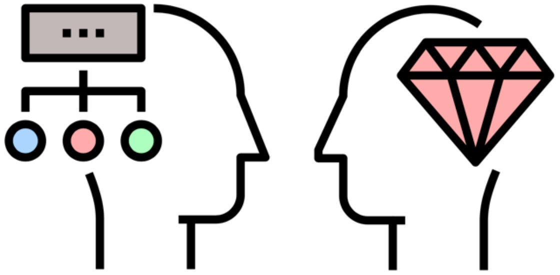
Imperial College London

Machine Learning for Materials



Zhenzhu Li

Department of Materials

Outline

- Optimization strategies
- RL in focus
- Alloy design
- Multi-objective

Optimisation strategies

Nature-inspired algorithms

Genetic Algorithm

Ant Colony Optimization

Artificial Bee Colony Optimization

Simulated Annealing

Gravitational Search Algorithm

Firefly Algorithm

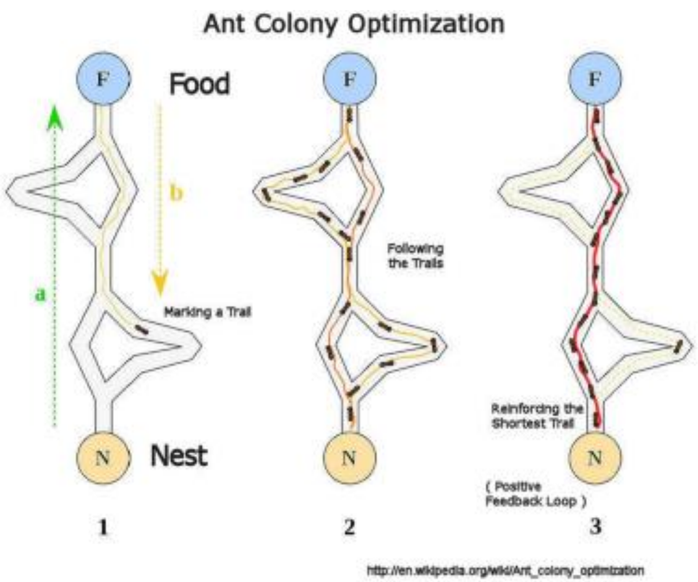
Fish Swarm Algorithm

Japanese Tree Frogs Algorithm

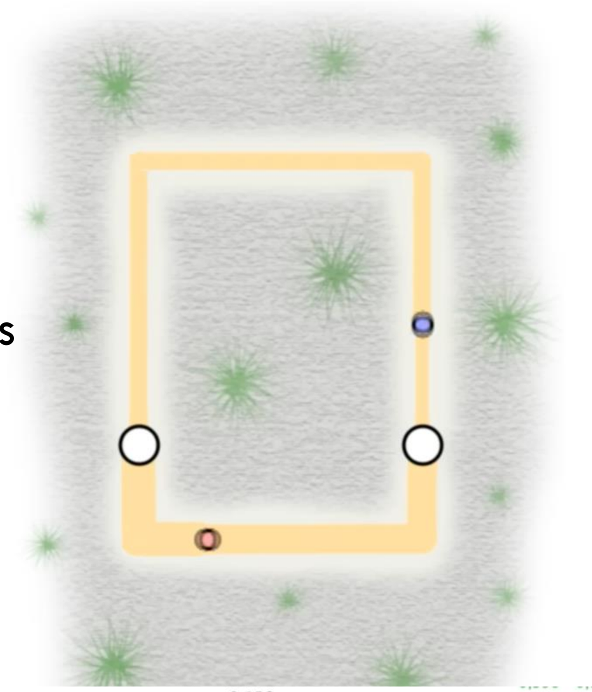
Optimisation strategies

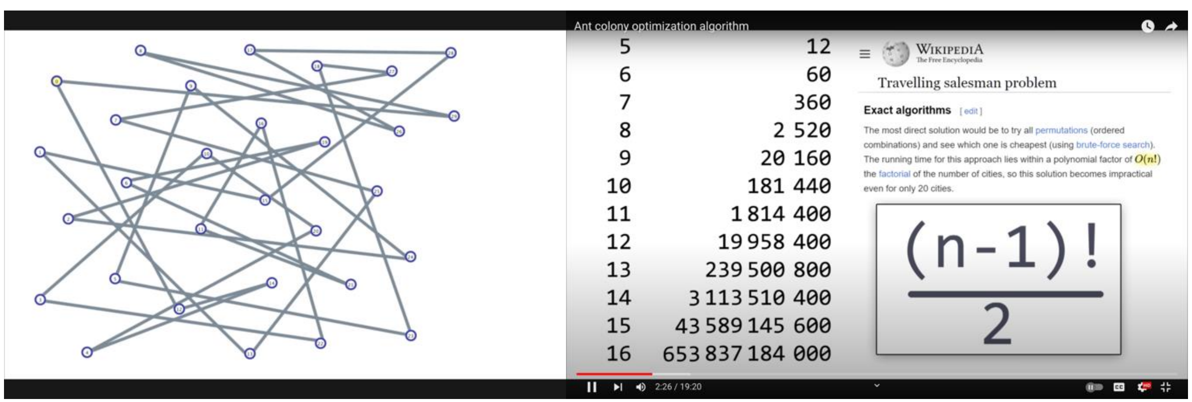
https://www.youtube.com/watch?app=desktop&v=u7bQomllcJw

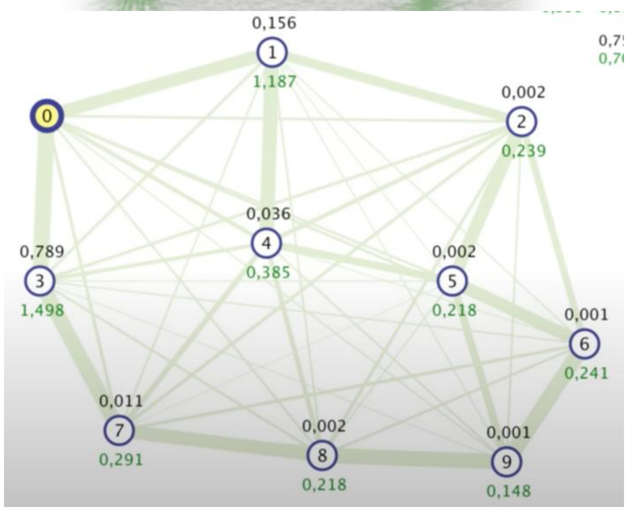
Optimisation strategies: ACO



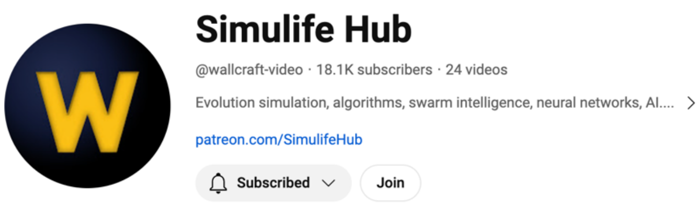
Update the hormones left with experiments



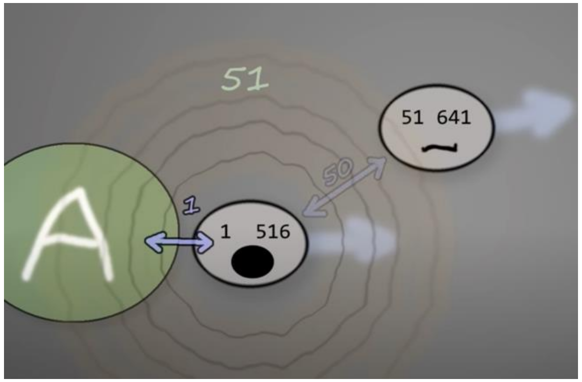


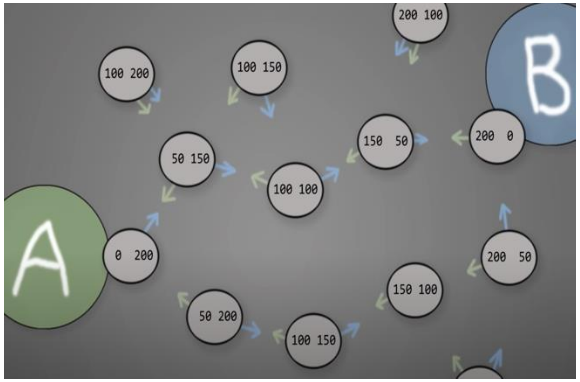


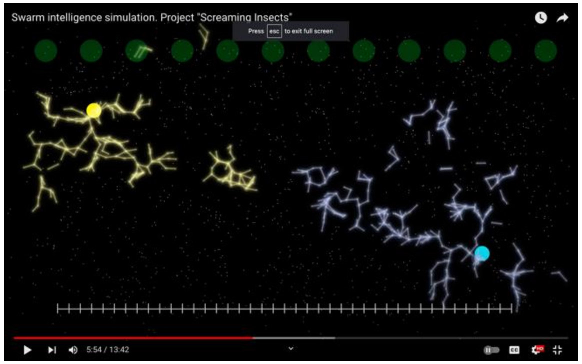
Optimisation strategies

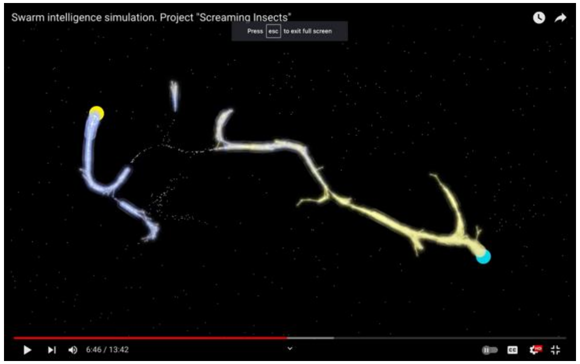


Swarm intelligence









Optimisation Problems

Optimization/Search Problems

Classical/Benchmark Problems

Examples:

