# Ethical Smart Grid: a Gym environment for learning ethical behaviours

Clément Scheirlinck[1], Rémy Chaput [1][¶], and Salima Hassas[1]

1 Univ Lyon, UCBL, CNRS, INSA Lyon, Centrale Lyon, Univ Lyon 2, LIRIS, UMR5205, F-69622 Villeurbanne, France ¶ Corresponding author

Summary

The `ethical-smart-grid` package is a reinforcement learning (RL) simulator of a smart grid, based on Gym (Brockman et al., 2016). It focuses on learning *ethical behaviours*, i.e., behaviours that include ethical considerations, and are aligned with moral values. The simulated smart grid is composed of multiple agents representing inhabitants, which have to learn how to consume and distribute energy while taking into account several ethical considerations, such as the equity of comforts between inhabitants.

This simulator is somewhat simplified compared to existing (co-)simulators (Karnouskos & De Holanda, 2009; Mihal et al., 2022; Oliveira et al., 2012; Vogt et al., 2018); it focuses instead on the ethical aspects. In other words, agents that have been trained on this environment could not be deployed in a real-world smart grid, yet this package still serves as a suitable environment for learning ethical behaviours.

The purpose of this work is to present a simple simulator that allows to run (reproducible) simulations, and to compare on a common environment multiple (ethical) decision-making algorithms. The simulator is implemented in an extensible way to allow third-party researchers to adapt it to their own needs, to explore and improve ethical behaviours.

Statement of need

The field of *Machine Ethics* has recently received numerous contributions that try to implement so-called *ethical behaviours* (Tolmeijer et al., 2020), which target various domains and use-cases. For example, Anderson et al. (2019) propose a "value-driven robot" instantiated on the eldercare domain, considering several ethical considerations, such as respecting the patient's autonomy, or maximizing its well-being. Although the decision-making algorithms of such approaches are thoroughly defined, the environment in which they are tested is most often not available. Thus, Machine Ethics researchers cannot compare different contributions on a common environment, e.g., to observe their respective effect on the exhibited behaviours.

`ethical-smart-grid` is an open-source environment including ethical considerations, which we propose to the community as a first step. In addition, this package may facilitate new contributions to the Machine Ethics community, by providing a ready-to-use environment, such that researchers can focus on the decision-making algorithms themselves rather than building both the algorithm and the application environment.

A Smart Grid simulator

The simulator is composed of multiple *prosumer* (producer and consumer) agents that interact in a shared smart grid by consuming and exchanging energy. It follows the standard RL

interaction loop (Sutton & Barto, 2018): agents receive observations, and take actions that update the environment's state (see Figure 1).
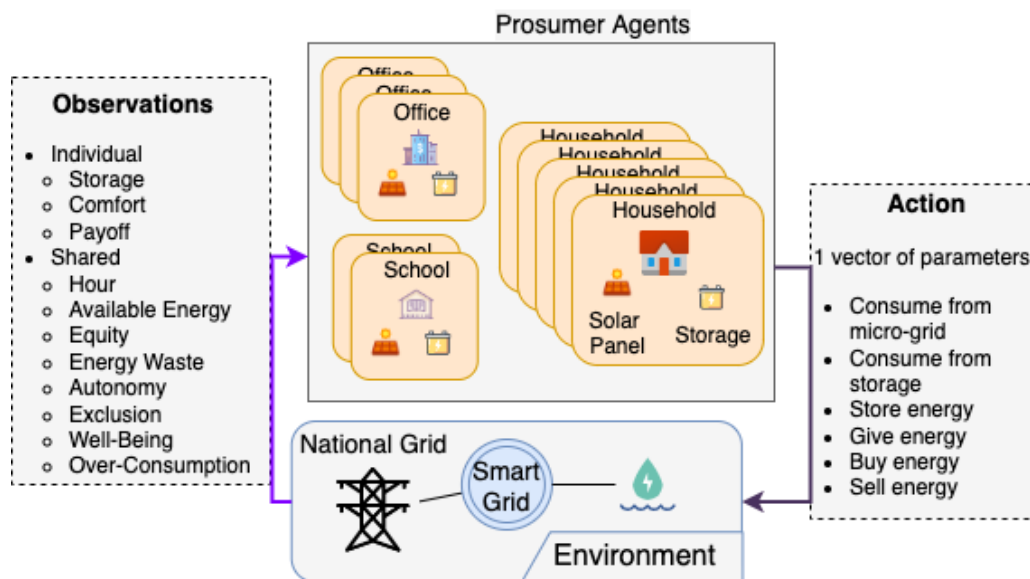

Figure 1: Simulator schema

A noteworthy aspect of this simulator is that both observations and actions are continuous and multi-dimensional, i.e., observations are represented as vectors in $\mathbb{R}^{11}$ and actions in $\mathbb{R}^6$. Some observations are shared by all agents whereas others are individual: other agents cannot access them. This is a design choice targeting the privacy of users who may be represented by such agents, if this simulator was deployed in the real world.

Four moral values taken from the literature (Boijmans, 2019; De Wildt et al., 2019; Milchram et al., 2018) are targeted in this environment: *security of supply*, *affordability*, *inclusiveness*, and *environmental sustainability*. Depending on the reward function used in simulations, these moral values can be focused individually, or several at a time, and may conflict at some time steps, making this simulator a suitable environment for learning ethics.

