

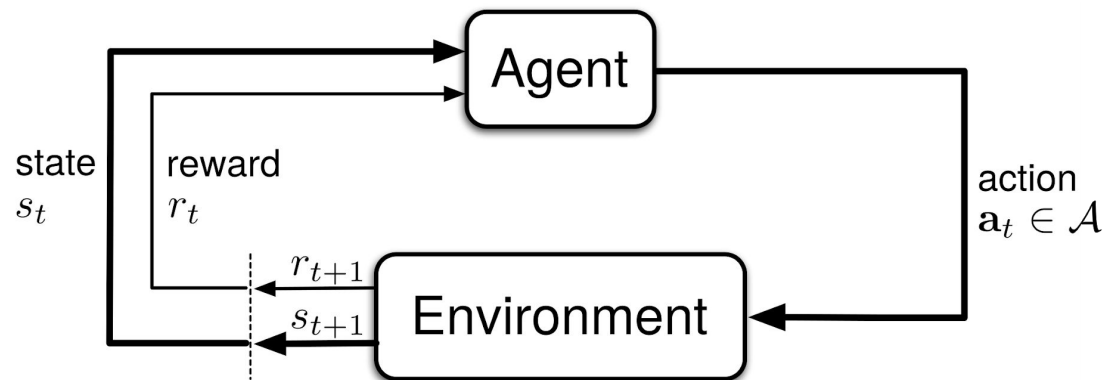


Learning to Represent Action Values as a Hypergraph on the Action Vertices

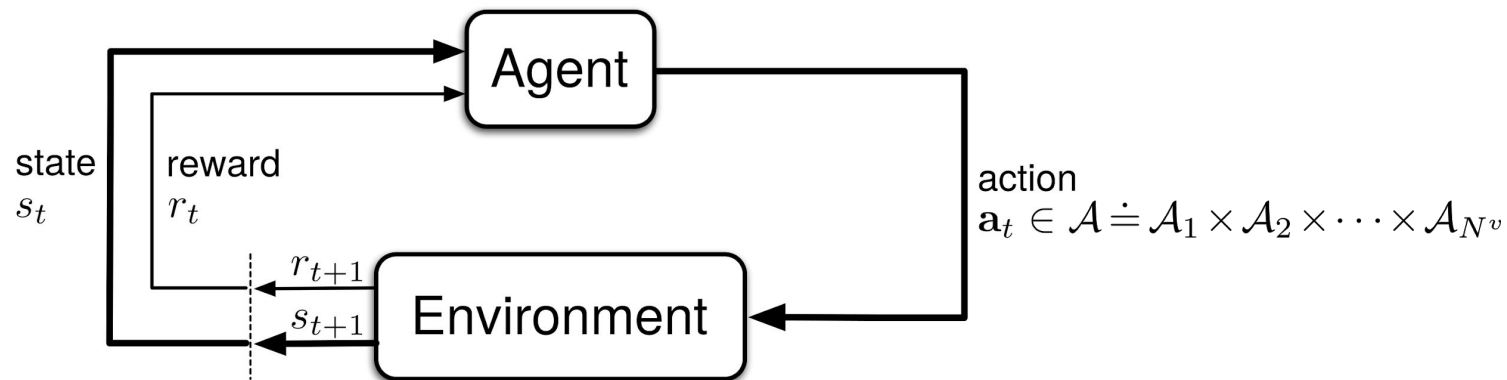
Arash Tavakoli, Mehdi Fatemi, Petar Kormushev