- Mathematical test functions (e.g., Ackley, Easom, EggCrate, Eggholder, Rastrigin, Schaffer, McCormick, Beale, Branin, Colville and Rosenbrock)
- n-Puzzle Problem
- n-queens Problem
- Grid Search Problem
- Bin Packing Problem
- Knapsack Problem
- Minimum Spanning Tree (MST)
- Travelling Salesman Problem (TSP)
- Chinese Postman Problem (CPP)
- Optimal Assignment Problem (OAP)
- Quadratic Assignment Problem (QAP)
- Job Shop Scheduling Problem (JSP)
- Graph Coloring Problem

Real-world Problems

Design Problems

Examples:

- Suspension/wheel Design
- VLSI Design
- Assembly Line Balancing (ALBP)
- Railway Scheduling
- PID Controller Design
- Voice Activity Detector (VAD)
- Timetabling Problem (TTP)
- Political Districting Problem
- Hospital Resource Planning
- Optimal placement of physical assets (e.g., cameras, EV charging stations, micromobility stations and walking/cycling routes/lanes)

Planning Problems

Examples:

- Motion Planning
- Ride-sharing/Ride-hailing
- Shifts Planning
- Task Allocation
- Vehicle Routing Problem (VRP)
- Appointment Scheduling
- Patient Admission Scheduling
- Fitness Planning
- Trip Itinerary Planning
- Eco-efficient Delivery
- Deadheading Problem

Control Problems

Examples:

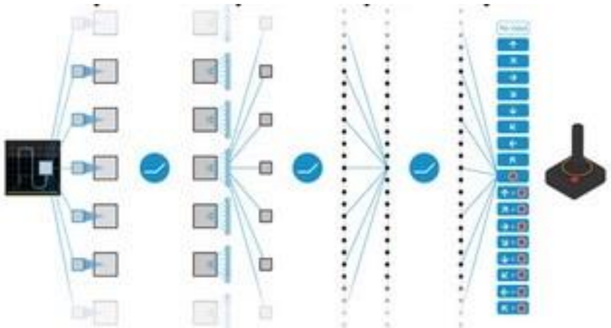
- Elevator Dispatching
- Communication Relaying
- Lateral and longitudinal Motion control
- Multirobot Control
- Targeted drug delivery using microrobots
- Multiple Target Clustering
- Dynamic Order Orchestration
- Self-driving vehicle (SDV) coordination in warehouses
- Truck Platooning

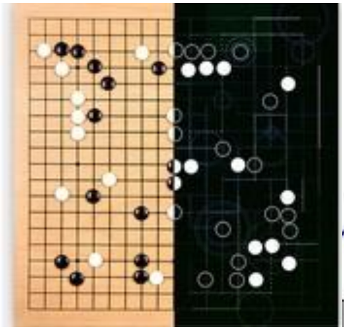
Optimisation strategies

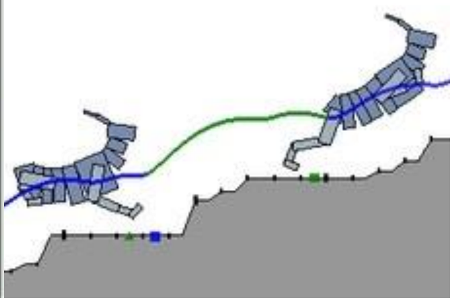
Ethology (the study of animal behavior) is the main source of inspiration of swarm intelligence algorithms such as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Artificial bee colony (ABC), Bat algorithm (BA), Social Spider Optimization (SSO), Firefly algorithm (FA), Butterfly Optimization Algorithm (BOA), Dragonfly Algorithm (DA), Krill Herd (KH), Shuffled Frog Leaping Algorithm (SFLA), Fish School Search (FSS), Dolphin Partner Optimization (DPO), Dolphin Swarm Opti-

Reinforcement learning in the wild

Boston Robotics

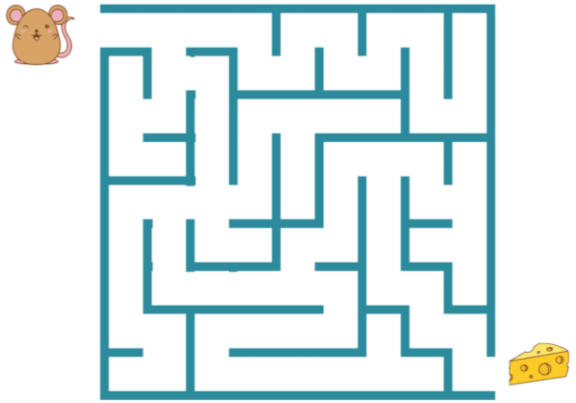


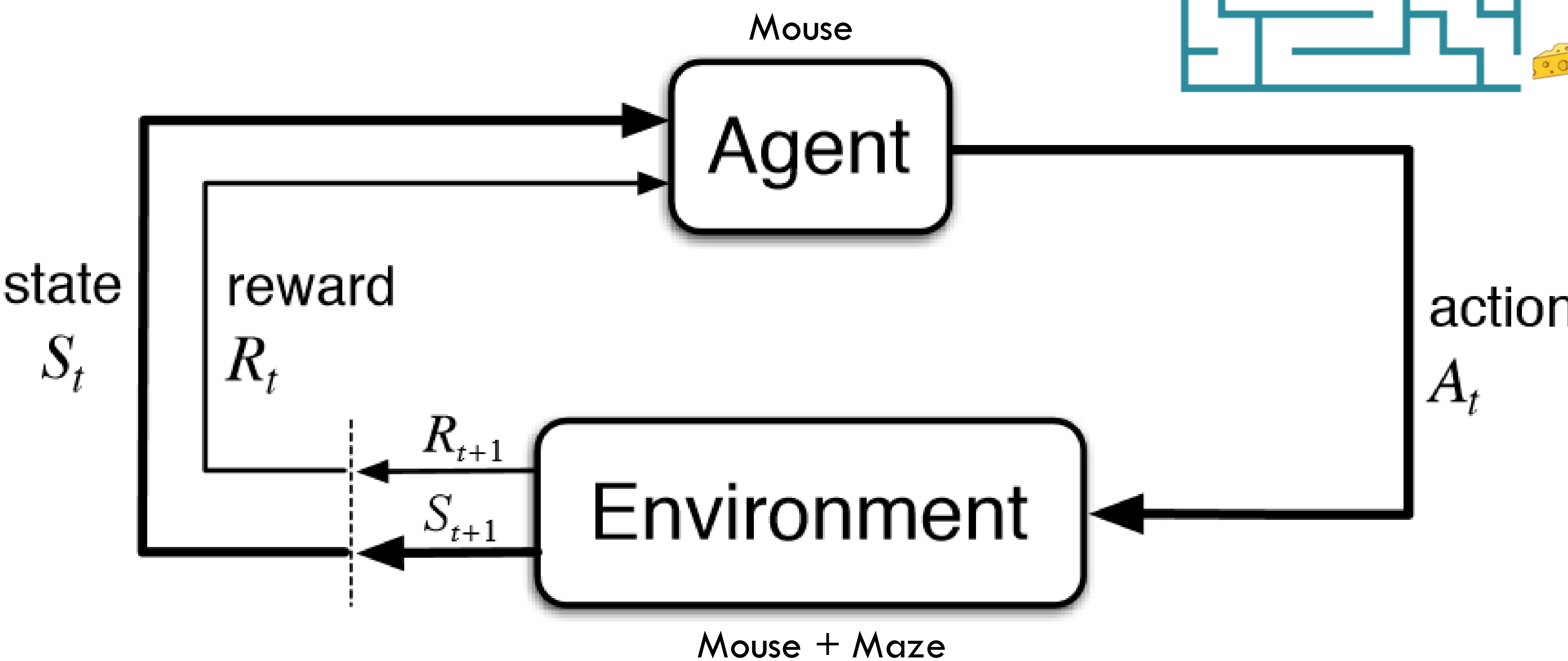




Deep Q Learning network playing ATARI, AlphaGo,
physically-simulated quadruped leaping over terrain.

