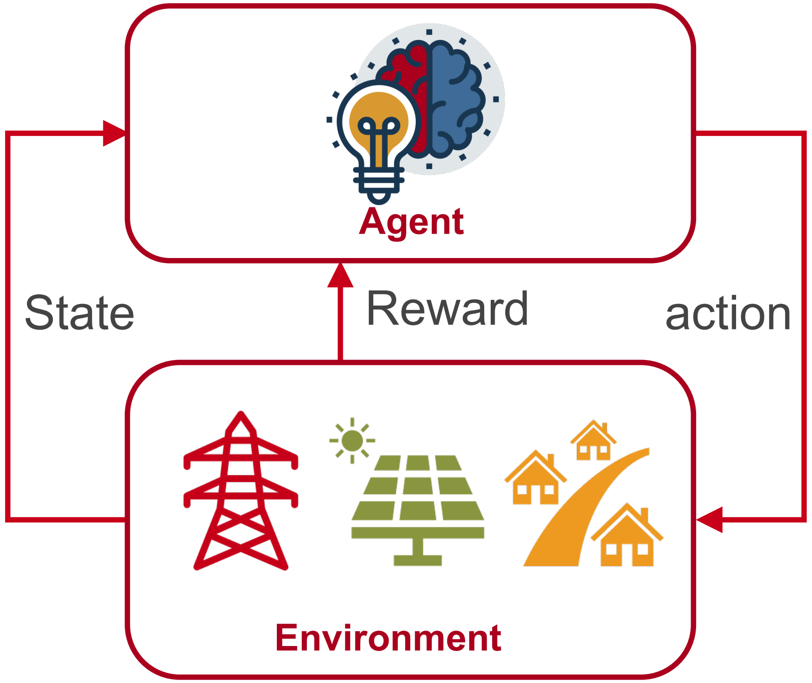
- Adapt from: Sony CSL, Inc. APIS https://github.com/SonyCSL/APIS, 2020, Dec.

Problem Statement and Challenges

- Could reinforcement learning methods work in energy management?
- Which reinforcement learning methods and components are to be considered?
- For multiple houses, current DCOES use a fixed rule-based control on energy exchanges.
  - How to combine the controller with reinforcement learning?
  - Can reinforcement learning agents do better?
- What state-action representations could influence the performances?

