Deep Reinforcement Learning in World-Earth System Models to Discover Sustainable Management Strategies

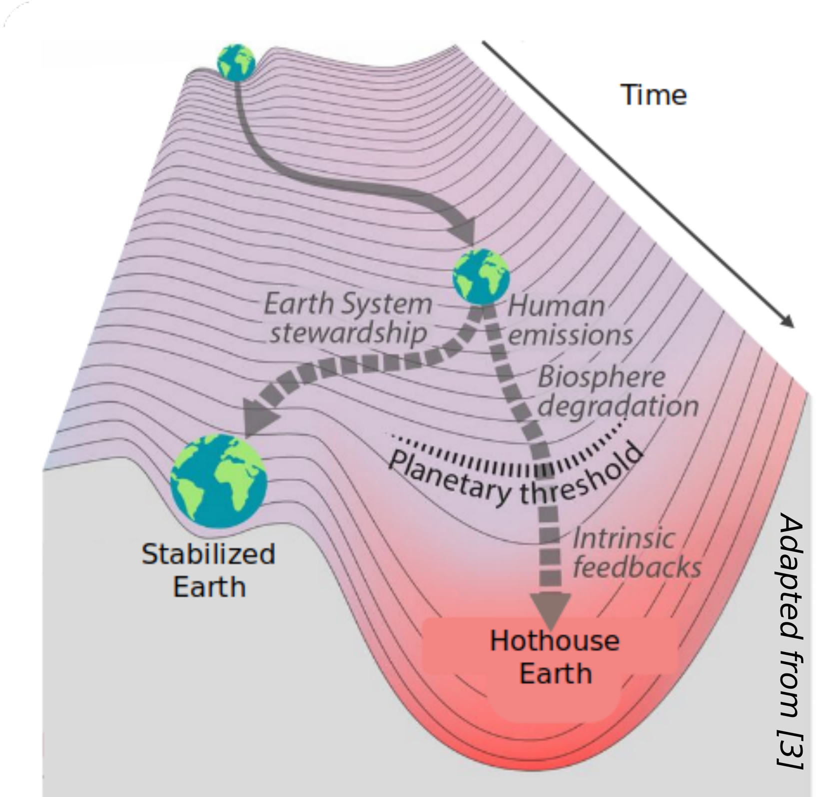
Felix M. Strnad[1,2,*], Wolfram Barfuss[1], Jonathan F. Donges[1,3], Jobst Heitzig[1]

1) Potsdam Institute for Climate Impact Research, 2) Georg-August Universität Göttingen, 3) Stockholm Resilience Center
* strnad@pik-potsdam.de

copan coevolutionary pathways

Motivation

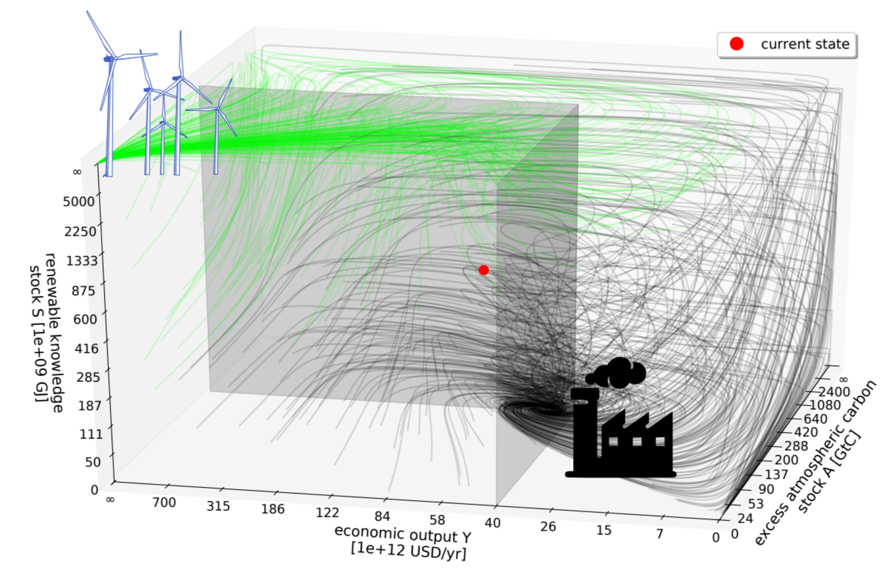
A Safe and Just Operating Space $\mathcal{O}$



How can the World be steered into a safe and just operating space for humanity [2]?