Reinforcement learning

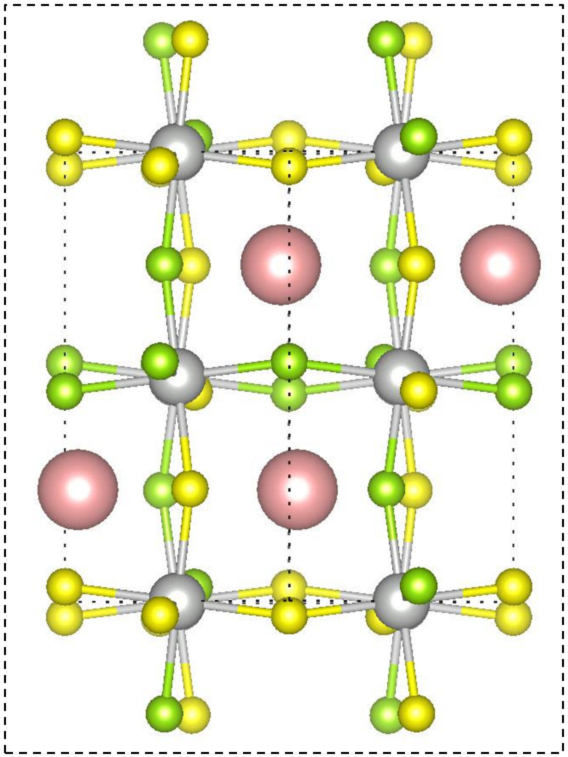


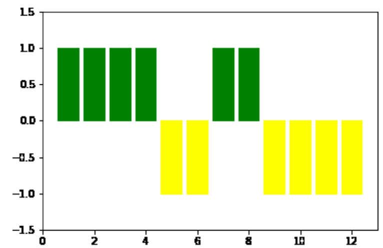


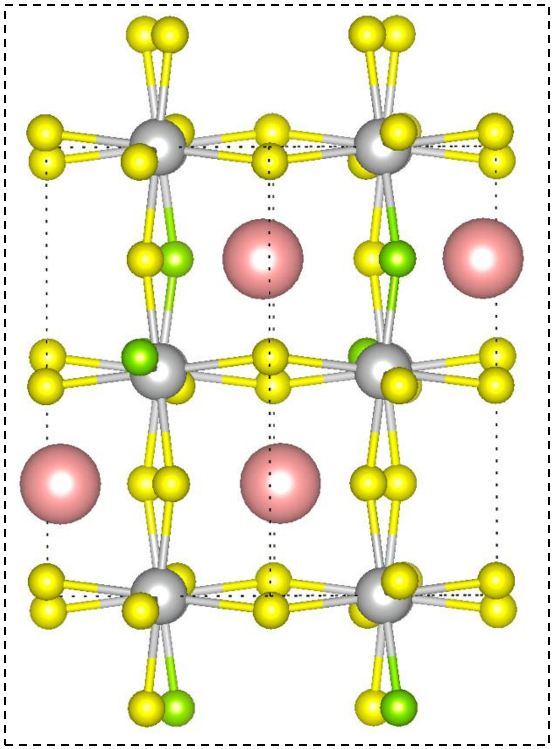
Materials search space

$(BaZrS_xSe_{3-x})_4$

Clean energy materials

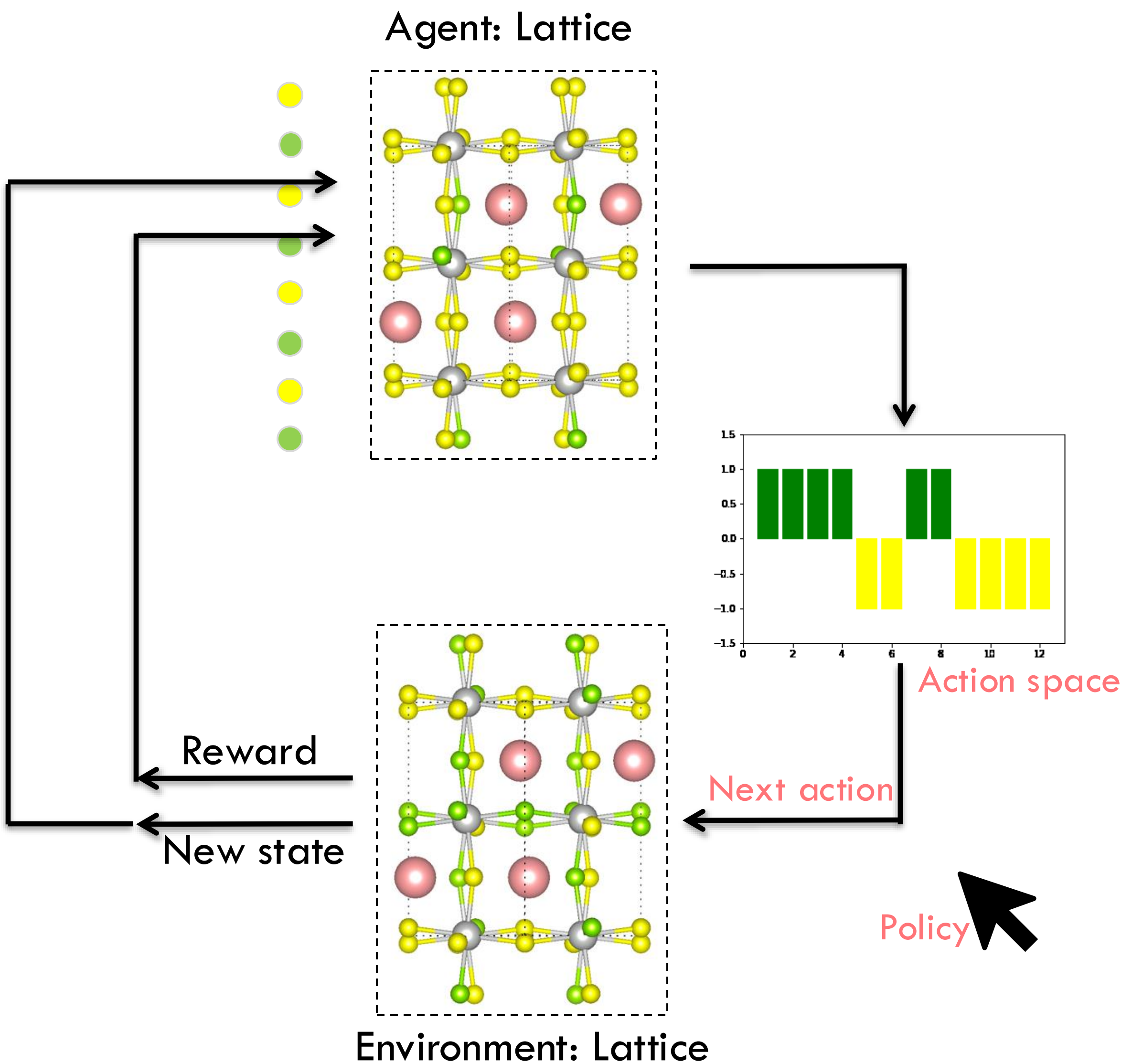






12 anion sites, how many possibilities?

