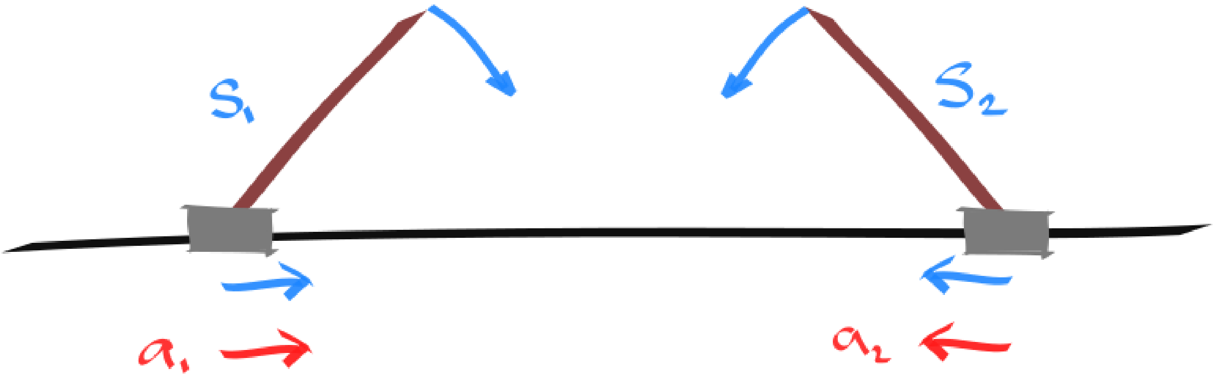
## Cart pole

### 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.

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

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

$$\forall a \text{ set } \ \pi(a|s_1) = \pi(-a|s_2) \tag{4}$$
$$\forall \gamma : Q^{\gamma}_{\pi}(s_1, a_1) = Q^{\gamma}_{\pi}(s_2, a_2) \tag{5}$$
$$\forall \gamma : Q^{\gamma}_{\pi}(s_1, a_2) = Q^{\gamma}_{\pi}(s_2, a_1) \tag{6}$$
$$\tag{7}$$

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

$$\{r(s, a, s') : \forall s \in \mathcal{R}(s_1, a_1)\} = \{r(s, a, s') : \forall s \in \mathcal{R}(s_2, a_2)\} \tag{8}$$

#### Translational symmetry

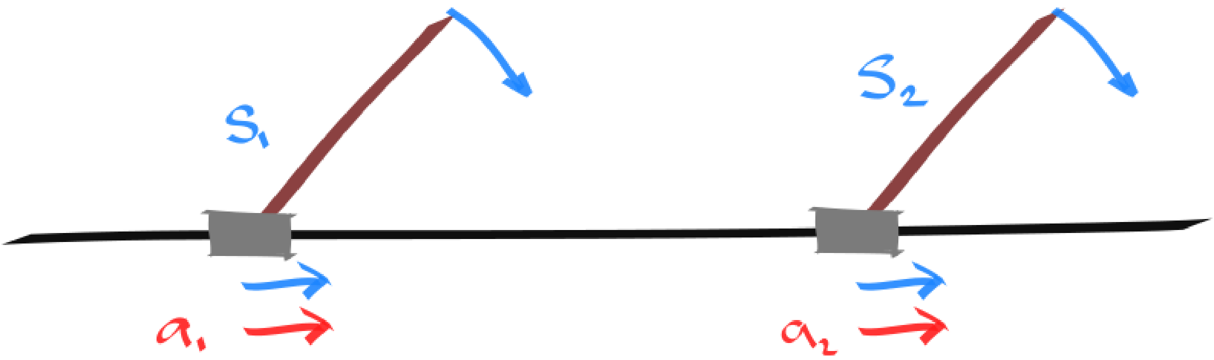



**Figure 1:** Each pair is similar, in a sense

(special case of regular actions)

$$\Delta(s_1, a_1) = \Delta(s_1, a_2) = \Delta(s_2, a_1) = \Delta(s_2, a_2) \tag{9}$$
$$\forall a \text{ set } \pi(a|s_1) = \pi(a|s_2) \tag{10}$$
$$\forall \gamma : Q^{\gamma}_{\pi}(s_1, a_1) = Q^{\gamma}_{\pi}(s_2, a_2) = Q^{\gamma}_{\pi}(s_1, a_2) = Q^{\gamma}_{\pi}(s_2, a_1), \tag{11}$$
$$\tag{12}$$