Management Pathways

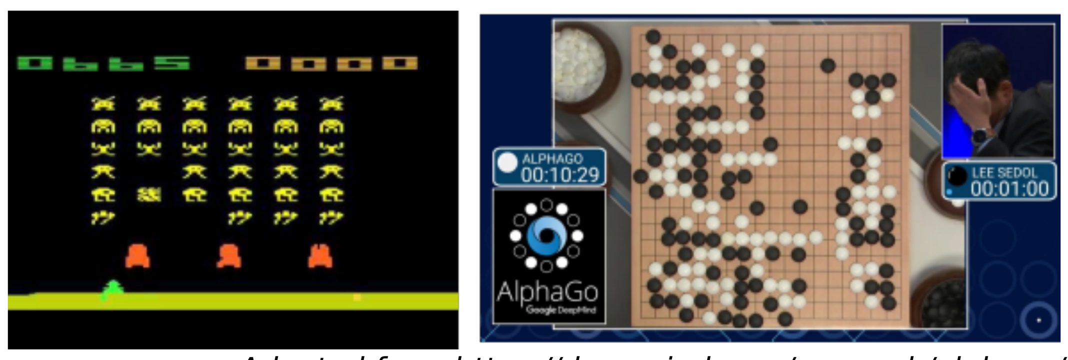
Computer models can be used to show possible pathways towards a sustainable future.



How to find an intelligent combination of management options in order to reach a sustainable future that stays within planetary boundaries at all times?

Deep Reinforcement Learning in Complex Systems

Deep Reinforcement Learning (DRL) algorithm has been proven to detect solutions up to super-human performance in various manageable environments



Adapted from https://deepmind.com/research/alphago/

Idea:
Use DRL to uncover previously unknown appropriate management strategies for the World-Earth system