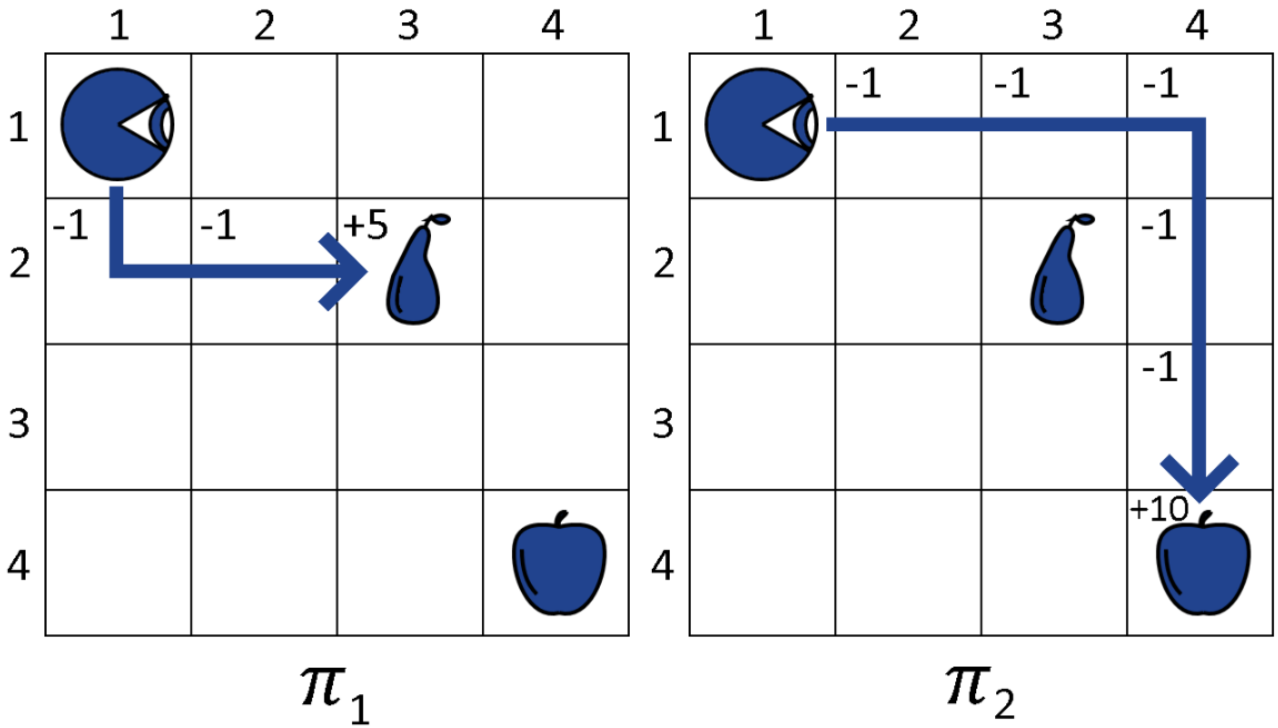
Reinforcement learning



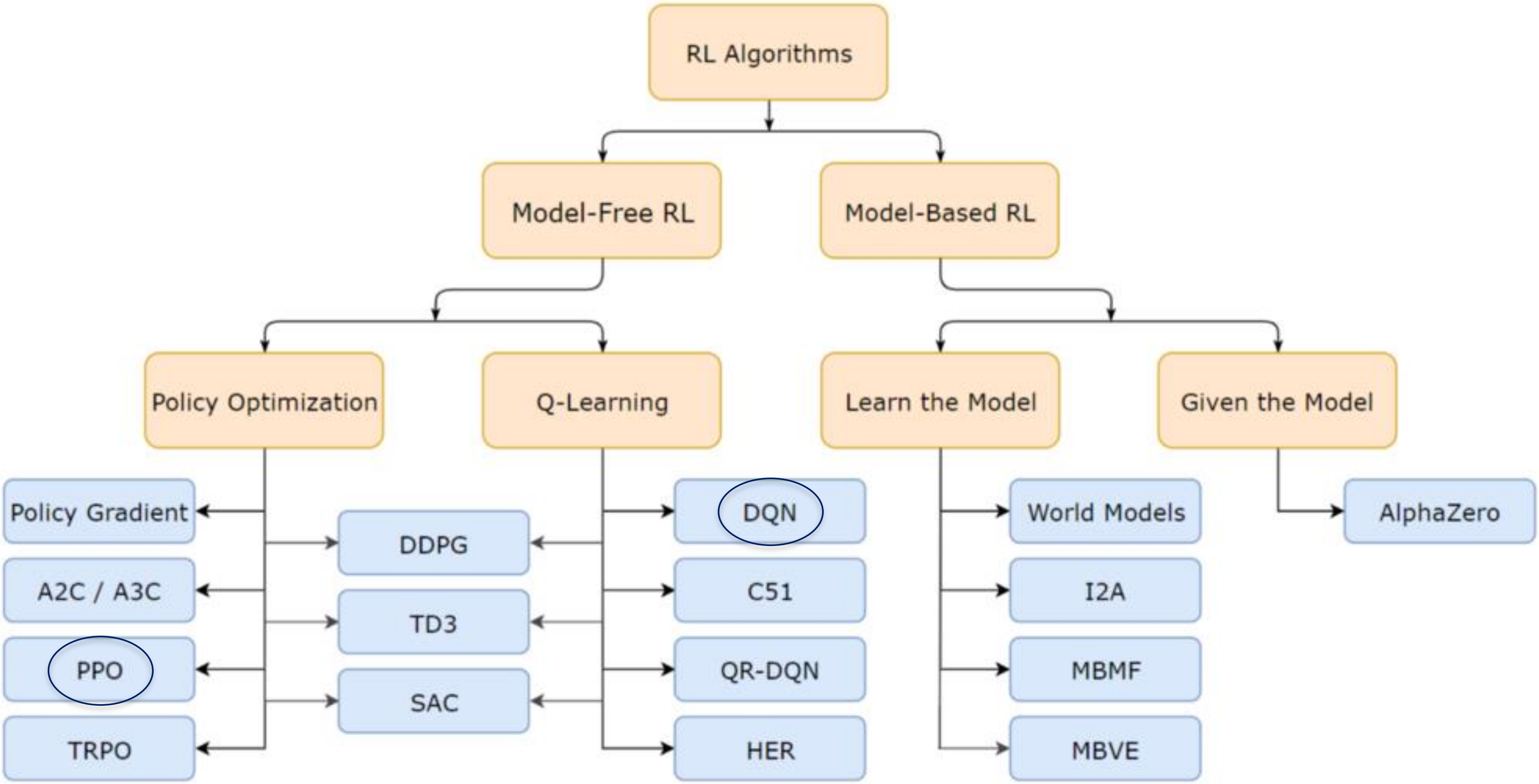
Policy

A policy π(s) comprises the suggested actions that the agent should take for every possible state s∈S.

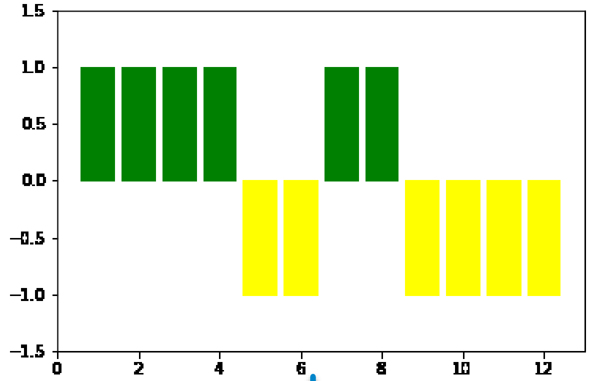
- $U(\pi_1) = -1 - 1 + 5 = +3$
- $U(\pi_2) = -1 - 1 - 1 - 1 - 1 + 10 = +5$



Policy


RL Algorithms
Model-Free RL
Model-Based RL
Policy Optimization
Q-Learning
Learn the Model
Given the Model
Policy Gradient
A2C / A3C
PPO
TRPO
DDPG
TD3
SAC
DQN
C51
QR-DQN
HER
World Models
I2A
MBMF
MBVE
AlphaZero

Action space



$2^{12} = 4096$

Discrete Action Space

Q-Learning

Deep Q-Networks (DQN)

SARSA (State-Action-Reward-State-Action)

Monte Carlo Methods

Continuous Action Space

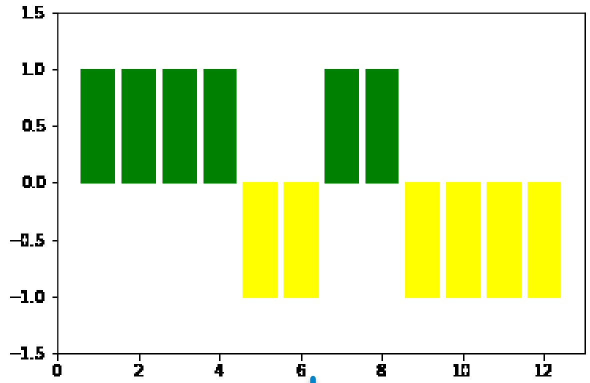
Deep Deterministic Policy Gradient (DDPG)

Proximal Policy Optimization (PPO)

Trust Region Policy Optimization (TRPO)

Soft Actor-Critic (SAC)

Action space



High-Dimensional Action Space

Hierarchical Reinforcement Learning (HRL)

Hindsight Experience Replay (HER)

