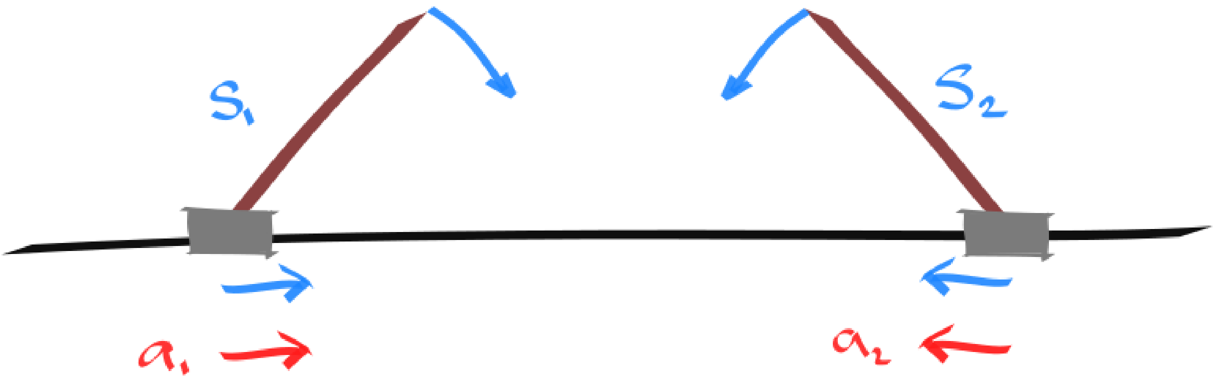
Cart pole

Mirror symmetry



Two indicators of this symmetry. It is reflected in the transition function, and the value function.

The change in state is conserved between the pair.

$$\Delta_\tau(s,a) = \mathop{\mathbb{E}}_{s'\sim p(\cdot|s,a)}(s'-s) \tag{1}$$
$$\Delta_\tau(s_1,a_1) = -\Delta_\tau(s_2,a_2) \tag{2}$$
$$\tag{3}$$

$$\Delta_T(s,a) = (T\circ Q)(s,a) - Q(s,a) \tag{4}$$
$$\Delta_T(s_1,a_1) = \Delta_T(s_2,a_2) \tag{5}$$
$$\tag{6}$$

The expected value is conserved between the pair (assuming we have a policy with mirror symmetry).

$$\forall a \text{ set } \ \pi(a|s_1) = \pi(-a|s_2) \tag{7}$$
$$\forall\gamma : Q^\gamma_\pi(s_1,a_1) = Q^\gamma_\pi(s_2,a_2) \tag{8}$$
$$\forall\gamma : Q^\gamma_\pi(s_1,a_2) = Q^\gamma_\pi(s_2,a_1) \tag{9}$$
$$\tag{10}$$

The (discounted) reachable rewards are conserved between the pair. (!!!)

$$\{r(s,a,s') : \forall s \in \mathcal{R}(s_1,a_1)\} = \{r(s,a,s') : \forall s \in \mathcal{R}(s_2,a_2)\} \tag{11}$$

