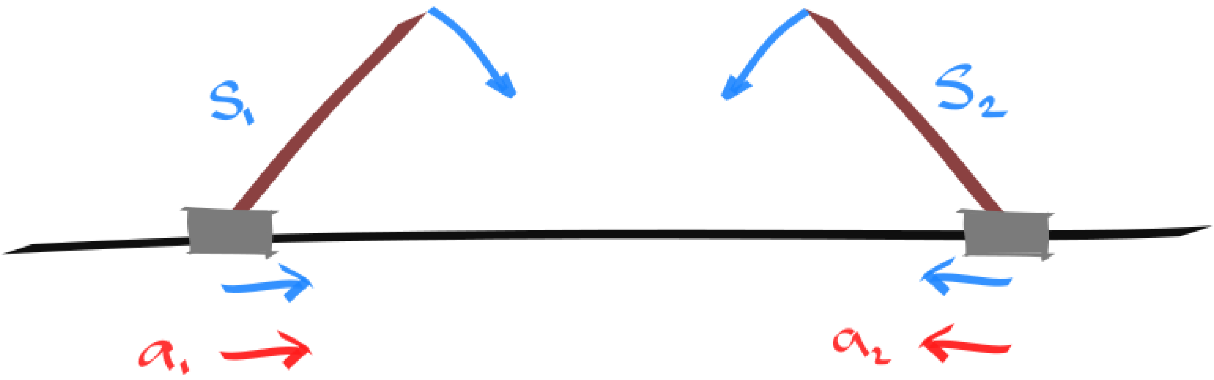
Cart pole

Mirror symmetry



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

(*this assumes we have a "nice" state representation where differenes make sense*)

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

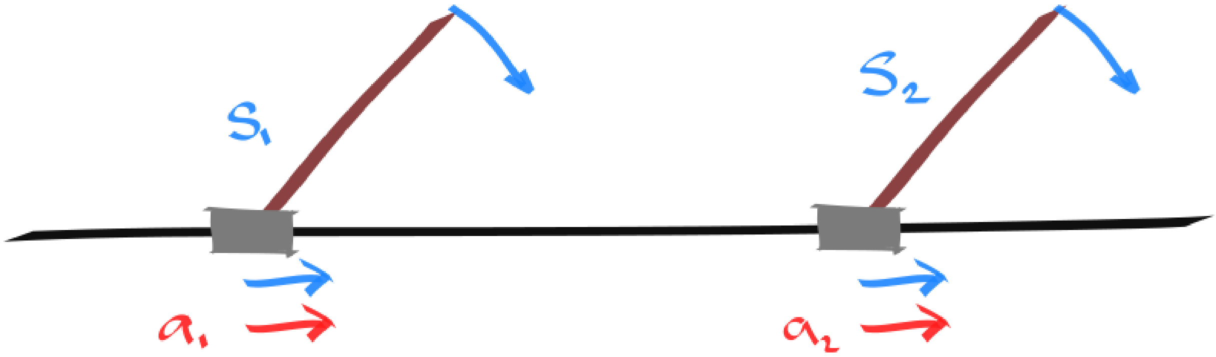
The expected value is conserved between the pair (assuming we have a policy with mirror symmetry).

$$\text{set} \quad \pi(a|s) = \pi(-a|-s) \tag{7}$$
$$Q_\pi(s_1,a_1) = Q_\pi(s_2,a_2) \tag{8}$$
$$Q_\pi(s_1,a_2) = Q_\pi(s_2,a_1) \tag{9}$$
$$\tag{10}$$

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



(special case of regular actions)

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

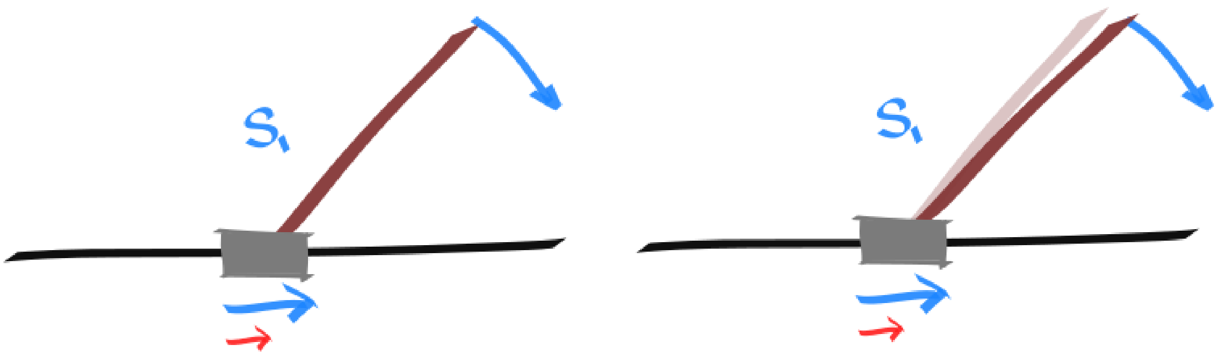
$$\text{if } \forall a \;\; \pi(a|s_1) = \pi(a|s_2) \tag{13}$$