Action Space Reduction

Actor-Critic Methods

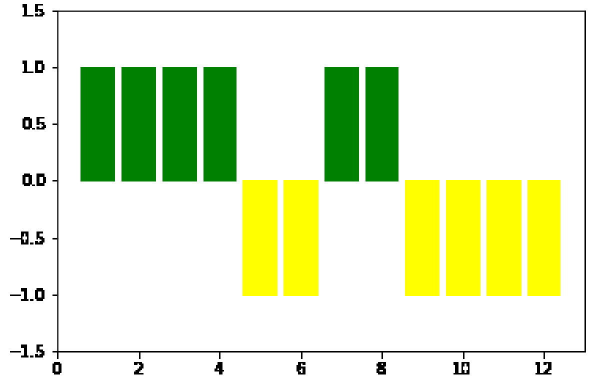
Sparse Reward Engineering

Policy Gradient Methods

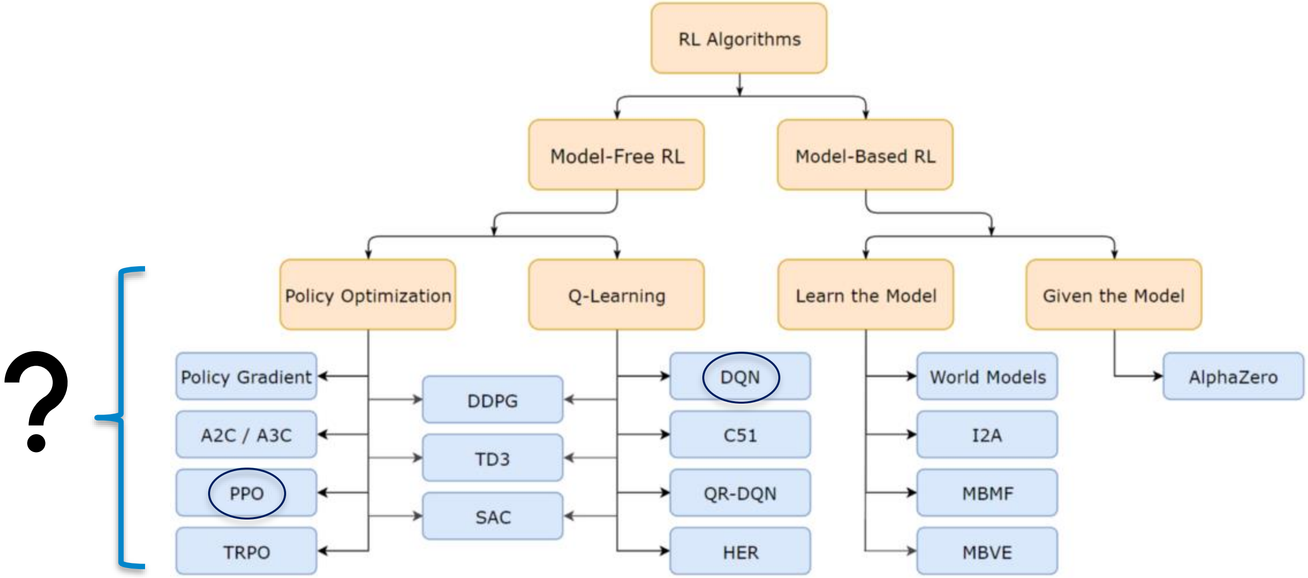
Function Approximation Techniques

Structured Action Space

Action space



High-Dimensional Action Space

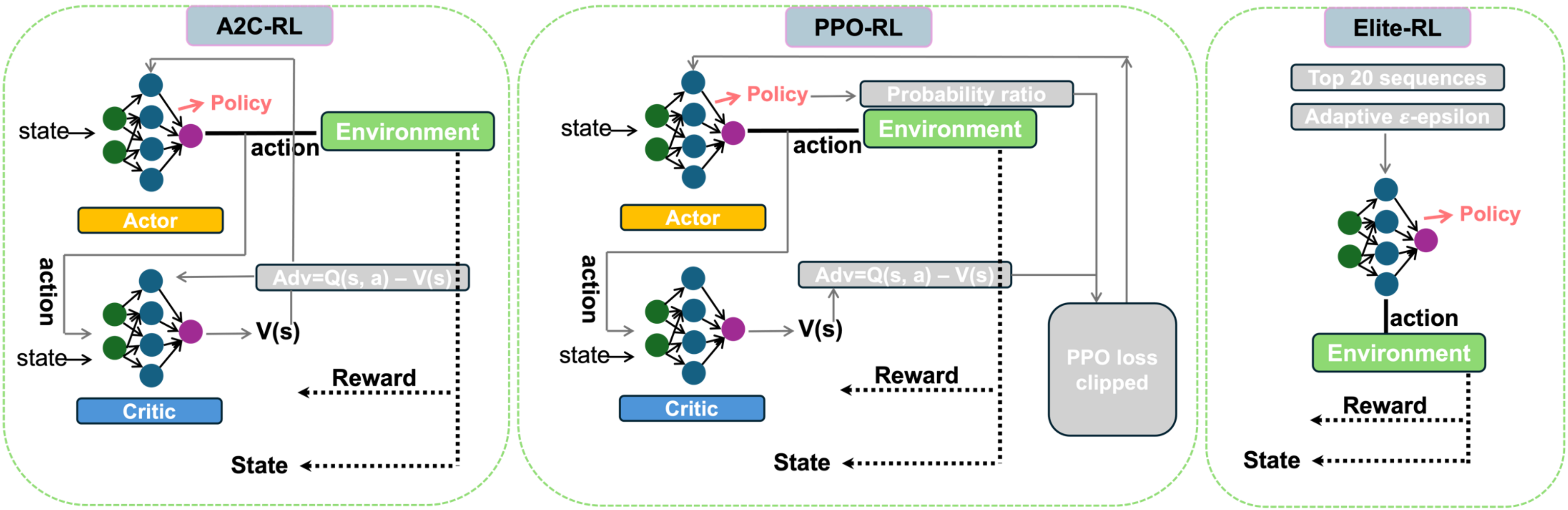


How to learn from memory

$$\text{Deterministic: } \pi(s) = a$$

$$\text{Stochastic: } \pi(a|s) = \mathbb{P}(A = a|S = s)$$

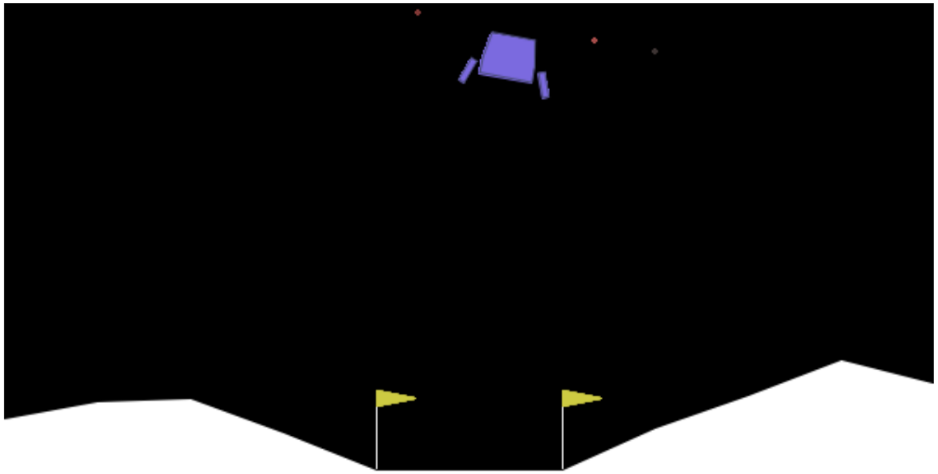
Policy:
Epsilon greedy action selection + next action is always the best action in previous 20 memories.



A2C-RL
state
Policy
Environment
action
Actor
action
Adv=Q(s, a) – V(s)
state
V(s)
Critic
Reward
State
PPO-RL
state
Policy
Probability ratio
Environment
action
Actor
action
Adv=Q(s, a) – V(s)
state
V(s)
PPO loss clipped
Critic
Reward
State
Elite-RL
Top 20 sequences
Adaptive ε-epsilon
Policy
action
Environment
Reward
State

Gymnasium

An API standard for reinforcement learning with a diverse collection of reference environments

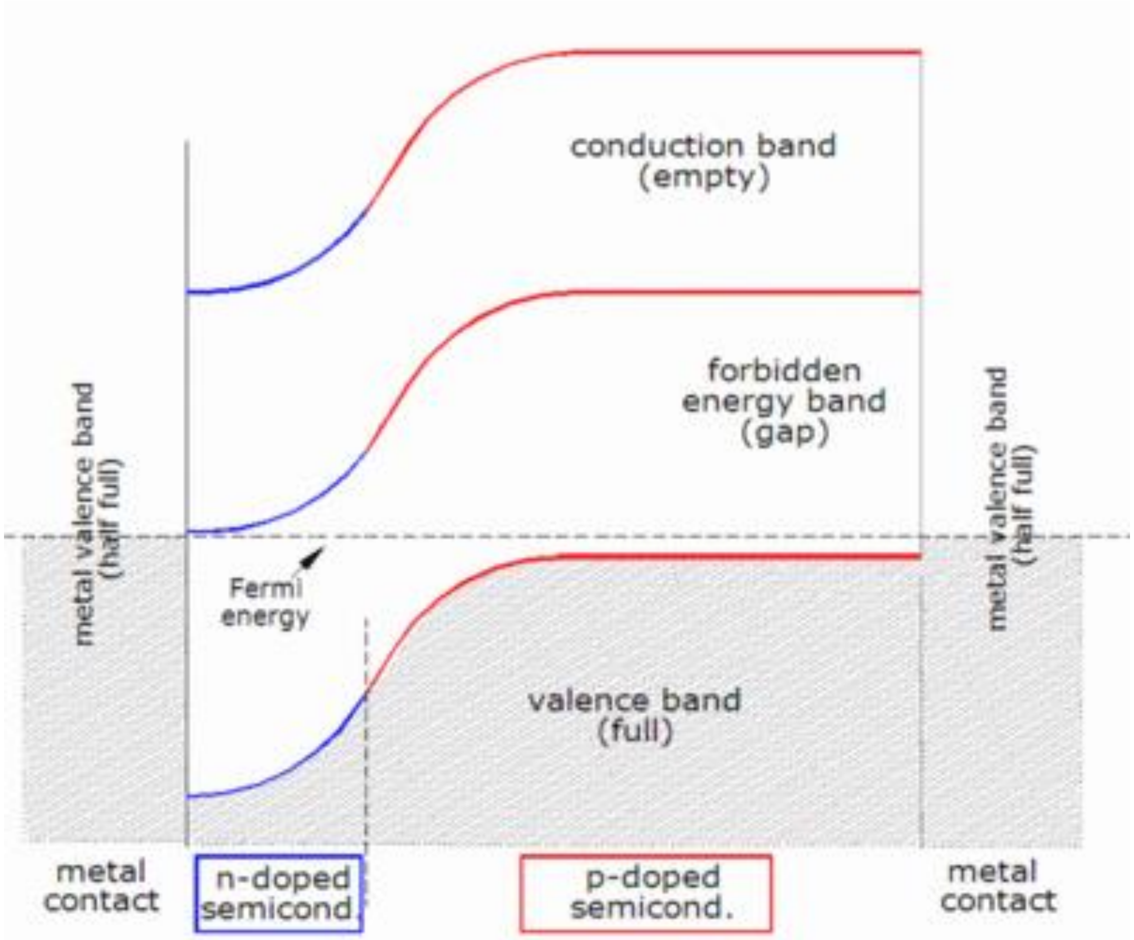


https://gymnasium.farama.org/

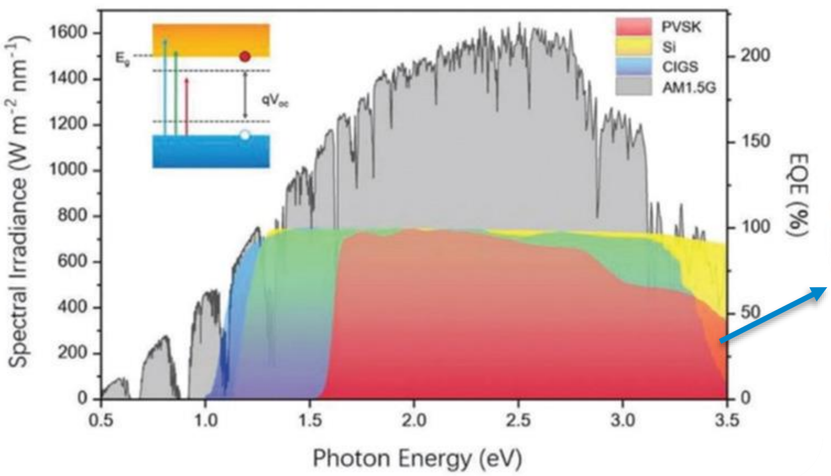
```python
import gymnasium as gym
env = gym.make("LunarLander-v2", render_mode="human")
observation, info = env.reset(seed=42)
for _ in range(1000):
   action = env.action_space.sample()  # this is where you would insert your policy
   observation, reward, terminated, truncated, info = env.step(action)

   if terminated or truncated:
      observation, info = env.reset()

env.close()
```