Reinforcement Learning Background

Diagram of Reinforcement Learning

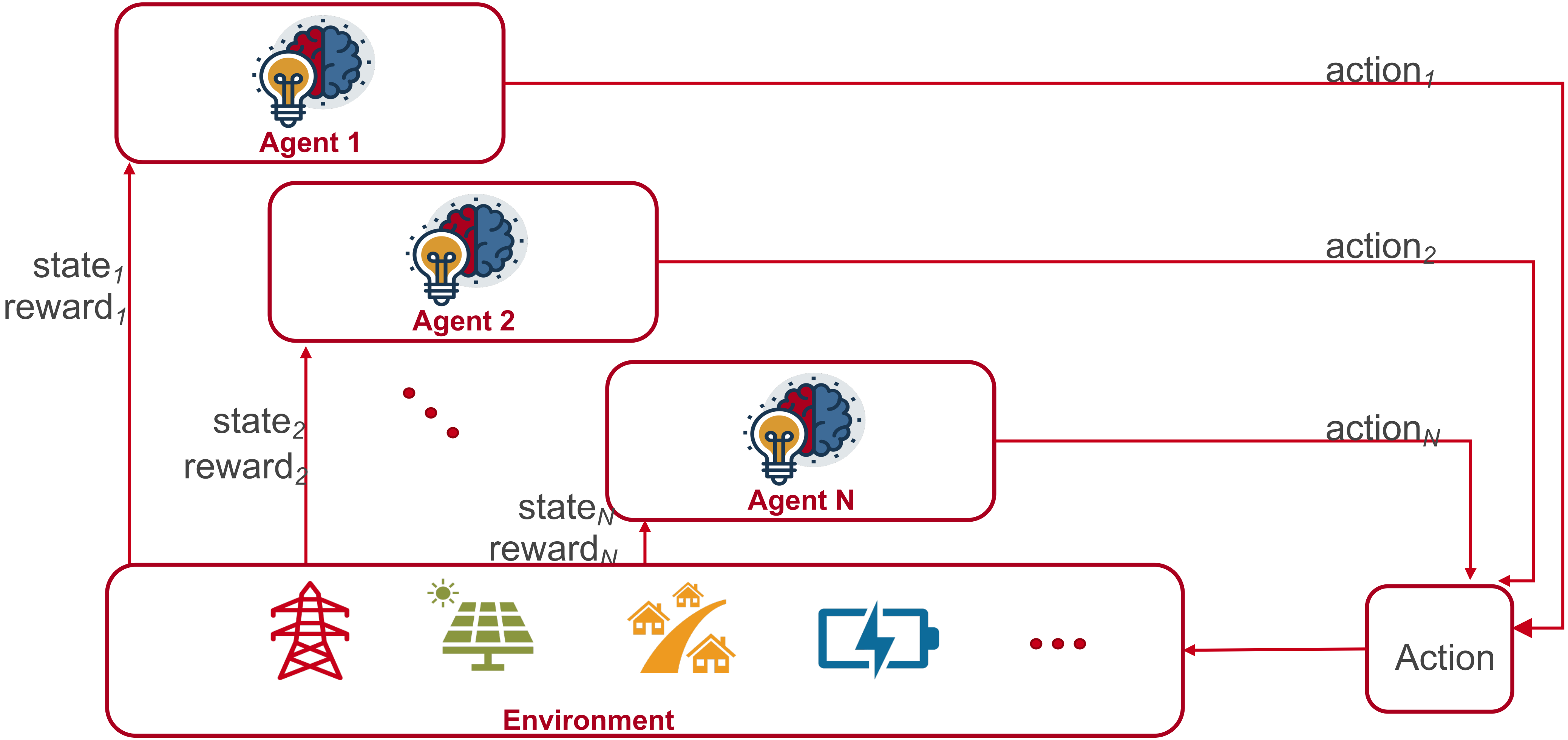


Learns policy to maximize reward

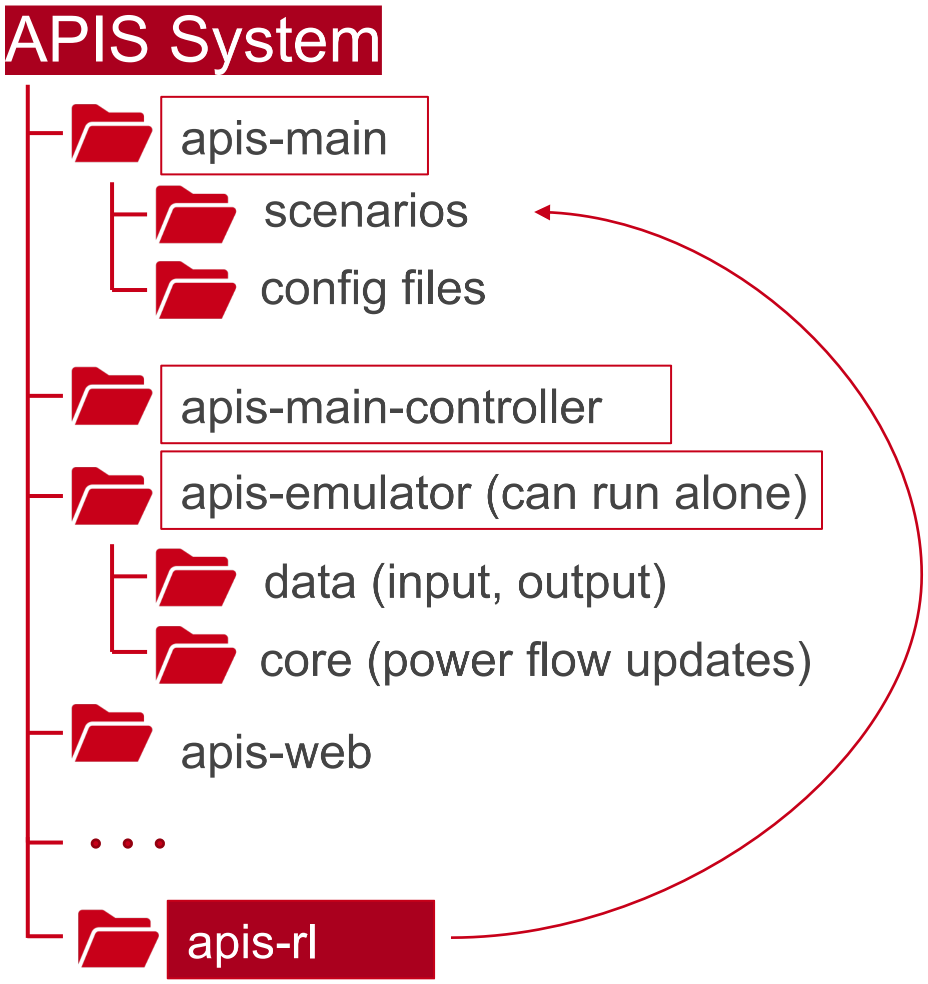
$$Q^*(s,a) = max_\pi E[\sum_t^T \gamma^t r_k | s_t = s, a_t = a, \pi]$$