$$Q_\pi(s_1,a_1) = Q_\pi(s_2,a_2) = Q_\pi(s_1,a_2) = Q_\pi(s_2,a_1) \tag{14}$$

(15)

Local symmetry

(*this is approximately a symmetry*)



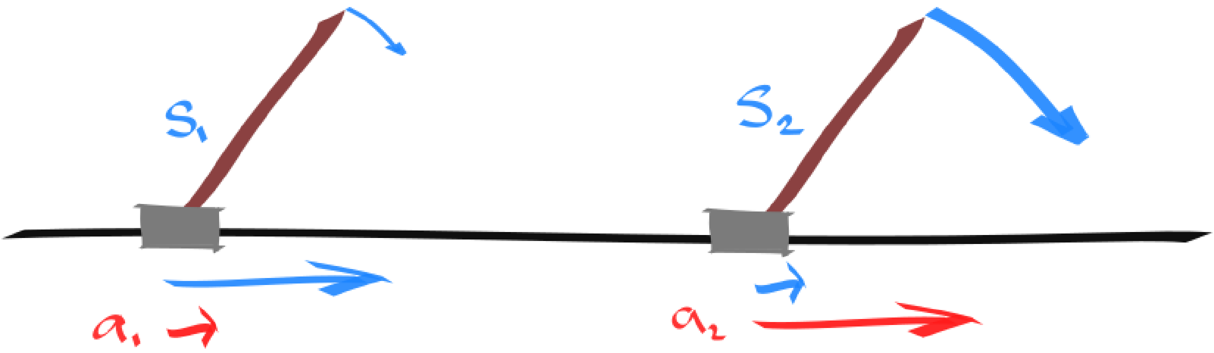
$$\Delta(s_1, a_1) = \Delta(s_1, a_2) = \Delta(s_2, a_1) = \Delta(s_2, a_2) \tag{16}$$

$$\forall a \text{ set } \pi(a|s_1) = \pi(a|s_2) \tag{17}$$

$$Q_\pi(s_1, a_1) \approx Q_\pi(s_2, a_2) \tag{18}$$

$$\tag{19}$$

Future translational symmetry



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.

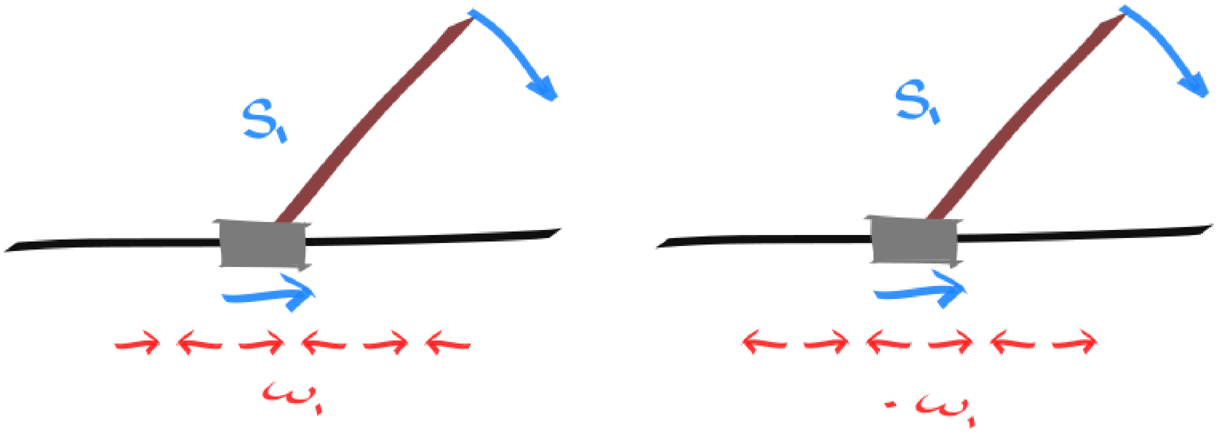
$$\forall a : \mathop{\mathbb{E}}_{s' \sim p(\cdot|s_1, a_1)} [\Delta(s', a)] = \mathop{\mathbb{E}}_{s' \sim p(\cdot|s_2, a_2)} [\Delta(s', a)] \tag{20}$$