Agent-Environment Interface

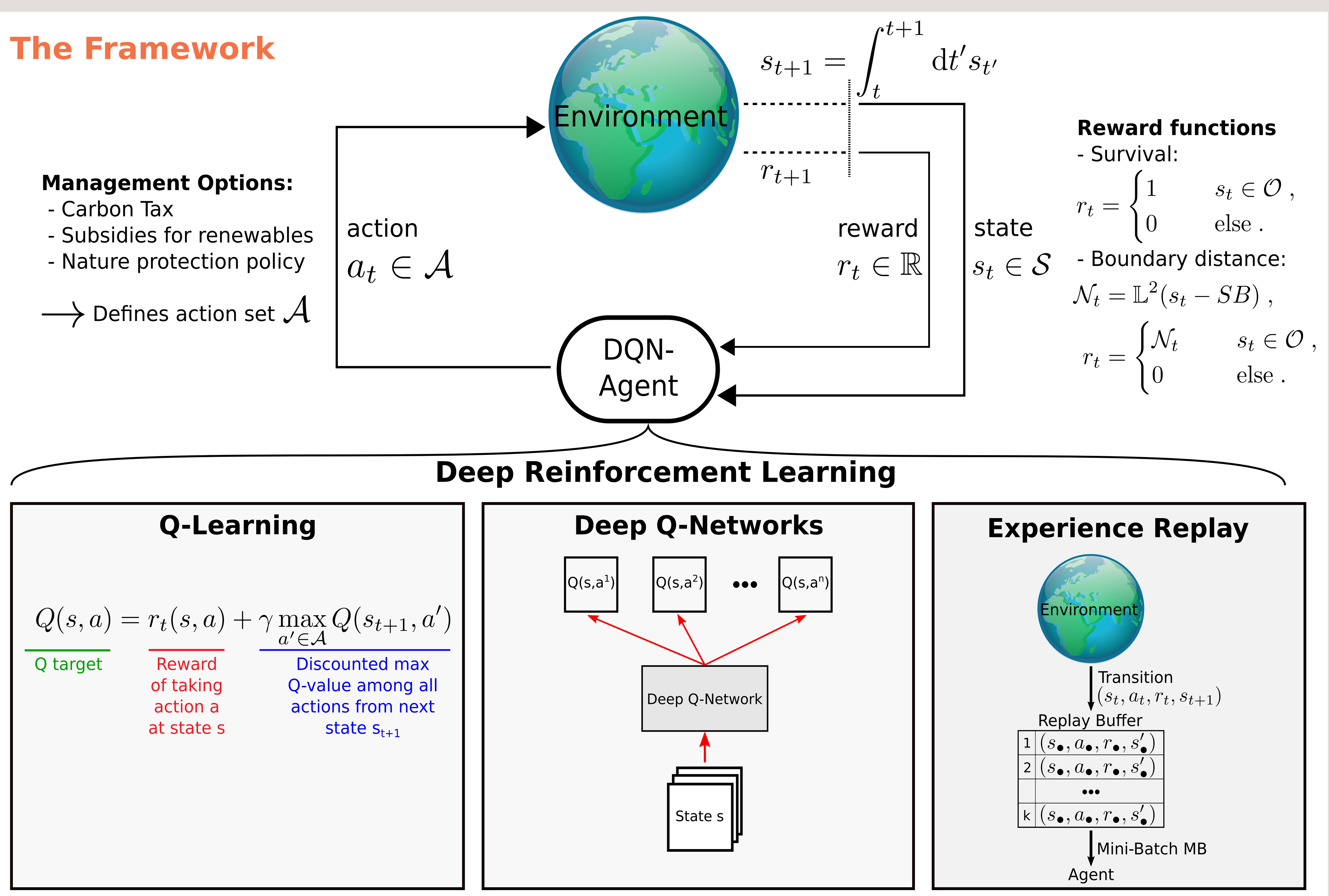
Interface Design

Agent: Deep Reinforcement Learner
- Acts only based on the information of the current state and the rewards signal
- Uses DQN - Algorithm for Learning [4], i.e. the combination of Q-learning, neural networks and experience replay

Environment: World-Earth Models
- Combine socio-economic World dynamics with biophysical Earth system dynamics [5]
- Models are stylized, their main focus is set on a qualitative understanding of the complex dynamics of these systems rather than to be used for quantitative predictions.

Interface: Combine Agent and Environments
- Map the World-Earth systems to a Markov Decission Process in terms of concrete states in the environment, actions in the action set and reward functions [3]
- Implementation of action set and reward function just depends on the developer and is independent of the agent

The Framework



Key Findings

1.) Even though dynamics of the environments are unknown to the agent in advance, it is able to find previously undiscovered trajectories within planetary boundaries.