Translational symmetry

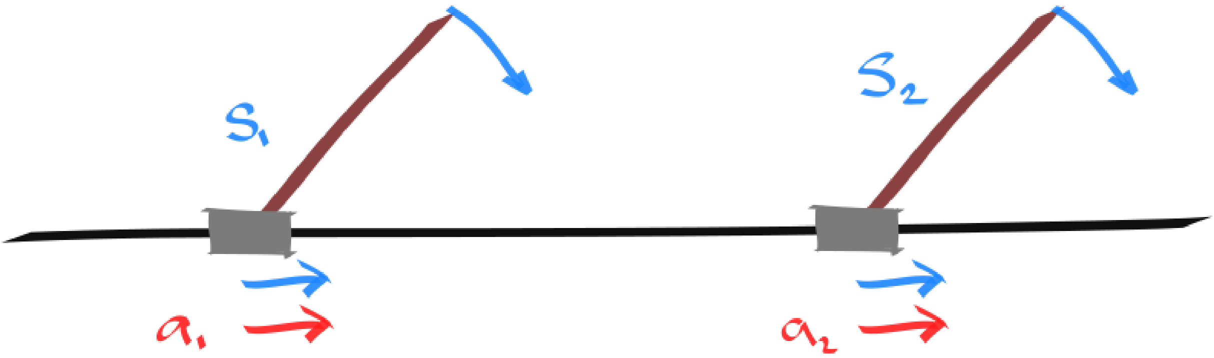


Figure 1: Each pair is similar, in a sense

(special case of regular actions)

$$\Delta(s_1,a_1) = \Delta(s_1,a_2) = \Delta(s_2,a_1) = \Delta(s_2,a_2) \tag{12}$$

$$\forall a \text{ set } \pi(a|s_1) = \pi(a|s_2) \tag{13}$$

$$\forall\gamma : Q^\gamma_\pi(s_1,a_1) = Q^\gamma_\pi(s_2,a_2) = Q^\gamma_\pi(s_1,a_2) = Q^\gamma_\pi(s_2,a_1), \tag{14}$$

$$\tag{15}$$

Future translational symmetry

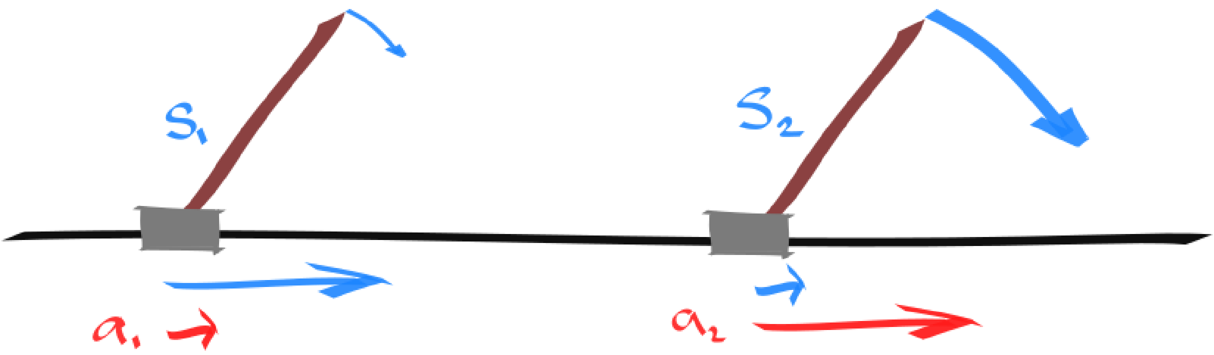


Figure 2: Each pair is similar, in a sense

different states, different actions. but maps into translational symmetry.

After this action. All future actions will have the same effect. In this sense, these two state-actions are similar.

$$\forall a : \mathop{\mathbb{E}}_{s' \sim p(\cdot|s_1,a_1)}[\Delta(s', a)] = \mathop{\mathbb{E}}_{s' \sim p(\cdot|s_2,a_2)}[\Delta(s', a)] \tag{16}$$

$$\tag{17}$$

Temporal mirror symmetry

This is simply a result of the eariler mirror symmetry?!? (want to show this!)

permutations of actions that yield similar outcomes.