$$\tag{21}$$

Temporal mirror symmetry

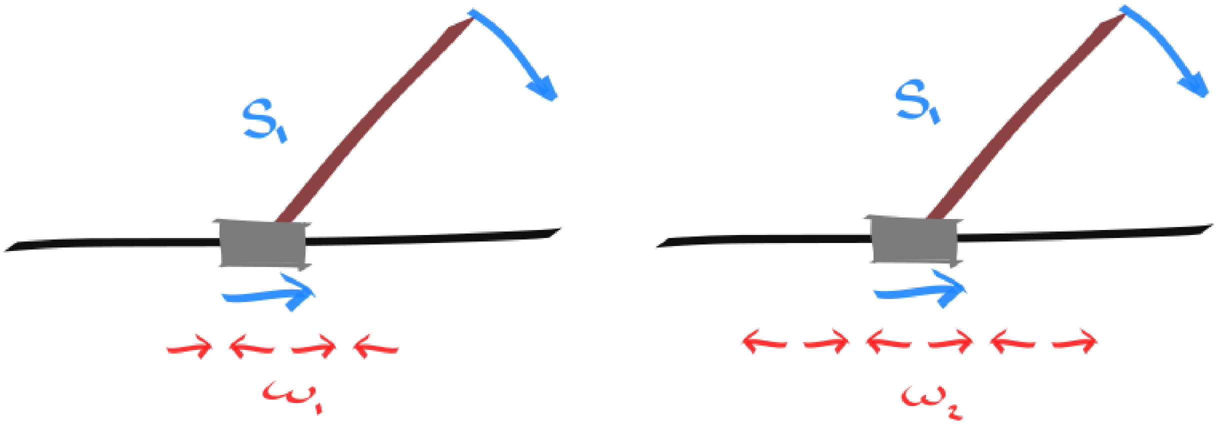
This is simply a result of the earlier mirror symmetry?!? (want to show this!)

permutations of actions that yield similar outcomes.



$$
\begin{aligned}
p(s'|s,\omega) &= \prod p(s|s,a)\omega(a|s) && (22)\\
p(\cdot|s_1,\omega_1) &= p(\cdot|s_1,-\omega_1) && (23)\\
Q_\pi(s_1,\omega_1) &= Q_\pi(s_1,-\omega_1) && (24)\\
& && (25)
\end{aligned}
$$

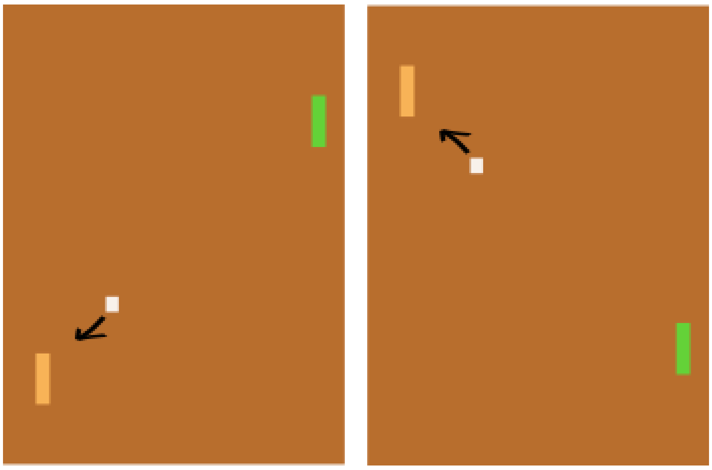
Temporal symmetry



$$
\begin{aligned}
p(\cdot|s_1,\omega_1) &= p(\cdot|s_1,\omega_2) && (26)\\
Q_\pi(s_1,\omega_1) &= Q_\pi(s_1,\omega_2) && (27)\\
& && (28)
\end{aligned}
$$

Pong

Mirror symmetry (vertical)



(*how can a change in state be evaluated? we need a representation…*)