#### Future translational symmetry

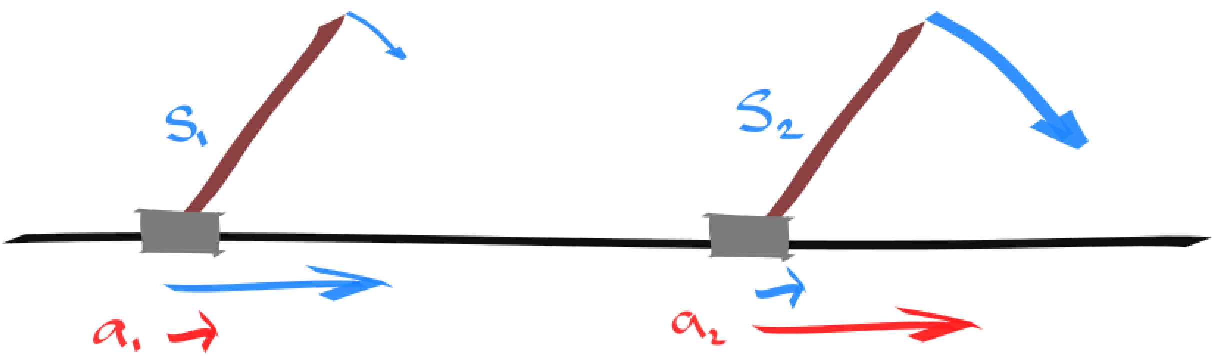



**Figure 2:** Each pair is similar, in a sense

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.

$$\forall a : \underset{s' \sim p(\cdot|s_1,a_1)}{\mathbb{E}}[\Delta(s',a)] = \underset{s' \sim p(\cdot|s_2,a_2)}{\mathbb{E}}[\Delta(s',a)] \tag{13}$$

$$\tag{14}$$

**Temporal mirror symmetry**

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

permutations of actions that yield similar outcomes.

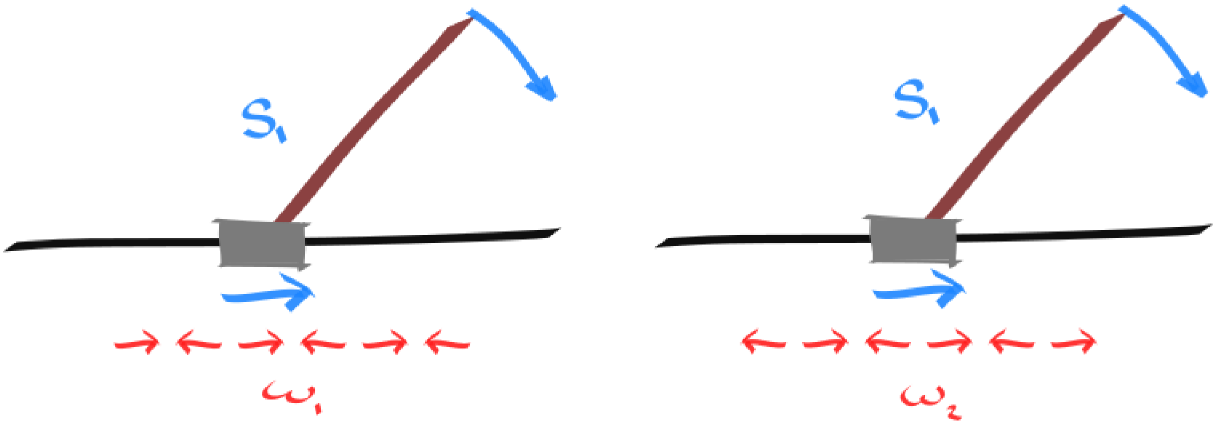



**Figure 3:** Each pair is similar, in a sense

$$p(s'|s,\omega) = \prod p(s|s,a)\omega(a|s) \tag{15}$$

$$p(\cdot|s_1,\omega_1) = p(\cdot|s_1,\omega_2) \tag{16}$$

$$Q(s_1,\omega_1) = Q(s_1,\omega_2) \tag{17}$$

$$\tag{18}$$