Imperial College London, Microsoft Research

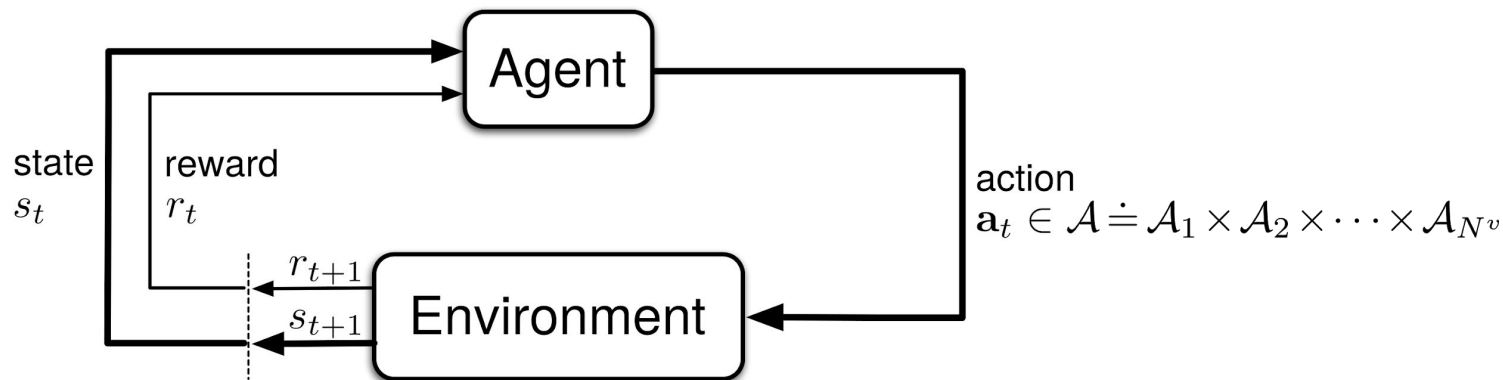
Problem Setting



Problem Setting