- Markov decision processes (MDP) components
  - **State set**
    - Relative state of charge (RSOC)
    - Weather (solar radiation, wind direction/speed)
    - Load
    - Time
    - …
  - **Action set**
    - Discharge / Charge / Idle
    - Send / Receive
    - …
  - **Environmental dynamics**
    - Power production, house usage
    - Battery dynamics (RSOC)
    - …
  - **Reward**:
    - External purchased energy (p2)
    - RSOC level
    - Peak of load
    - …

Multi-agent Reinforcement Learning



Reinforcement Learning in APIS



apis-rl software

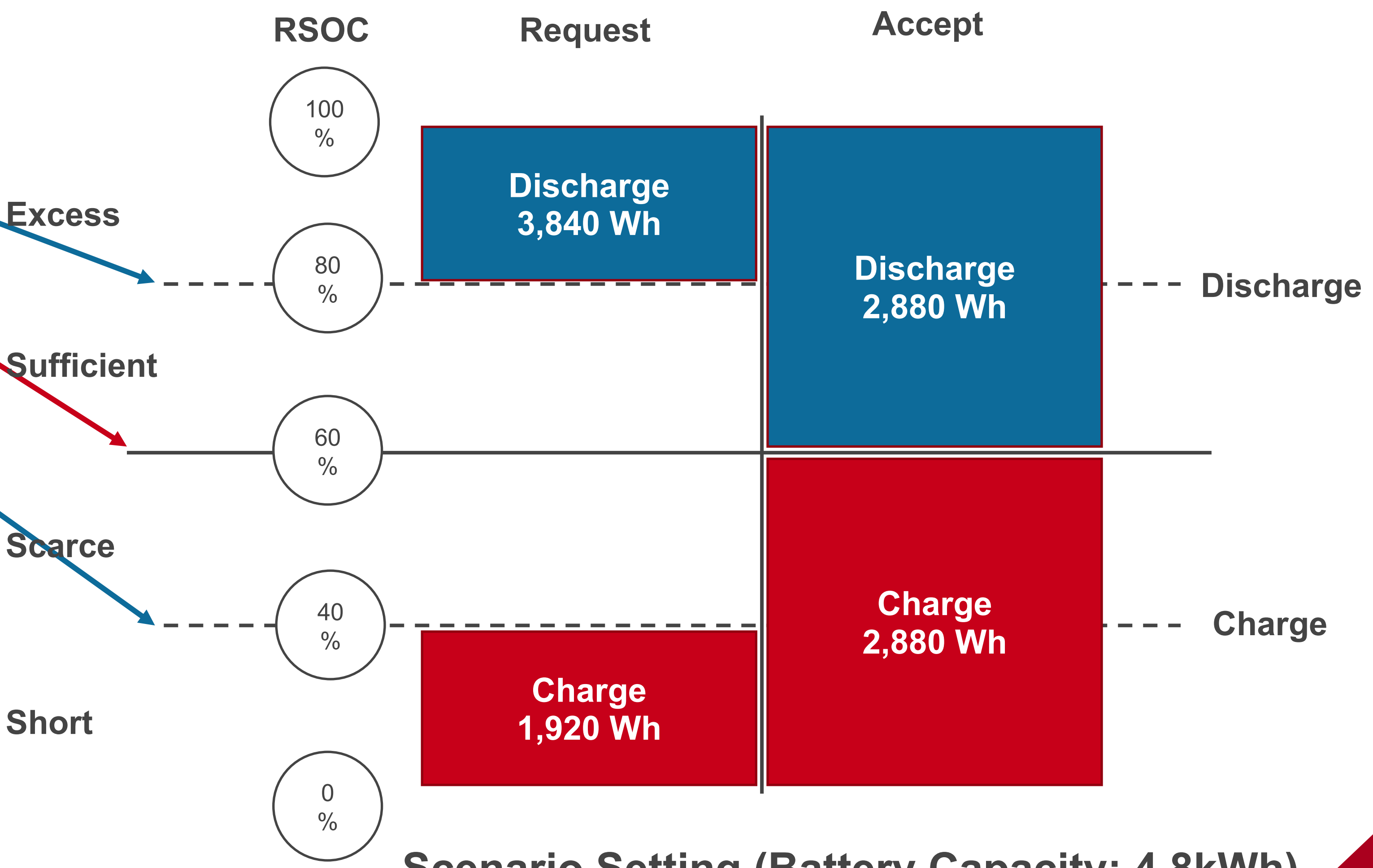
- Learn policy to control the sending/receiving thresholds with respects to RSOC

