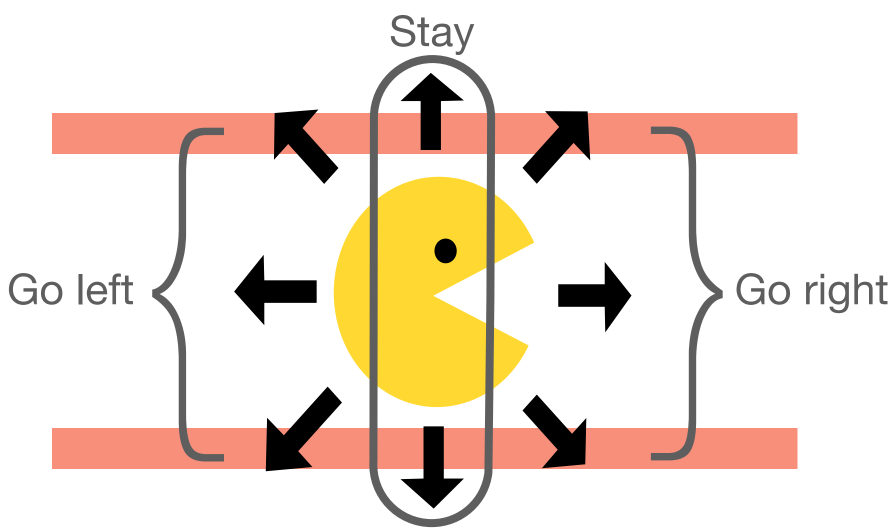
Characterizing the Action-Generalization Gap in Deep Q-Learning

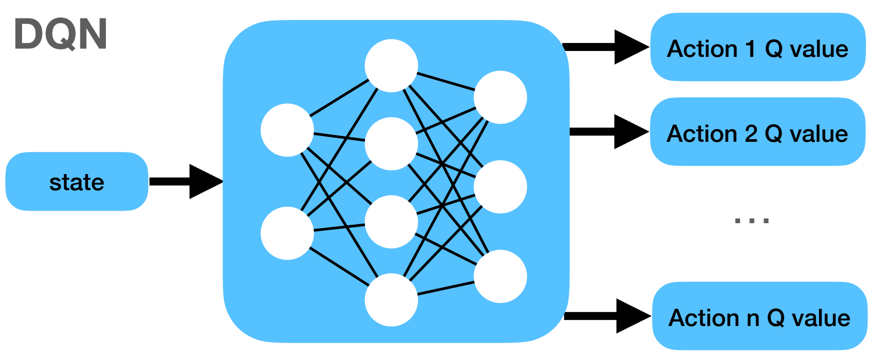
Zhiyuan Zhou, Cameron Allen, Kavosh Asadi, George Konidaris
Brown University, Amazon Web Services
zhouzy@brown.edu zhouzypaul.github.io

Paper on Arxiv



To what degree can DQN generalize over actions? How do we evaluate it?

DQN



DQN can **generalize** over discrete actions **in small action spaces,** but not larger ones.

Oracle: Evaluating Action Generalization

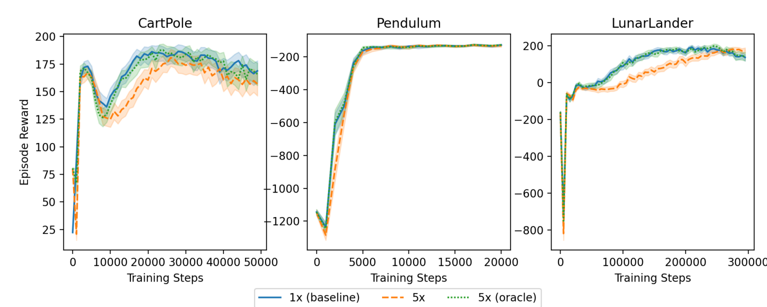
We use an **oracle** for characterizing perfect action generalization. The oracle uses **expert human knowledge** $K(a,\tilde{a}) \in [0,1]$ of **action similarity** to **adjust the Q-update process**: we not only update the experienced action a, but also every action according to its similarity to a.

$$Q(s,\tilde{a}) \rightarrow Q(s,\tilde{a}) + \alpha * K(a,\tilde{a}) * [(r + \gamma V(s')) - Q(s,\tilde{a})] \quad \forall\ \tilde{a} \in A$$

Duplicate Actions Env: ✓

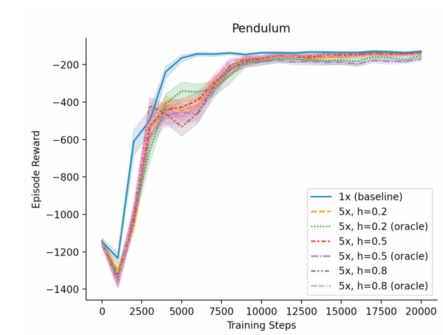
Make **5x copies** of every action

Action generalization indeed helps fast learning



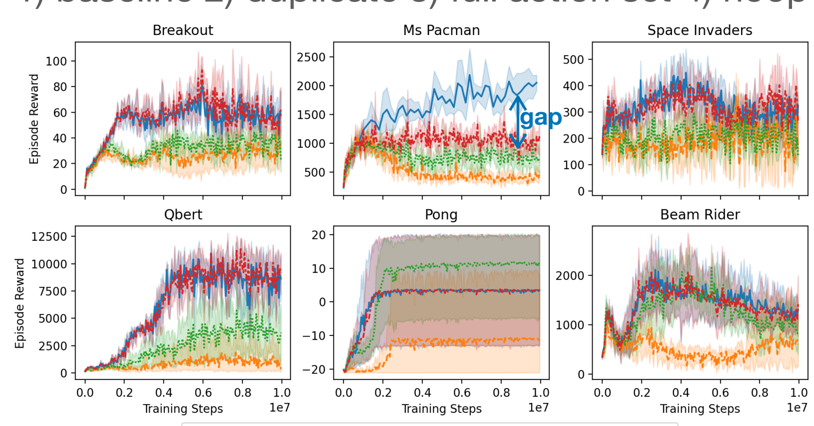
Semi-Duplicate Actions Env: ✓

Augment the original action space with 4x **reduced-magnitude actions**, where the magnitude is indexed by $h \in \{0.2, 0.5, 0.8\}$



Atari 2600: ✗

4 different sets of action spaces:
1) baseline 2) duplicate 3) full action set 4) noop



Large Duplicate Actions Env: ✗

Make **copies** of the original action set $n \in \{5, 15, 50\}$ times