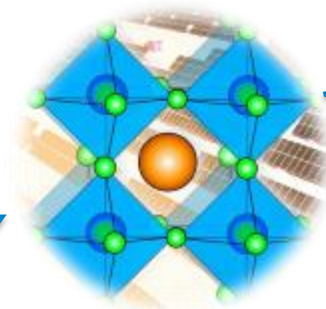
🎉 Built everything, now apply

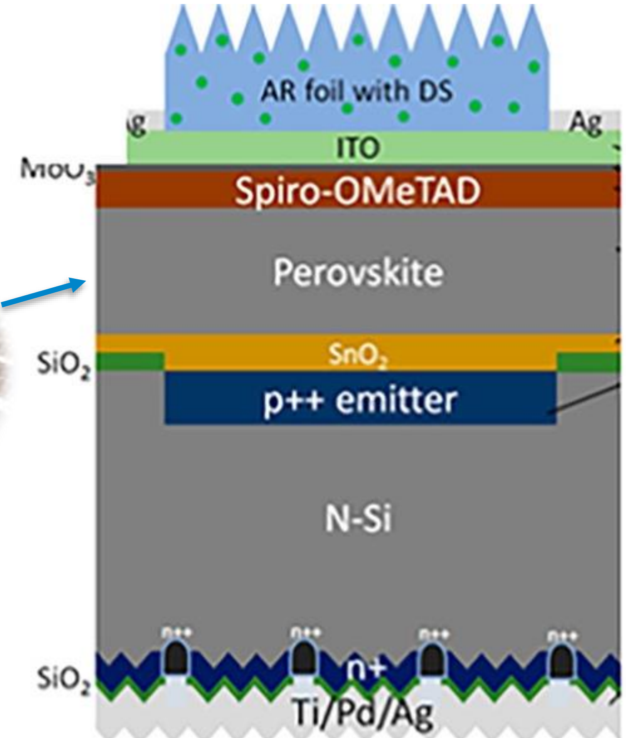
Photovoltaic effect



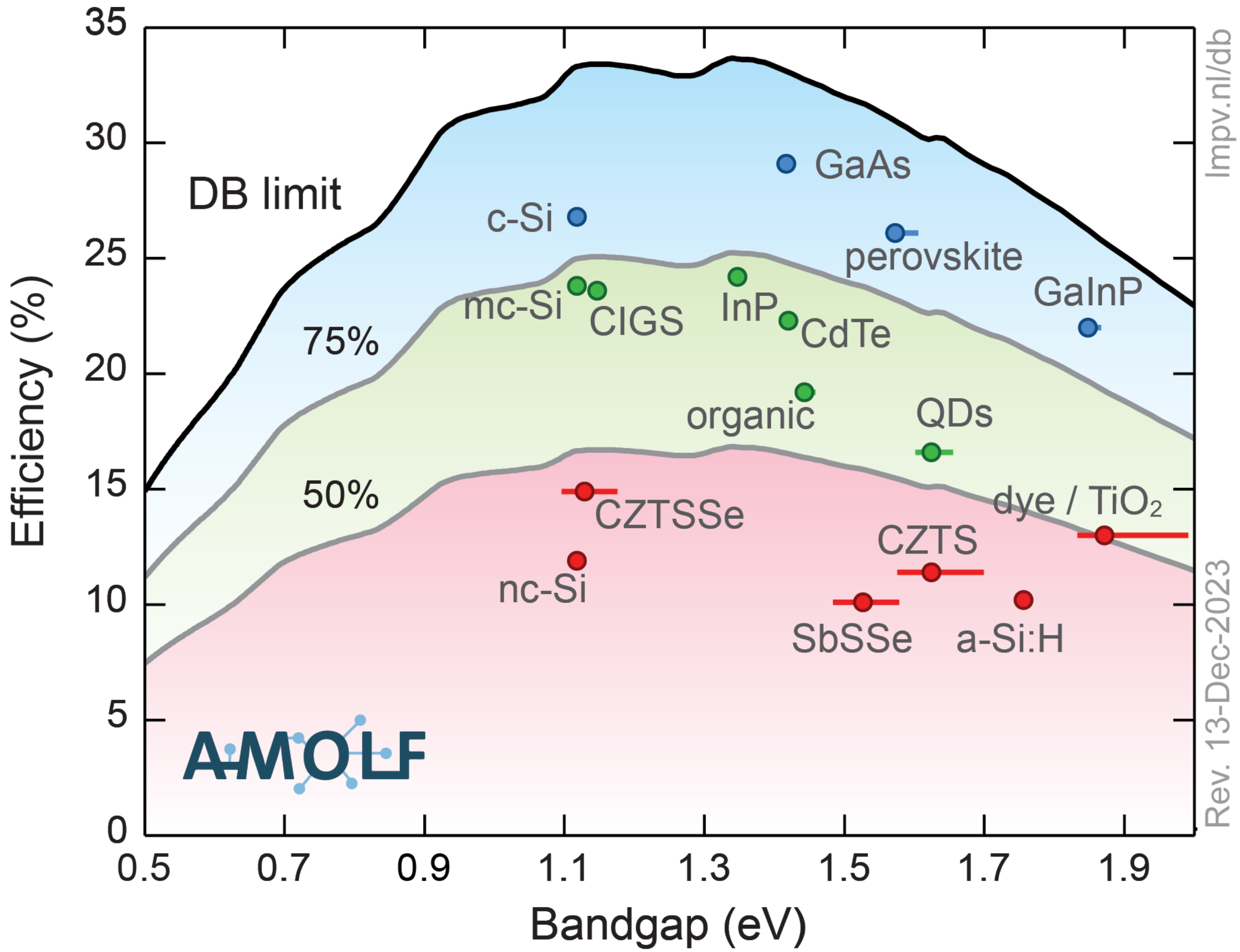
Photovoltaic device



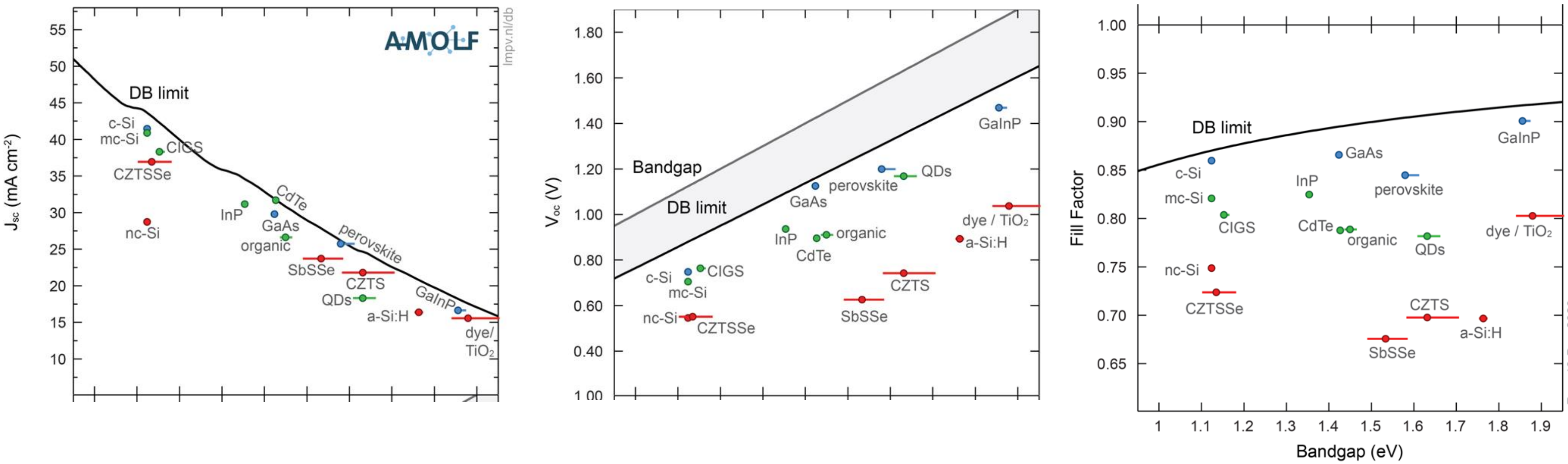




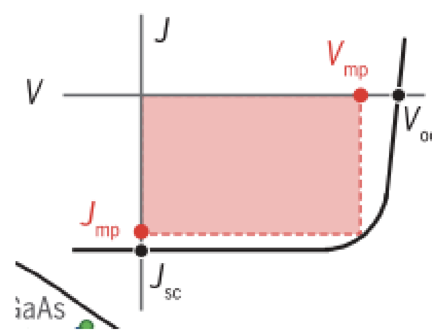
Detailed balance efficiency limit (Shockley Queisser Limit)



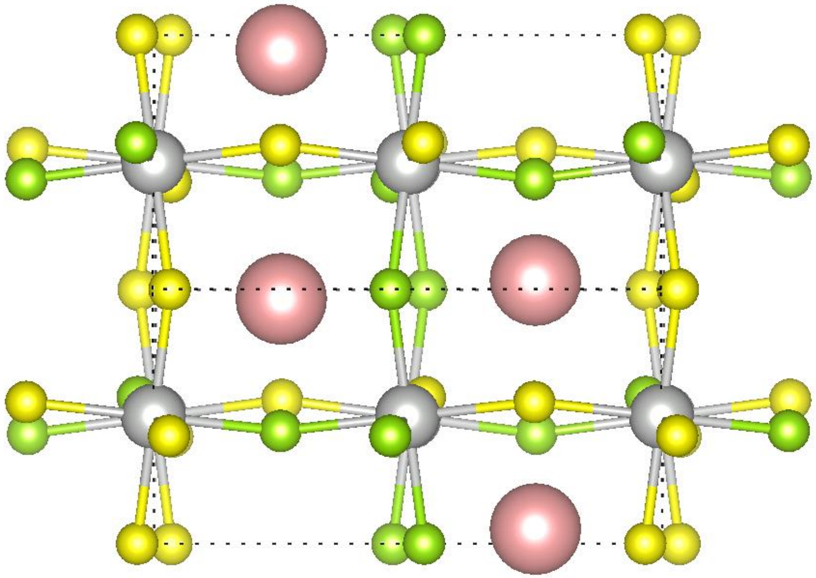
Detailed balance efficiency limit (Shockley Queisser Limit)



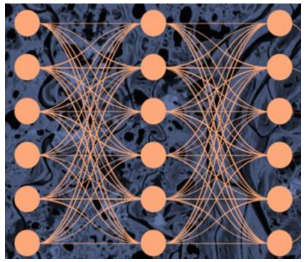
$$\eta = \frac{P_{\mathrm{max}}}{I_{\mathrm{in}}} = \frac{J_{\mathrm{mpp}} V_{\mathrm{mpp}}}{I_{\mathrm{in}}} = \frac{J_{\mathrm{sc}} V_{\mathrm{oc}} \mathrm{FF}}{I_{\mathrm{in}}}$$