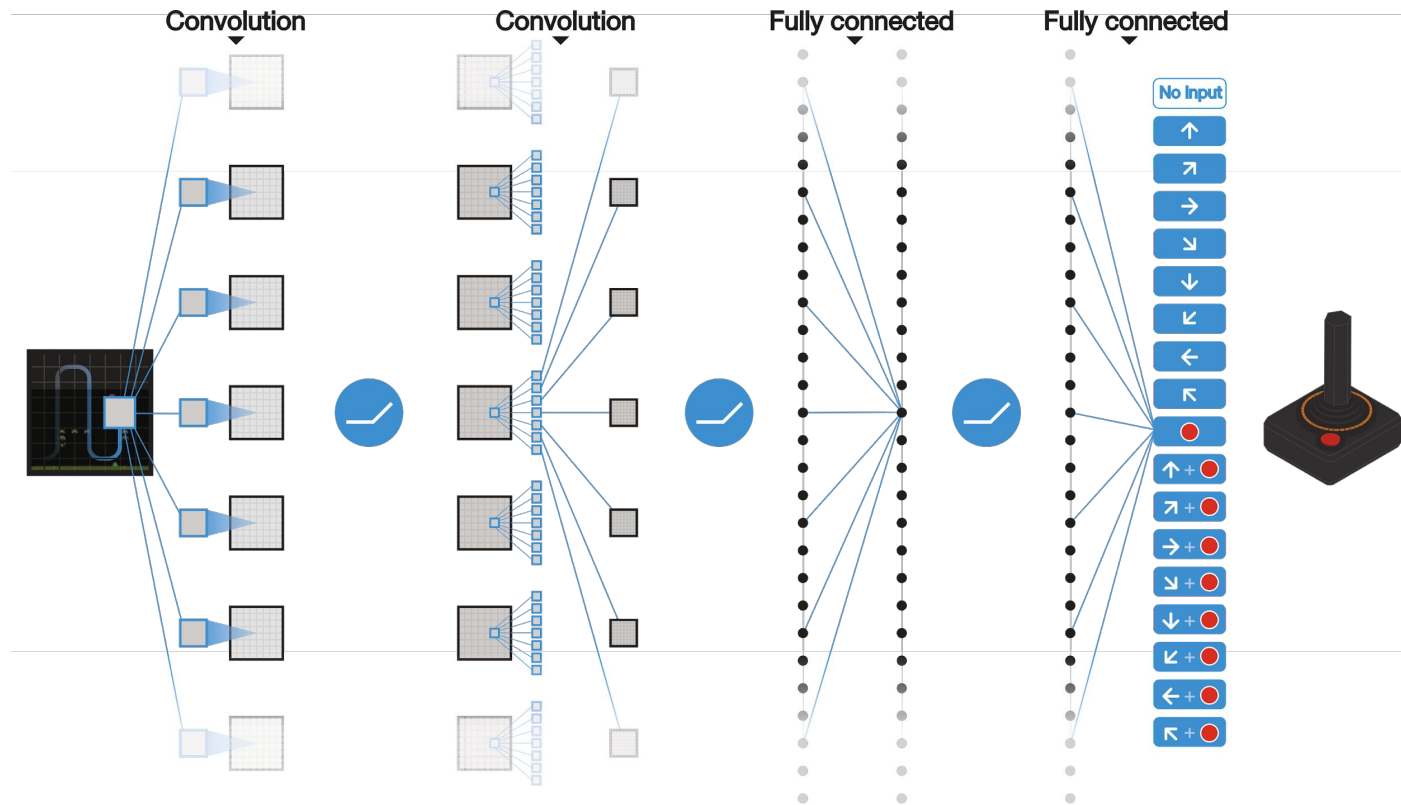


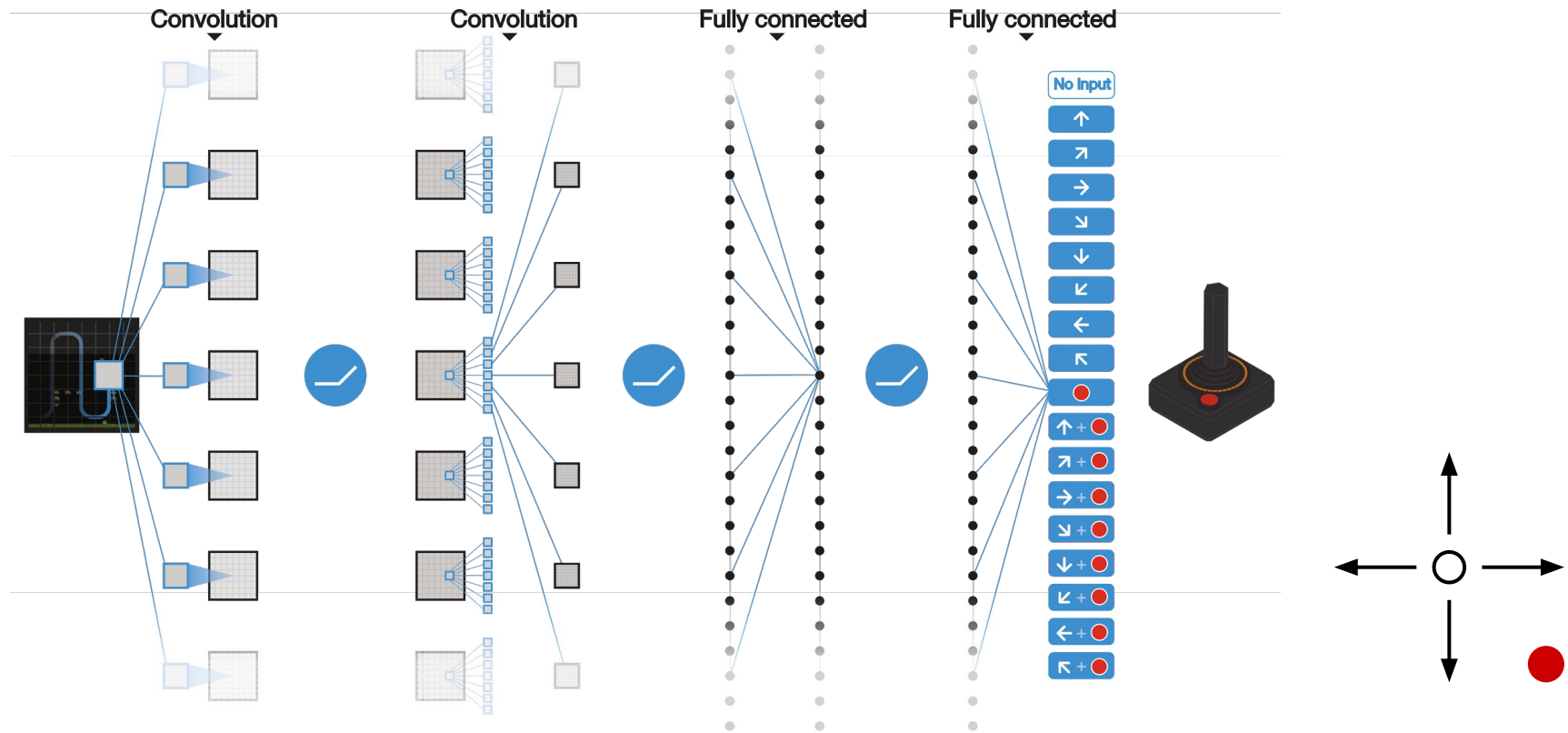
Problem Setting

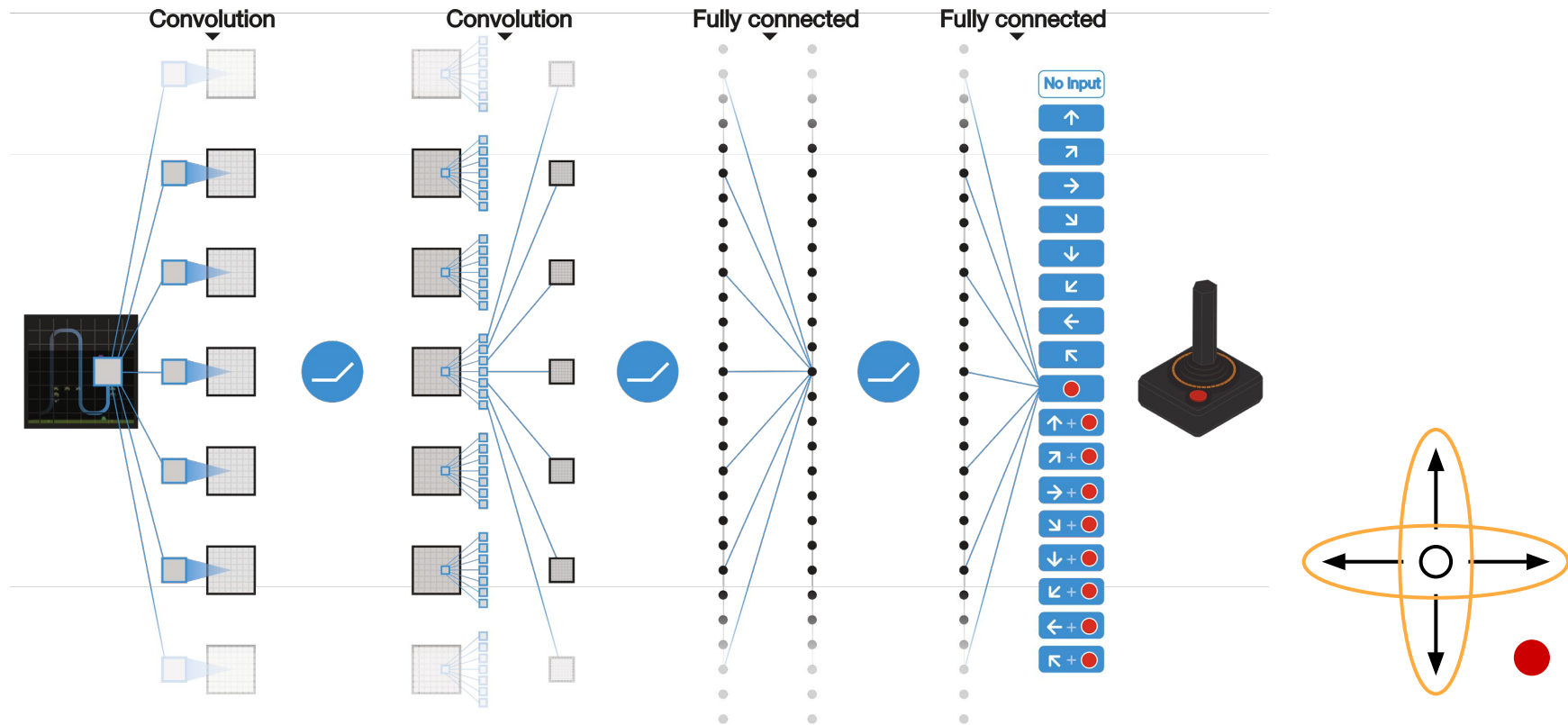