$$\Delta_T(s,a) = (T \circ Q)(s,a) - Q(s,a) \tag{29}$$

$$\Delta_T(s_1,a_1) = \Delta_T(s_2,a_2) \tag{30}$$

$$\tag{31}$$

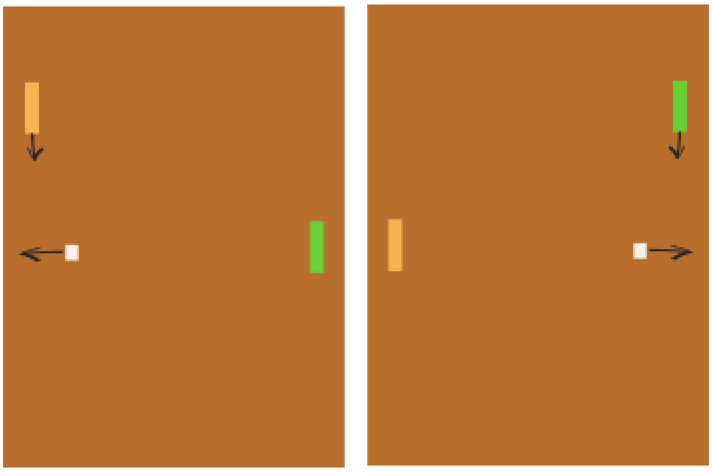
$$\forall a \text{ set } \pi(a|s_1) = \pi(-a|s_2) \tag{32}$$

$$Q_\pi(s_1,a_1) = Q_\pi(s_2,a_2) \tag{33}$$

$$Q_\pi(s_1,a_2) = Q_\pi(s_2,a_1) \tag{34}$$

$$\tag{35}$$

Mirror symmetry (horizontal)



Upon pretending to play as your opponent (flipping the image and inverting the colors, via $\rho : O \to O$, and ??? the actions)

$$Q(s_1, a_1) = -V(\rho(s_2)) \tag{36}$$

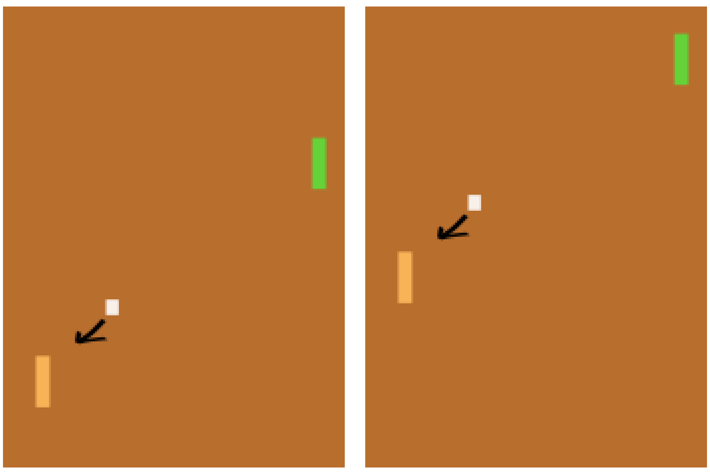
$$\tag{37}$$

(*this really requires you to disentangle the model from your opponent!?*)

$$\tau(s'|s, a_{p=0}) = f_{p=0}(s''|s, a) \cdot f_{p=1}(s'|s'', \hat{a}) \cdot \pi_{p=1}(\hat{a}|s'') \tag{38}$$

$$Q_{p=0}(s, a) = -Q_{p=1}(s, a) \tag{39}$$

Translational symmetry

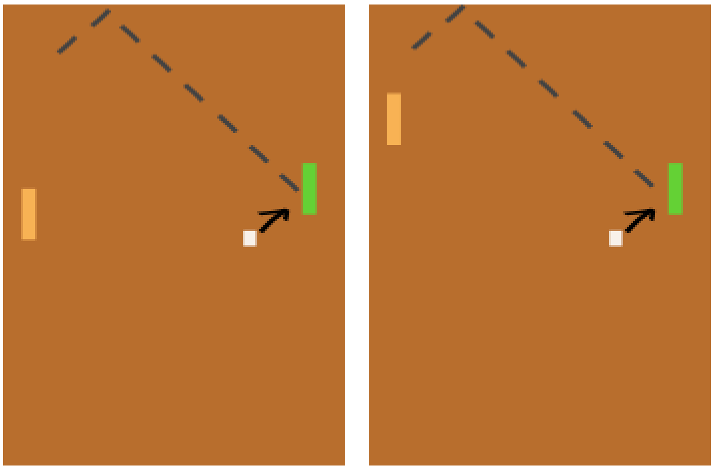


(*although there are boundary cases which cannot be ignored. how can they be dealt with!?*)

$$\forall a : Q_{\pi_1}(s_1, a) = Q_{\pi_2}(s_2, a) \tag{40}$$
$$\forall a, t : \pi_1(a|s^t_{s^0=s_1}) = \pi_2(a|s^t_{s^0=s_2}) \tag{41}$$

If we take the same actions, in translated states, we get the same outcome (up to the boundary conditions).

Temporal symmetries



$$\exists \pi_1, \pi_2 \text{ s.t. } Q^{\pi_1}(s_1, a_1) = Q^{\pi_2}(s_2, a_2) \tag{42}$$

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