Goal

- Minimize the purchase power from the powerline (power supply company)

States, Action, Reward Settings (For Each House)

- States (possible variable)
  - pvc, load, p2, rsoc, rsoc_ave, ig (dcdc_exchange_current), time (hour_sin, hour_cos)
- Actions
  - 8 levels : [0.2, 0.3, …, 0.9]
  - action_request_discharge
  - action_accept_dis/char
  - action_request_char
- Reward
  - - p2 (acin power)
- Done
  - Update each day



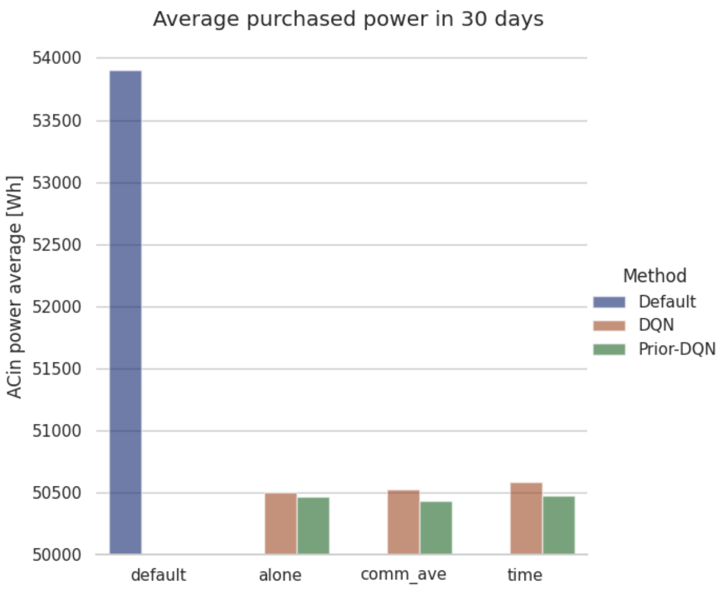
Scenario Setting (Battery Capacity: 4.8kWh)