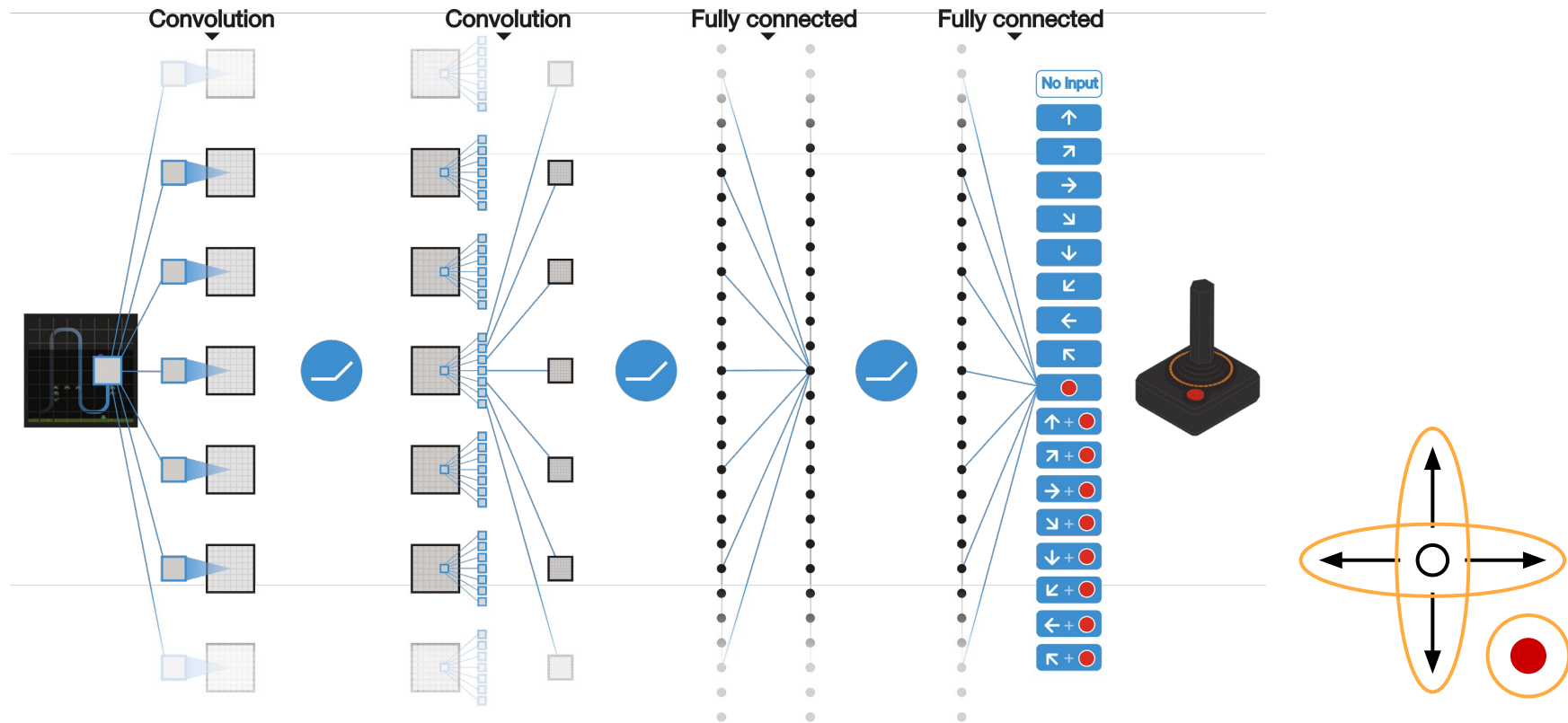
$$Z_{\pi}(s, \mathbf{a}) = \sum_{t=0}^{\infty} \gamma^t r_{t+1} \quad \text{where} \quad 0 \leq \gamma \leq 1$$