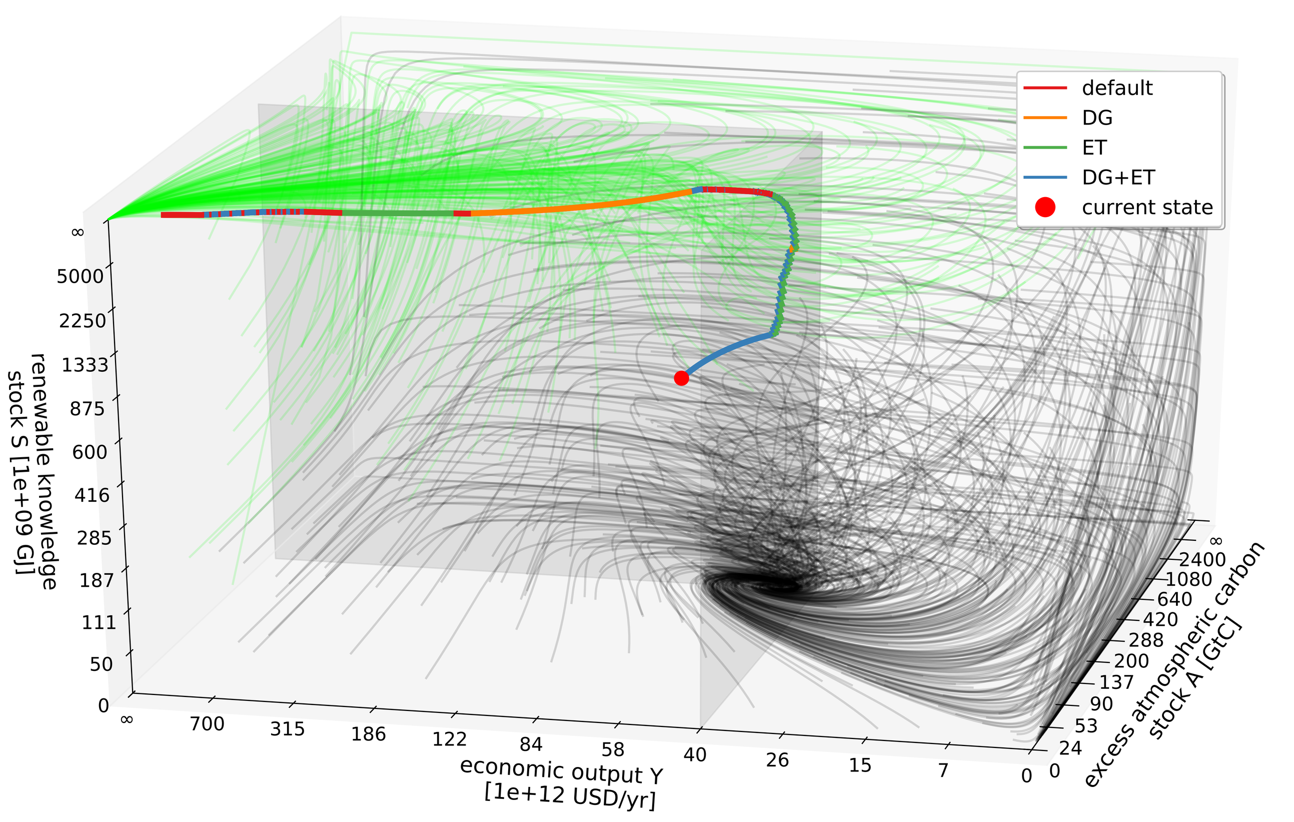


Figure: Example pathway in a stylized World-Earth system model. Green lines: attraction basin of sustainable fix point. Black lines: attraction basin of carbon based economy without renewables. In color: Example trajectory for path towards sustainable future. The different management options are: DG: Degrowth=Restrict resources. ET: Energy-Transformation=Carbon tax + Subsidies on renewables.

2.) Learning is stable and successful management is possible in various environments.