Tested Cases in Multi-agent RL

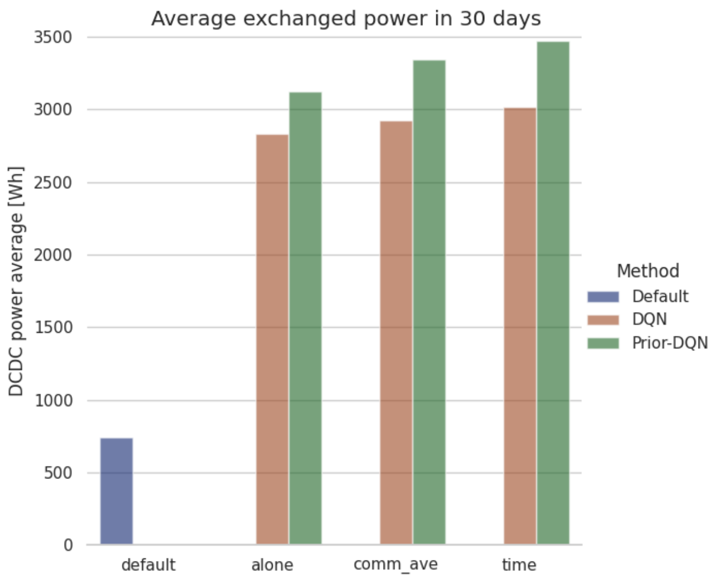
- Using 4 houses
  - House 212, 213, 214, 215 in Route B
  - One month data in May, 2019
- Different options of state
  - Stand alone
    - pvc, load, p2, rsoc
  - Community average
    - pvc, load, p2, rsoc, rsoc_ave, dcdc_exchange_current
  - Time-of-day information
    - pvc, load, p2, rsoc, rsoc_ave, dcdc_exchange_current, time (hour_sin, hour_cos)
- DQN and Prior-DQN
- Reward settings
  - Individual purchased power
  - Sum purchased power