$$Q_{\pi}(s, \mathbf{a}) = \mathbb{E}[Z_{\pi}(s, \mathbf{a})]$$

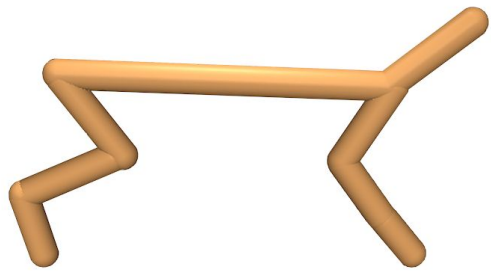




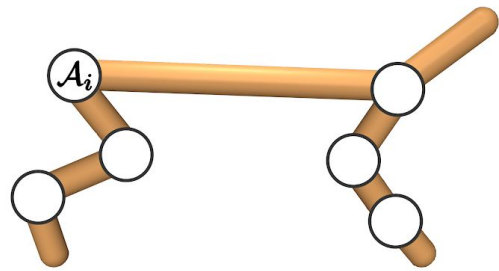




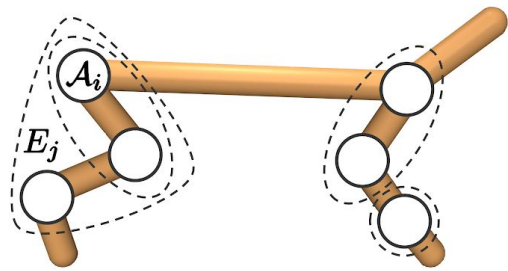
(a)



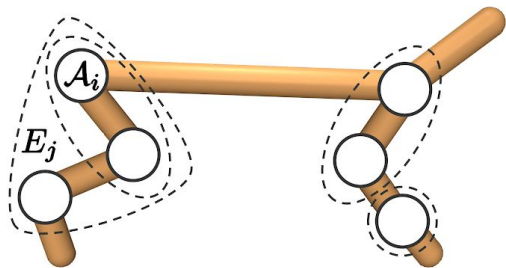
(a)



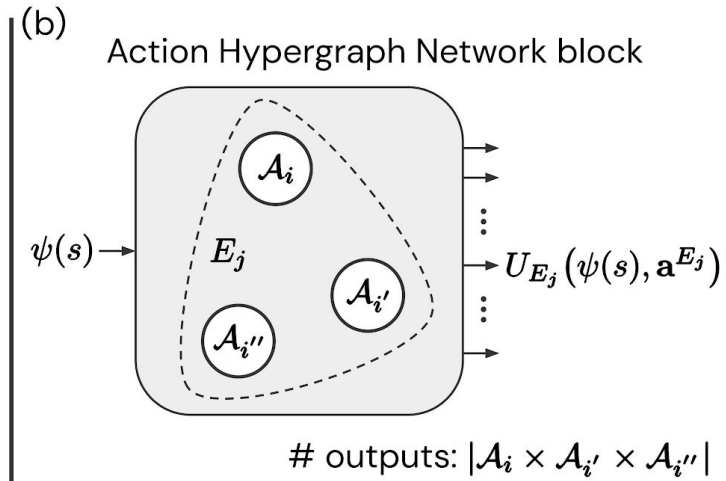
(a)



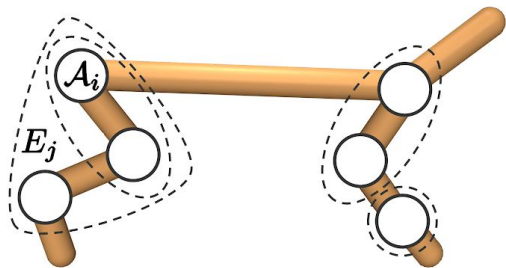
(a)