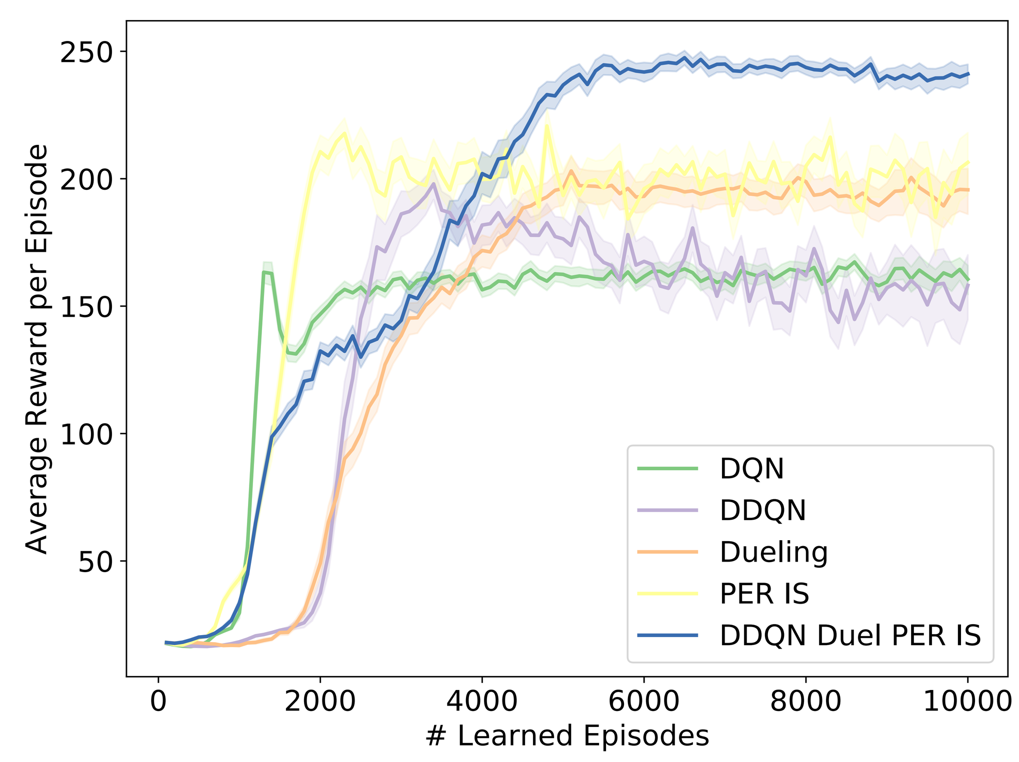


Figure: Development of total average reward per episode. Different Deep-Q-Network architectures are analyzed: DQN=Deep Q Networks, DDQN=Double DQN, DDQN Duel=Dueling Network Architecture with DDQN, DDQN Duel PER IS = DDQN Duel using prioritized experience replay with importance sampling.

Find us on GitHub: https://github.com/fstrnad/pyDRLinWESM

3.) Even under partial observability of state s=(LAGTPKS) and noisy measurements, the agent is still capable of detecting solutions.

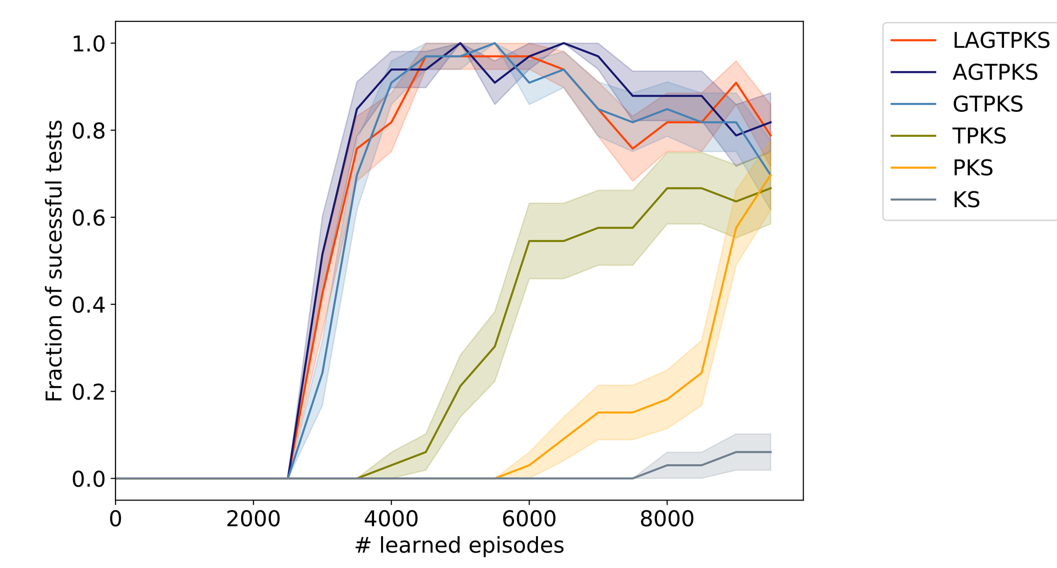


Figure: Development of percentage of successful tests for agents with different knowledge about the state variables. The different variables denote:L=terrestrial carbon, A=atmospheric carbon, G=geological carbon, T=temperature, P=population, K=capital, S=renewable energy knowledge.

References

[1] Strnad et al. (2019). Deep reinforcement learning in World-Earth system models to discover sustainable management strategies, *Chaos*
[2] Rockström et al. (2009). A safe operating space for humanity. *Nature*
[3] Steffen et al. (2018). Trajectories of the Earth System in the Anthropocene, PNAS
[4] Mnih et al. (2015). Human-level control through deep reinforcement learning. *Nature*
[5] Donges et al. (2017). Closing the loop: Reconnecting human dynamics to Earth System science, *The Anthropocene Review*