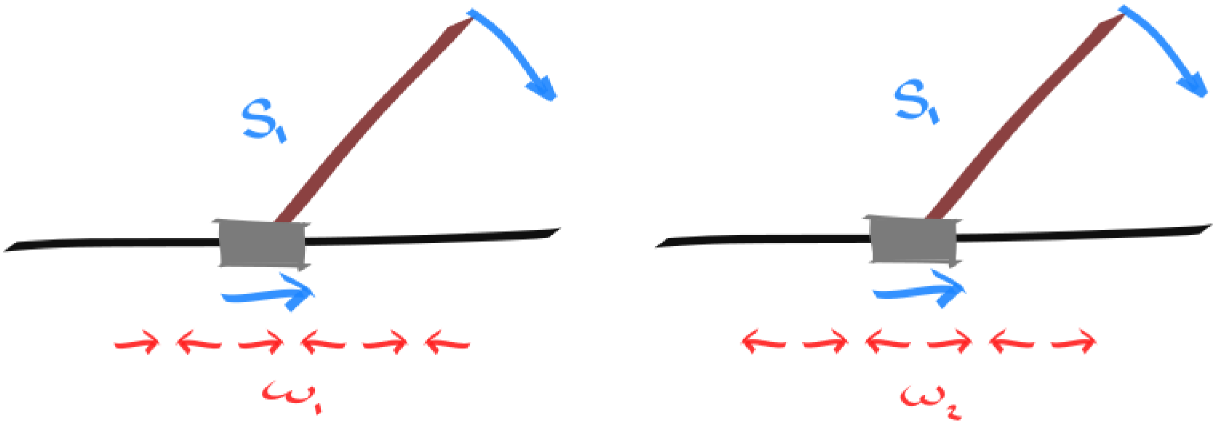
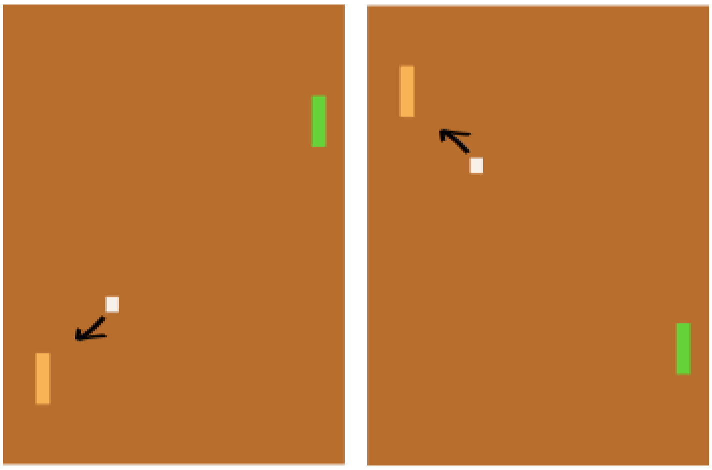


Figure 3: Each pair is similar, in a sense

$$p(s'|s, \omega) = \prod p(s|s, a)\omega(a|s) \tag{18}$$
$$p(\cdot|s_1, \omega_1) = p(\cdot|s_1, \omega_2) \tag{19}$$
$$Q(s_1, \omega_1) = Q(s_1, \omega_2) \tag{20}$$
$$\tag{21}$$

Pong

Mirror symmetry (vertical)



$$\Delta_\tau(s,a) = \mathop{\mathbb{E}}_{s' \sim p(\cdot|s,a)} (s' - s) \tag{22}$$

$$\Delta_\tau(s_1, a_1) = -\Delta_\tau(s_2, a_2) \tag{23}$$

$$\tag{24}$$

$$\Delta_T(s,a) = (T \circ Q)(s,a) - Q(s,a) \tag{25}$$

$$\Delta_T(s_1, a_1) = \Delta_T(s_2, a_2) \tag{26}$$

$$\tag{27}$$

The expected value is conserved between the pair (assuming we have a policy with mirror symmetry).