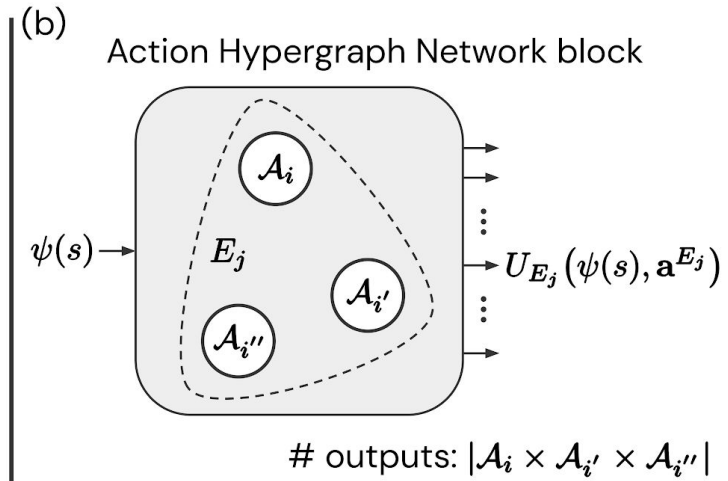
(b)



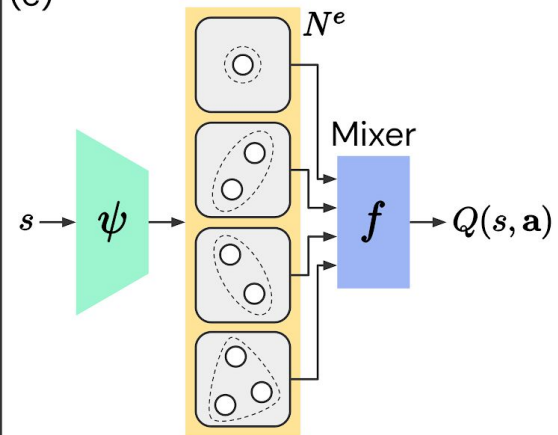
(a)



(b)

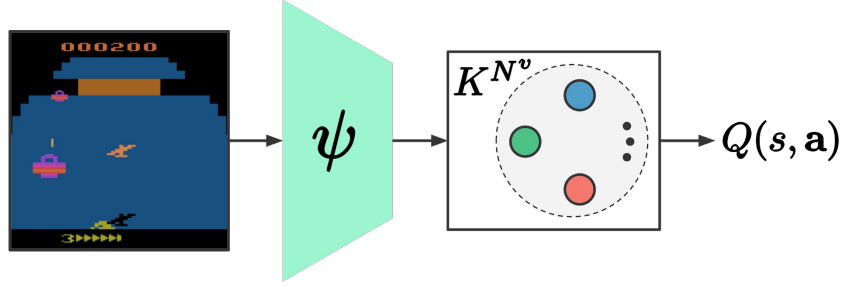


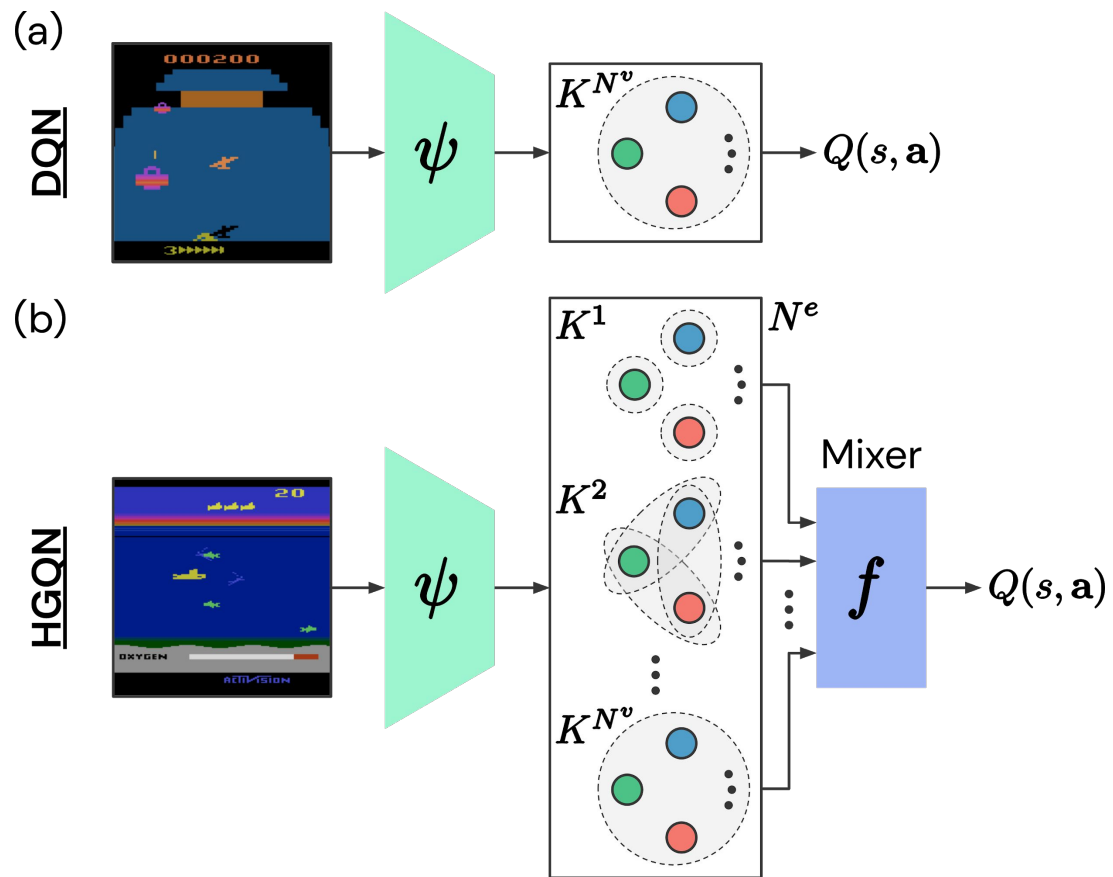
(c)