DQN vs. Prioritized DQN

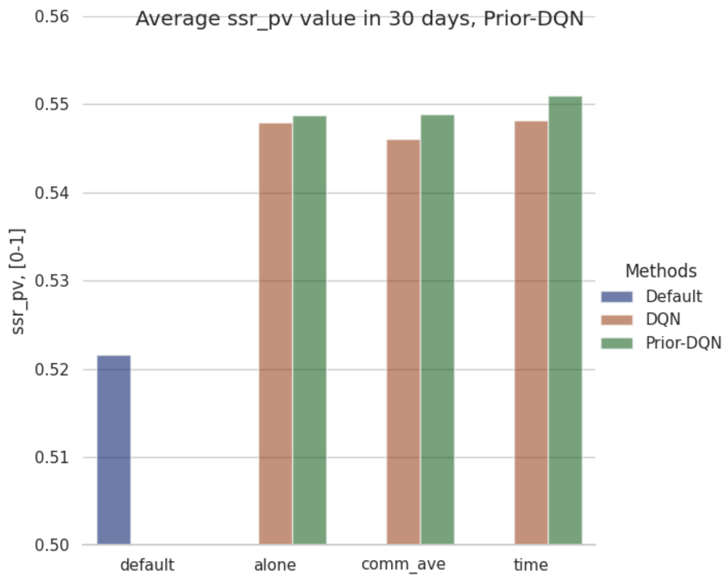
Purchased power



Exchanged power



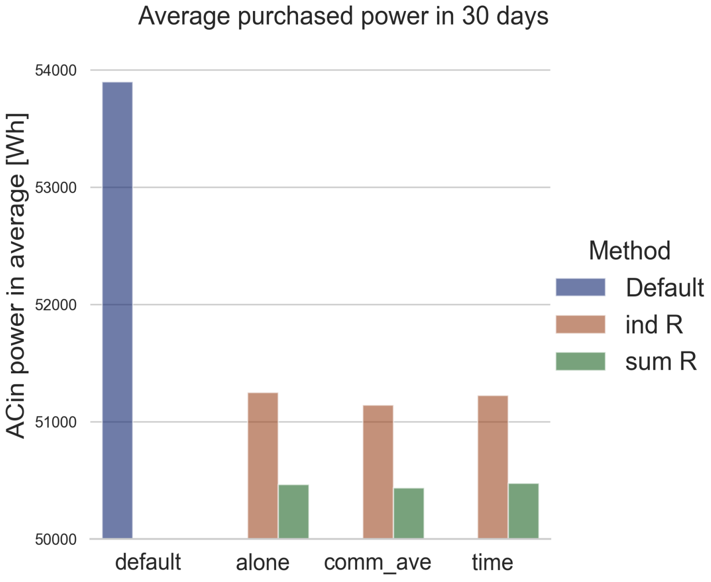
SSR



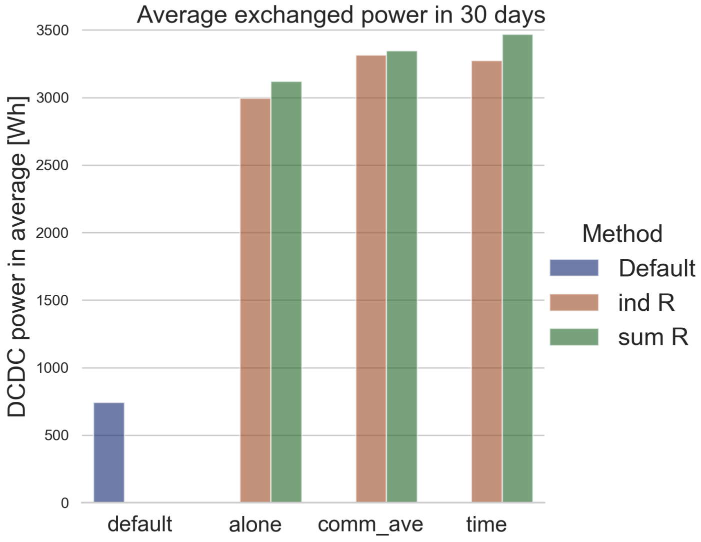
- Purchased power decreases in all training conditions
- Exchanged power increases in all training conditions
- SSR improves in all training conditions
- Prioritized DQN performances better than DQN

Individual vs. Sum reward

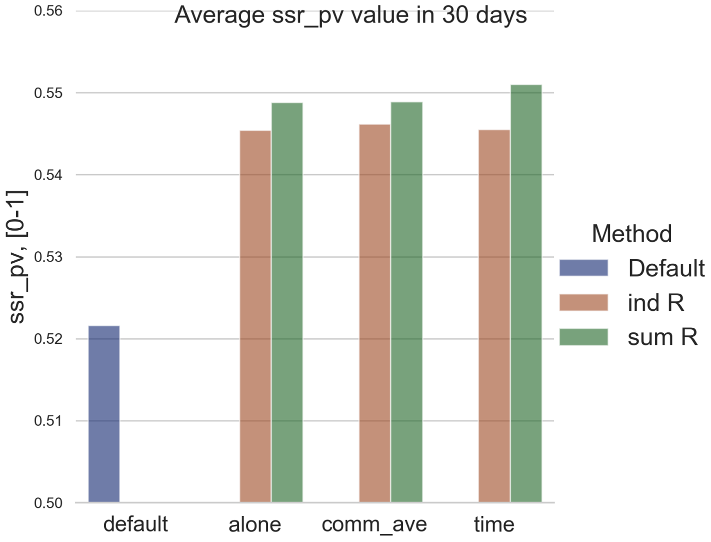
Purchased power



Exchanged power



SSR



- Using sum purchased power reward setting performs better than individual reward

Conclusions

- For multiple houses, RL methods outperform than fixed rule-based controller
- For multiple houses, Prior-DQN method outperform DQN
- Having community average information could further improve the performance
- Using sum purchased power for reward outperforms using individual purchased power.