In order to propose a simple, yet extensible, simulator, we follow the Gym (now Gymnasium) library standard (Brockman et al., 2016). This standard is well-known in the Reinforcement Learning community, which makes our simulator easily compatible with many existing decision-making algorithms that have an implementation available for Gym environments.

However, the simulator slightly differs with Gym by accepting multiple agents instead of a single one. The main step function thus takes a list of actions, and returns a list of observations and a list of rewards. This modification should be compatible with existing multi-agent algorithms; we explain nonetheless how to adapt custom models to our simulator in the documentation.

Multiple scenarios can be used to parametrize the simulator, with several parameters such as the number of agents, their types, the quantity of energy available, and so on. Several components can be extended, and new scenarios can be implemented (see Open to extensions).

Using the simulator

This package can be used through the standard Gym interaction loop:

```python
from smartgrid import make_basic_smartgrid
```

```python
env = make_basic_smartgrid()
obs = env.reset(seed=None)  # Set the seed here for reproducible simulations
done = False
while not done:
    # Replace the `actions` array with your own learning algorithm here
    actions = [
        agent.profile.action_space.sample()
        for agent in env.agents
    ]
    obs, reward_n, terminated_n, truncated_n, info_n = env.step(actions)
    done = all(terminated_n) or all(truncated_n)
```

More complex scenarios can be created to fully customize the simulator, by configuring existing components instead of using the `make_basic_smartgrid()` function; the documentation provides a tutorial on this subject.

Open to extensions

Although this package is completely usable as-is, it was designed to be open to extensions, in particular by third-party researchers. This is particularly important within the field of Machine Ethics, in which ethical considerations are not always the same between different groups or cultural contexts. To allow other members of the community to bring their own moral values, ethical considerations, or even data sets for realistic simulations, we explain in the documentation how to extend several aspects, including:

- Agents' profiles, which determine a few common characteristics of agents, such as their needs for each time step, or the quantity of energy they produce. These characteristics can be based on external datasets for a realistic simulation, e.g., using real data corresponding to existing buildings.
- The environment's conditions, such as the quantity of energy available at each time step. This can be used to simulate, for example, a scenario of scarcity, or on the contrary an abundance. More complex scenarios can also be created, e.g., where agents first experience abundance and then have to adapt to scarcity.
- Reward functions that determine the rewards received by agents, i.e., the degree of "correctness" of their actions, with respect to the moral values represented by the function. Several moral values are already available, and other ones can be implemented through new reward functions. These functions can focus on one or several objectives, and potentially be aggregated to cope with single-objective decision-making algorithms.

Mentions

A previous (closed-source) version was used by one of the authors (Rémy Chaput) during his PhD studies to support experiments on several learning algorithms. However, the simulator was not easily re-usable by the community; this open-source version is a port, as close as possible to the old one, but with an emphasis on extensibility and re-usability.

Acknowledgments

References

Anderson, M., Anderson, S. L., & Berenz, V. (2019). A value-driven eldercare robot: Virtual and physical instantiations of a case-supported principle-based behavior paradigm. *Proceedings of the IEEE*, *107*(3), 526–540. https://doi.org/10.1109/JPROC.2018.2840045

Boijmans, A. (2019). *The acceptability of decentralized energy systems: Identifying value conflicts through simulations of decentralized energy systems for city districts* [Master's thesis]. TU Delft.

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). OpenAI gym. *arXiv Preprint arXiv:1606.01540*.

De Wildt, T., Chappin, E., Van de Kaa, G., Herder, P., & Van De Poel, I. (2019). Conflicting values in the smart electricity grid a comprehensive overview. *Renewable and Sustainable Energy Reviews*, *111*, 184–196. https://doi.org/10.1016/j.rser.2019.05.005

Karnouskos, S., & De Holanda, T. N. (2009). Simulation of a smart grid city with software agents. *2009 Third UKSim European Symposium on Computer Modeling and Simulation*, 424–429. https://doi.org/10.1109/EMS.2009.53

Mihal, P., Schvarcbacher, M., Rossi, B., & Pitner, T. (2022). Smart grids co-simulations: Survey & research directions. *Sustainable Computing: Informatics and Systems*, *35*, 100726. https://doi.org/10.1016/j.suscom.2022.100726

Milchram, C., Van de Kaa, G., Doorn, N., & Künneke, R. (2018). Moral values as factors for social acceptance of smart grid technologies. *Sustainability*, *10*(8), 2703. https://doi.org/10.3390/su10082703

Oliveira, P., Pinto, T., Morais, H., & Vale, Z. (2012). MASGriP—a multi-agent smart grid simulation platform. *2012 IEEE Power and Energy Society General Meeting*, 1–8.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tolmeijer, S., Kneer, M., Sarasua, C., Christen, M., & Bernstein, A. (2020). Implementations in machine ethics: A survey. *ACM Computing Surveys (CSUR)*, *53*(6), 1–38.

Vogt, M., Marten, F., & Braun, M. (2018). A survey and statistical analysis of smart grid co-simulations. *Applied Energy*, *222*, 67–78. https://doi.org/10.1016/j.apenergy.2018.03.123