(a)

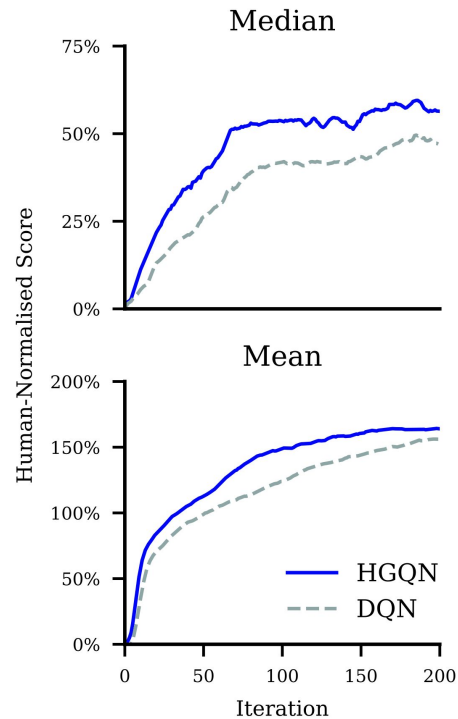
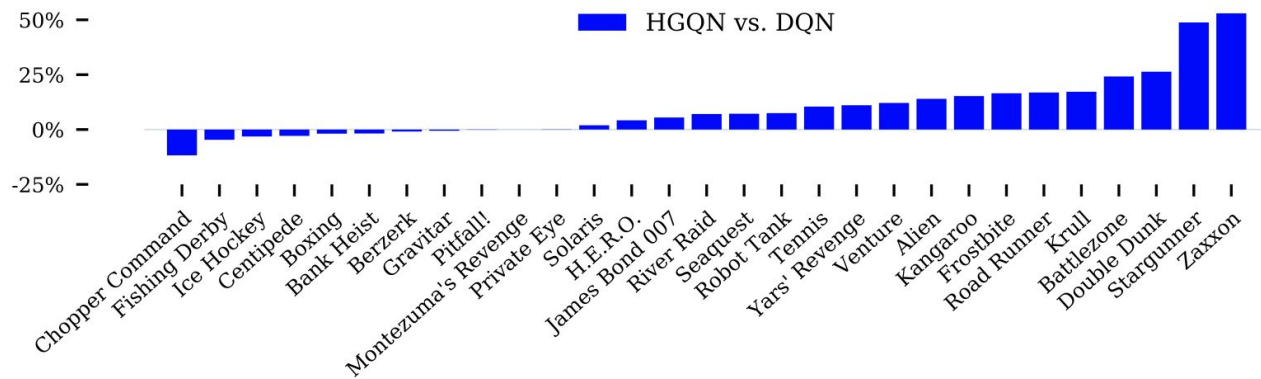
DQN