Further Work

- Different MARL techniques, and longer training/testing data
- Hardware connections
- Expanding in different local communities

Bibliographical References

- T. Sakagami, A. Werth, M. Tokoro, Y. Asai, D. Kawamoto, and H. Kitano. Performance of a dc-based microgrid system in Okinawa. In 2015 International Conference on Renewable Energy Research and Applications (ICRERA), pages 311–316. IEEE, 2015.
- T. Sakagami, Y. Asai, and H. Kitano. Simulation to optimize a dc microgrid in Okinawa. In 2016 IEEE International Conference on Sustainable Energy Technologies (ICSET), pages 214–219. IEEE, 2016.
- A. Werth, N. Kitamura, and K. Tanaka. Conceptual study for open energy systems: distributed energy network using interconnected dc nanogrids. IEEE Transactions on Smart Grid, 6(4):1621–1630, 2015.
- Sony Computer Science Laboratories, Inc. Autonomous power interchange system. https://github.com/SonyCSL/APIS, 2020.
- T. Schaul, J. Quan, I. Antonoglou, and D. Silver. Prioritized experience replay. ICLR 2016.

Acknowledgements

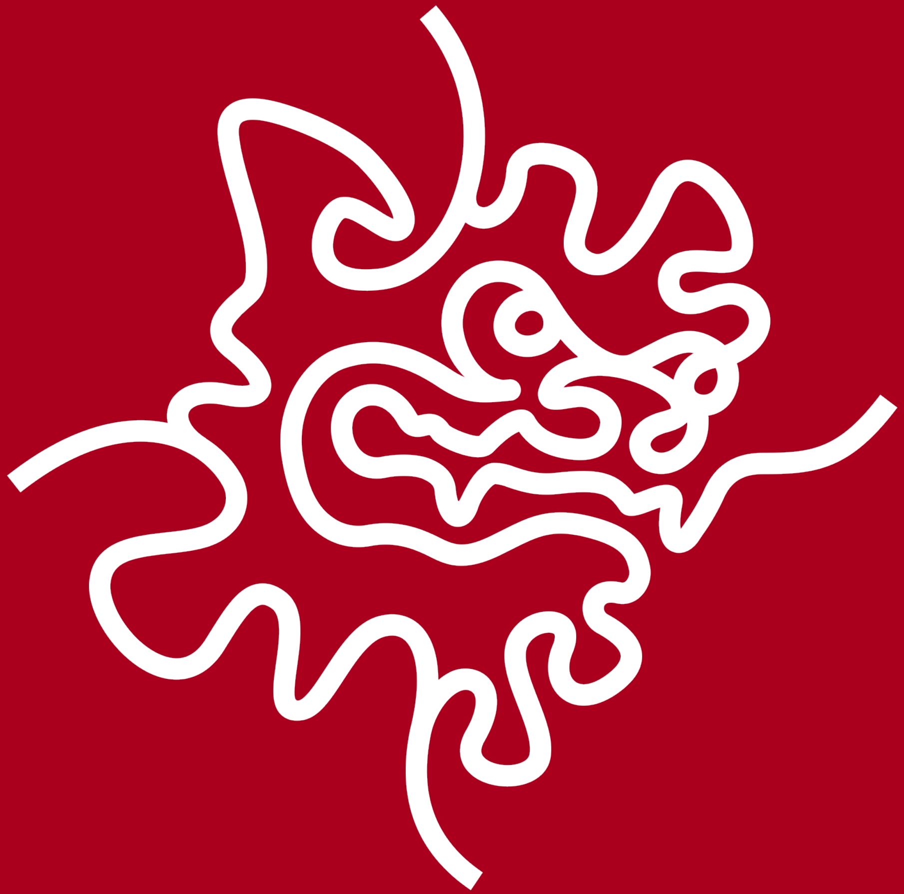
Prof. Kenji Doya
Supervisor @NC Unit

Prof. Hiroaki Kitano
Committee @IOS Unit

Mr. Kenichiro Arakaki
Technical @IOS Unit

Mr. Daisuke Kawamoto
Staff @Sony CSL

Neural Computational Unit



qiong.huang@oist.jp