$$\forall a \text{ set } \; \pi(a|s_1) = \pi(-a|s_2) \tag{28}$$

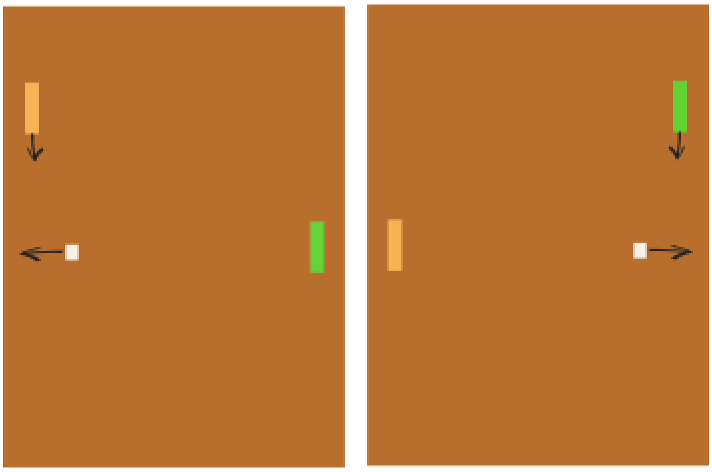
$$\forall \gamma : Q^\gamma_\pi(s_1, a_1) = Q^\gamma_\pi(s_2, a_2) \tag{29}$$

$$\forall \gamma : Q^\gamma_\pi(s_1, a_2) = Q^\gamma_\pi(s_2, a_1) \tag{30}$$

$$\tag{31}$$

(do these invariances uniquely define this mirror symmery?!)

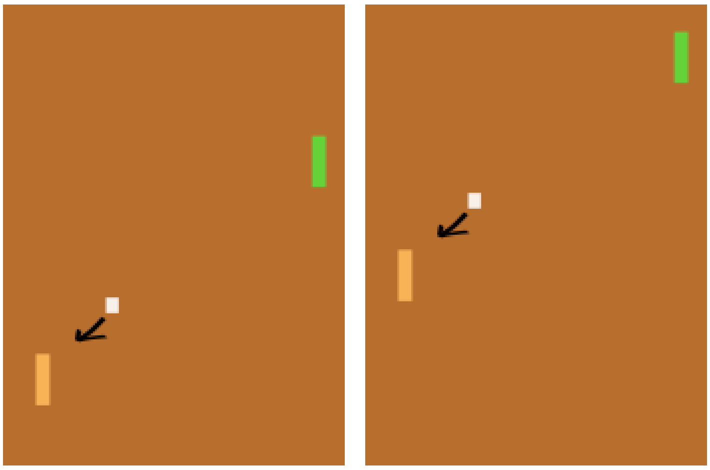
Mirror symmetry (horizontal)



$$Q(s_1, a_1) = -Q(s_2, a_2) \tag{32}$$

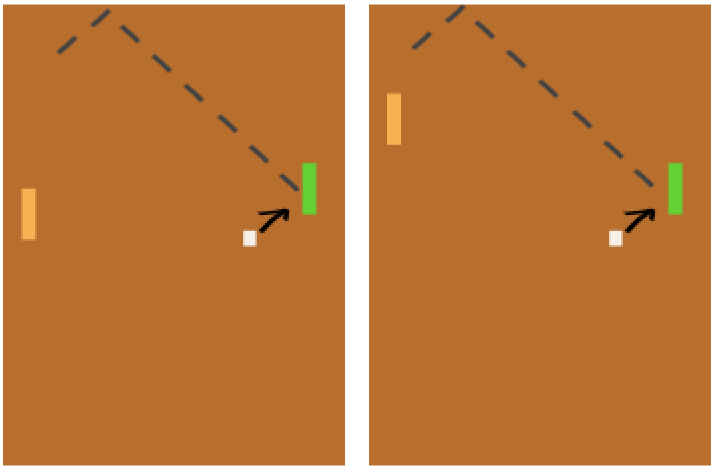
Translational symmetry

(although there are boundary cases which cannot be ignored. how can they be dealt with!?)



$$\forall a : Q(s_1, a) = Q(s_2, a) \tag{33}$$

Temporal symmetries



$$\exists \pi_1, \pi_2 \;\; \text{s.t.} \;\; Q^{\pi_1}(s_1, a_1) = Q^{\pi_2}(s_2, a_2) \tag{34}$$

The same future state can be reached, and thus the same rewards can be achieved.