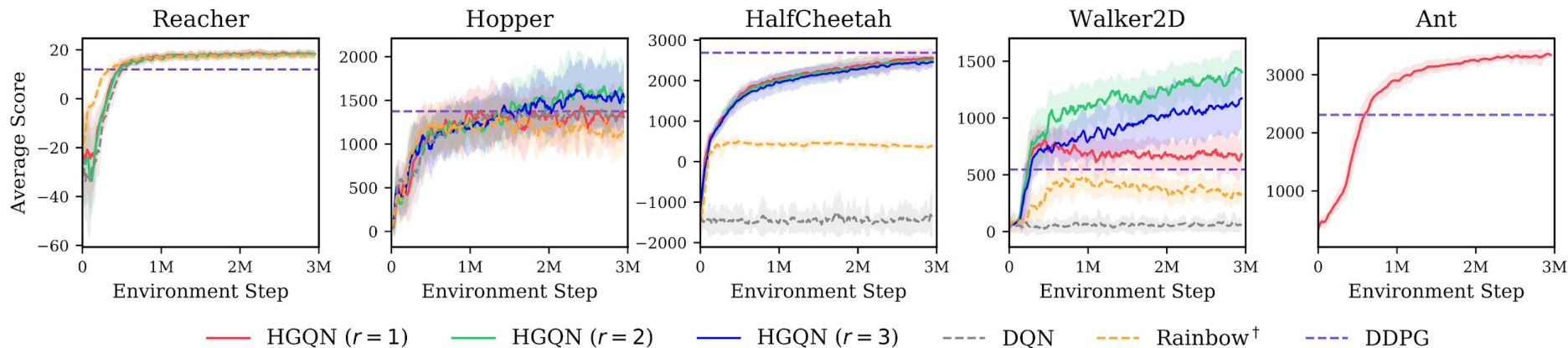


Results

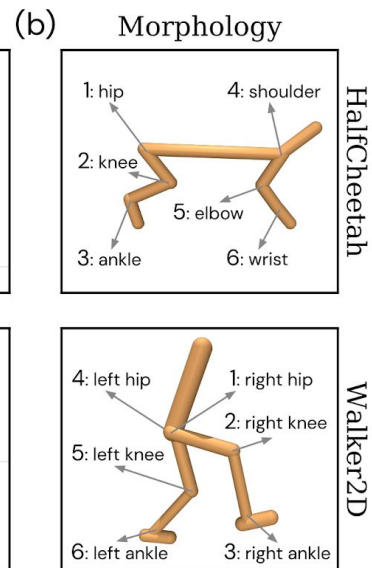
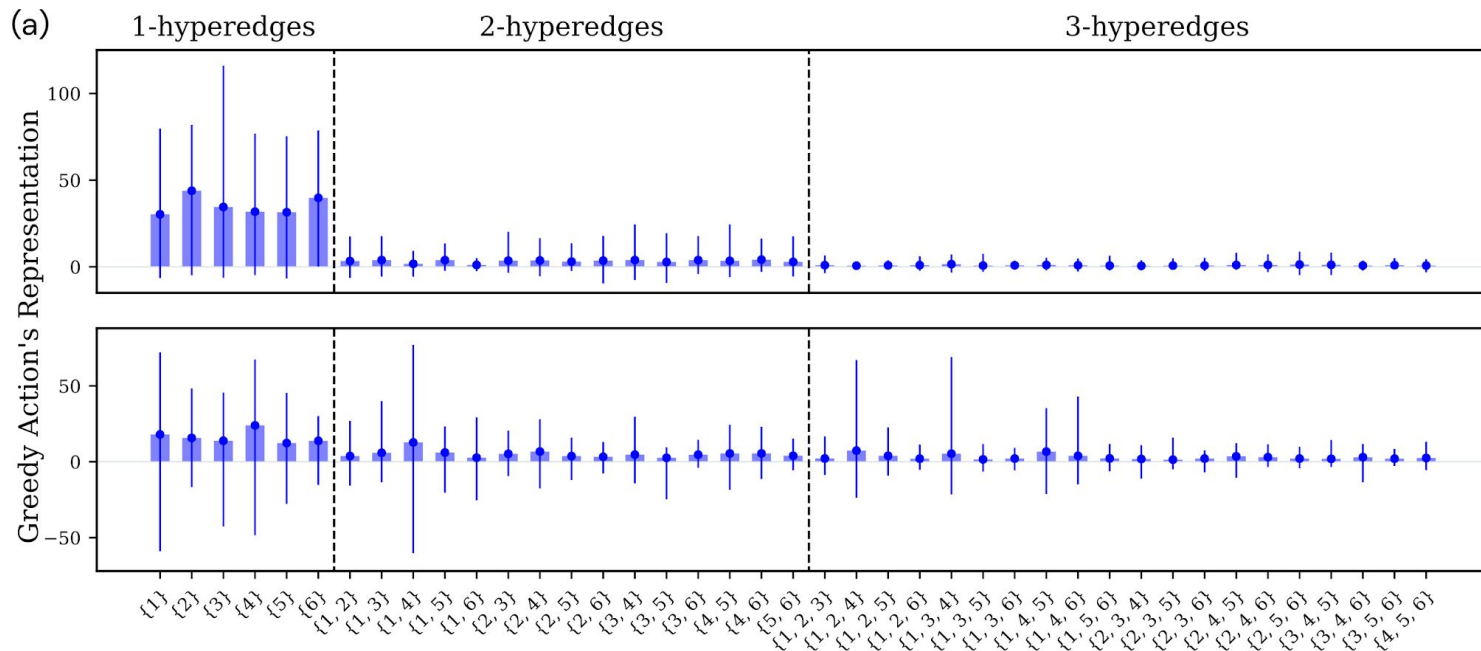
Atari 2600 games



Physical control benchmarks (PyBullet)



Analysis of learned representations



Correspondence to: Arash Tavakoli

Email: a.tavakoli@imperial.ac.uk

Webpage: atavakol.github.io

Twitter: [@arshtvk](https://twitter.com/arshtvk)