Stable photo-absorber



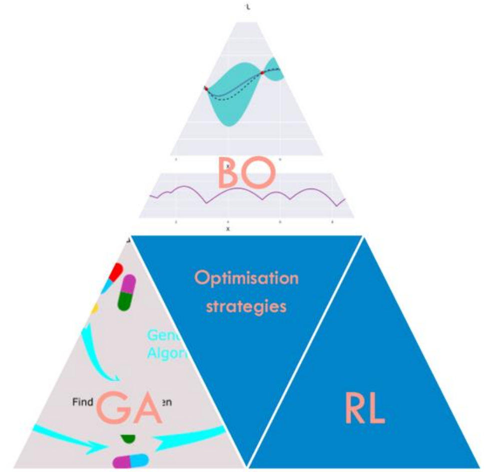
Model architecture



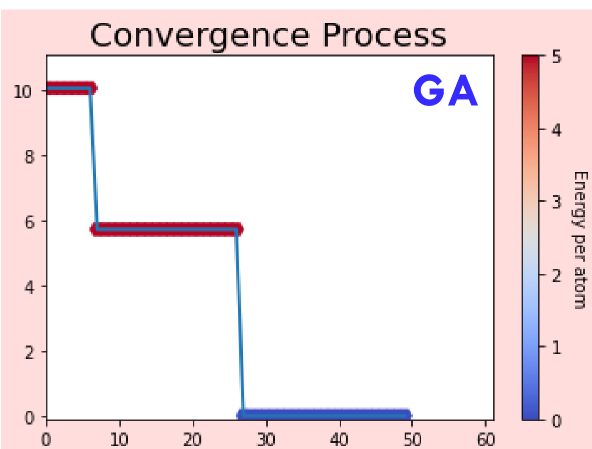
NN for Energy

M3GNet

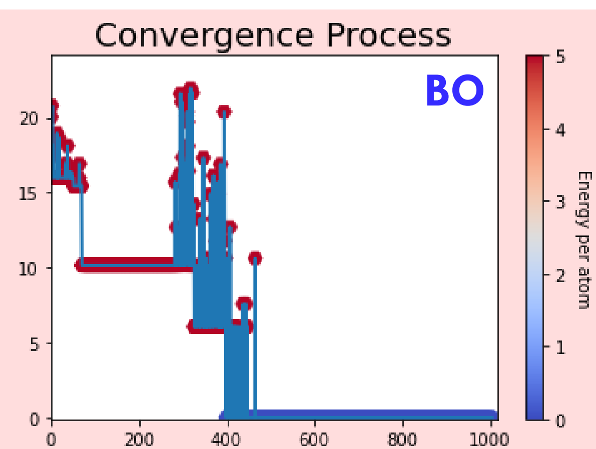
+



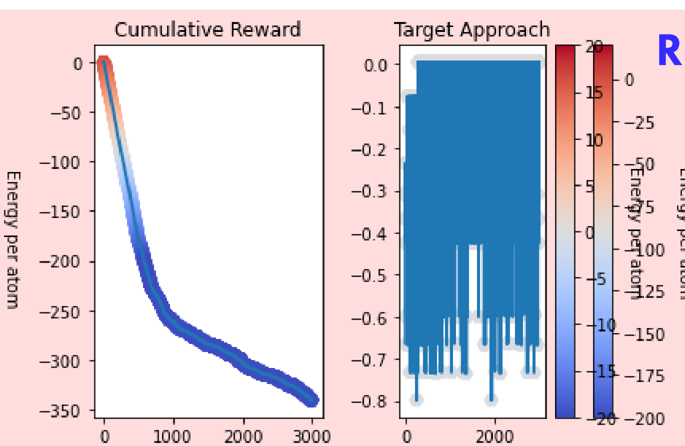
❖ Single-objective: find the lowest energy configuration



270 samples

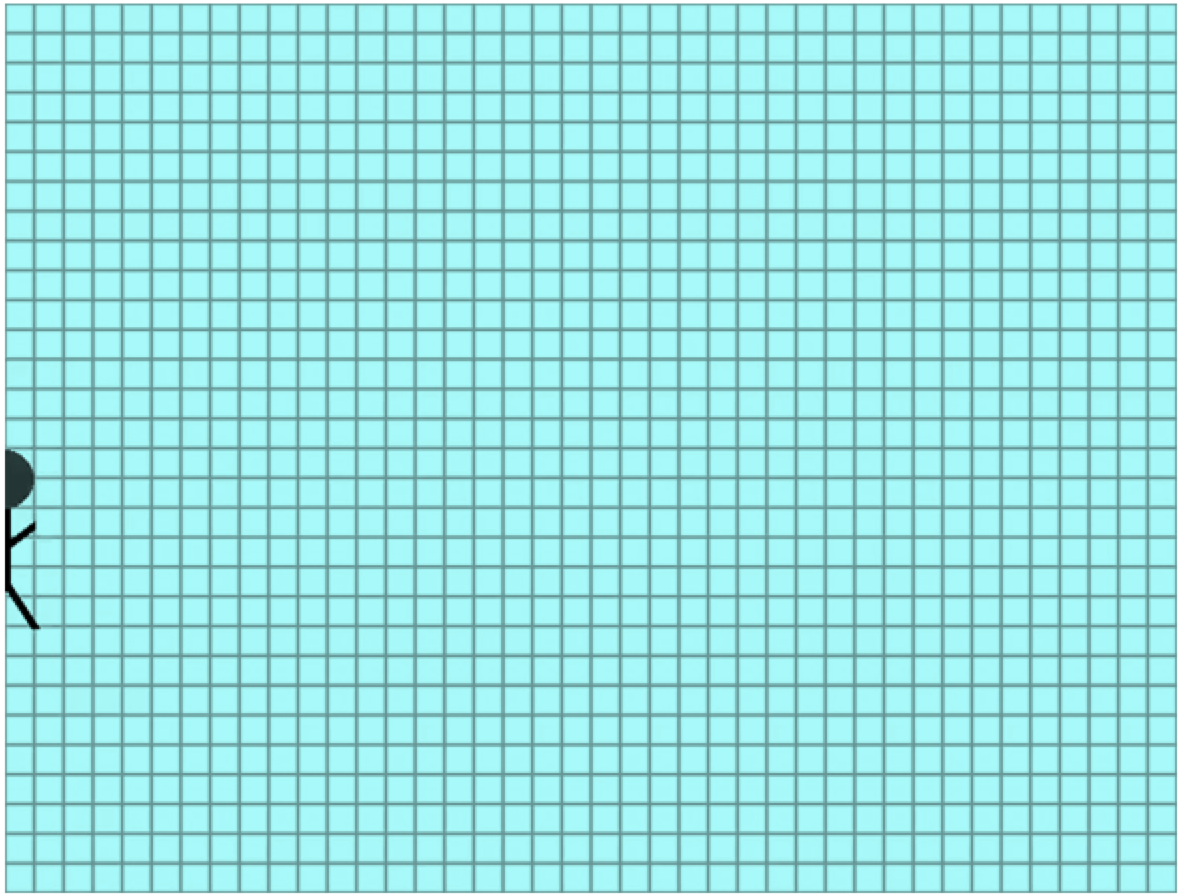


434 samples

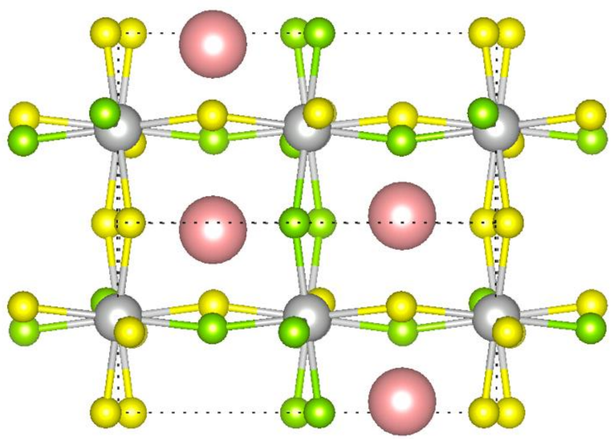


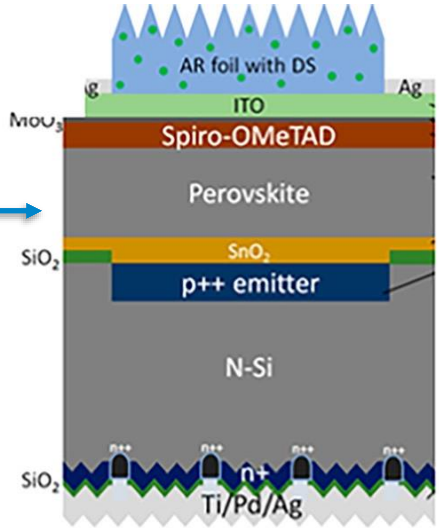
206 samples

Visualize the RL process



Stable and high power conversion efficiency photo-absorber



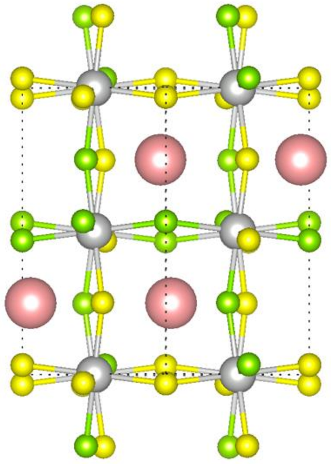


Multi-objective PV alloy design

$(BaZrS_xSe_{3-x})_4$: Find targeted composition with highest photo conversion efficiency

$$\eta = \frac{P_{\mathrm{max}}}{I_{\mathrm{in}}} = \frac{J_{\mathrm{mpp}} V_{\mathrm{mpp}}}{I_{\mathrm{in}}} = \frac{J_{\mathrm{sc}} V_{\mathrm{oc}} \mathrm{FF}}{I_{\mathrm{in}}}$$

Bandgap



Problem: large chemical search space

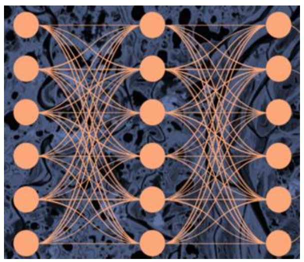
- Stability: energy
- Bandgap

- PCE: multi-objective

Multi-objective PV alloy design

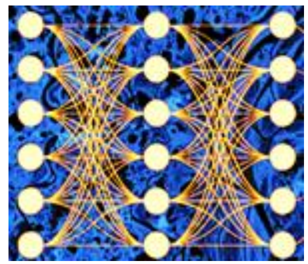
$(BaZrS_xSe_{3-x})_4$: Find targeted composition with highest photo conversion efficiency

ML model architecture



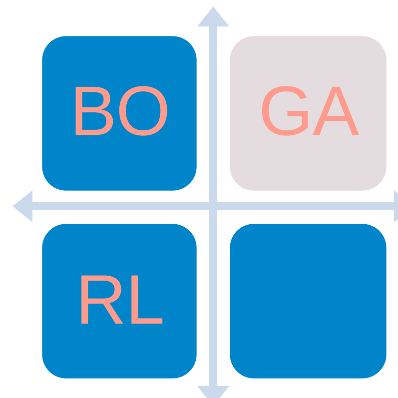
NN for Energy

\+



NN for Bandgap

\+



`M3GNet`

```
matgl.load_model("MEGNet-MP-2019.4.1-BandGap-mfi")
```

Shockley-Queisser-limit: PCE

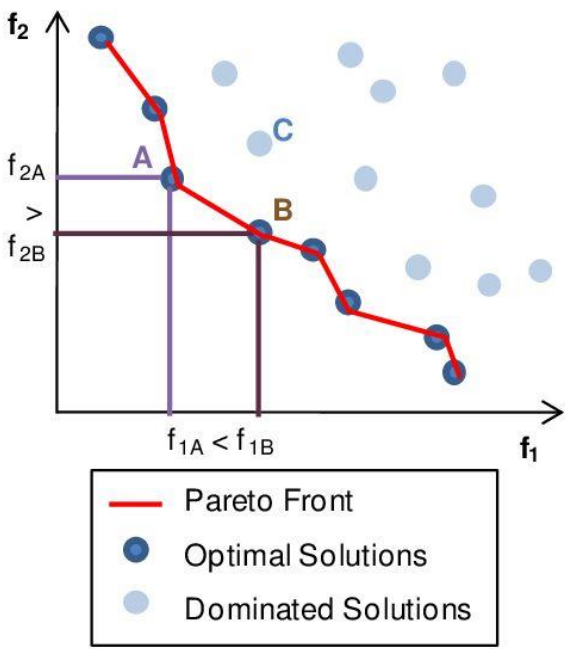
Multi-objective PV alloy design

Single Objective/Reward function

Pareto front

$$Objective = -E + PCE$$

- Cannot promise the increases of E and PCE are positively correlated.



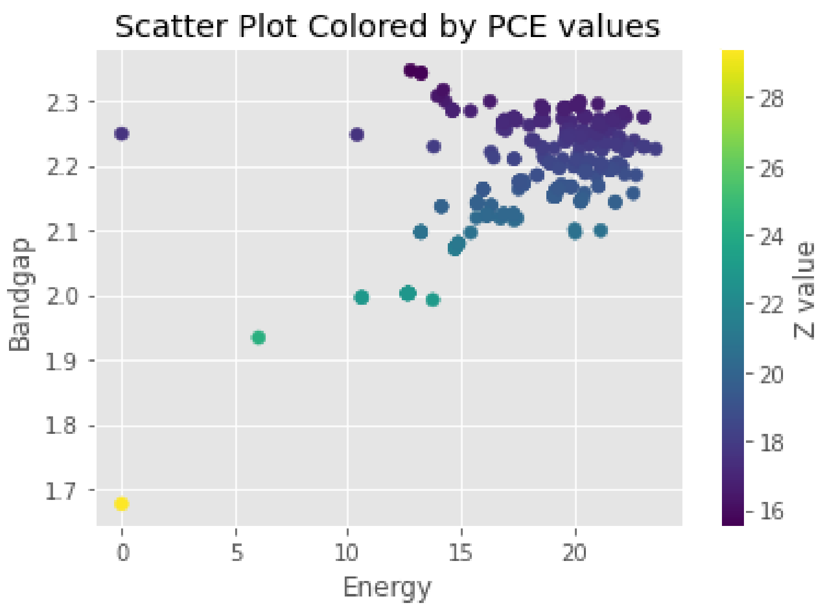
Multi-objective PV alloy design

Single Objective/Reward function — VS — Pareto front

DISTRIBUTED EVOLUTIONARY ALGORITHMS IN PYTHON

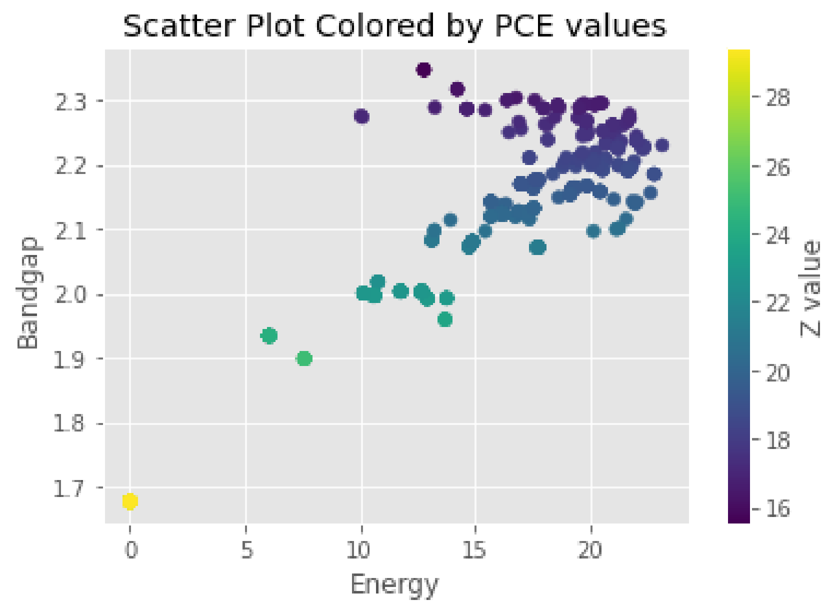
GA

Single objective function



272 samples

Pareto front



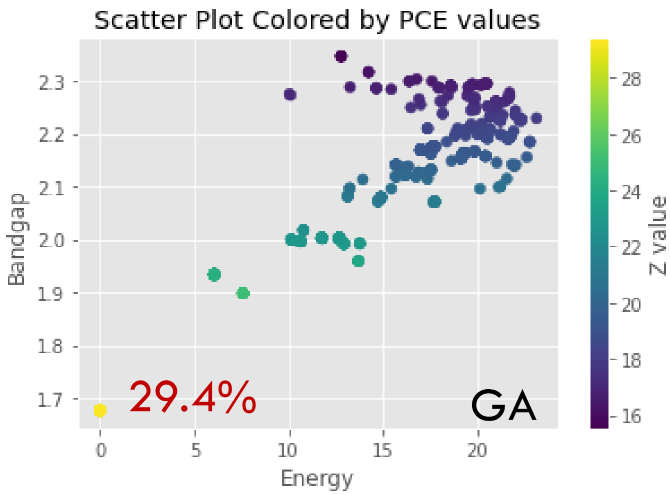
285 samples

```
pareto_front = tools.sortNondominated(population, len(population), first_front_only=True)[0]
```

Multi-objective PV alloy design

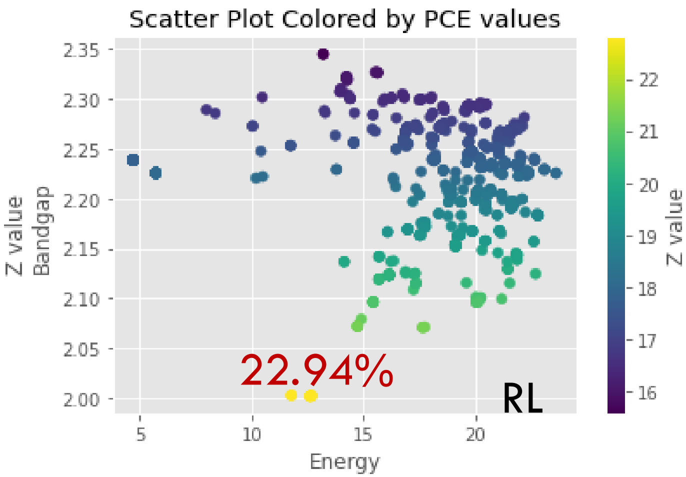
Single Objective/Reward function | VS | Pareto front

❖ Multi-objective: find the configuration with highest PCE, Pareto Front search



Pareto front

285 samples



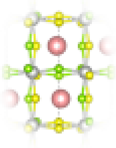
$$Objective = -E + PCE$$

336 samples

In recorded memory

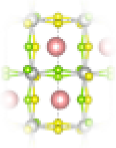
```
pareto_front = tools.sortNondominated(population,
len(population), first_front_only=True)[0]
```

Conclusions



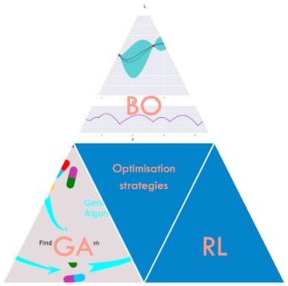
Navigated materials design

➢ Energy



Multi-objective materials design

➢ PCE & Stability



Acknowledgements

Imperial College London

Prof. Aron Walsh
Xia Liang
Sean Kavanagh
Xinwei Wang
Johan Klarbring
Youngwon Woo

Collaborators
Prof. Cecilia Mattevi, Imperial
Dr. Alex Ganose, Imperial
Dr. Ji-Sang Park, SKKU
Tian Xie, Microsoft
Alex Jen, CityU HK
Enzheng Shi, Westlake
Yunfan Guo, ZhejiangU

SCHMIDT FUTURES
NRF National Research Foundation of Korea
MATERIALS AND MOLECULAR MODELLING HUB
archer2
THE THOMAS YOUNG CENTRE
Imperial-X

Imperial College London

Thank you very much for your attention!