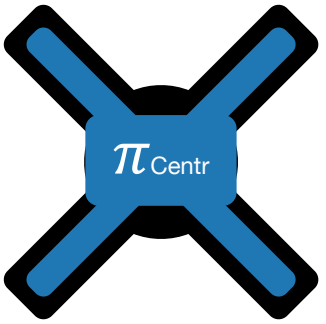
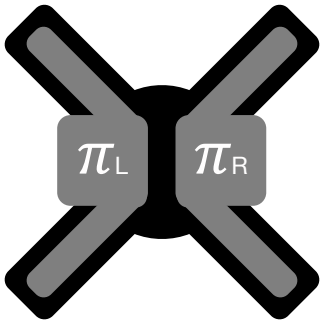
a) Concurrency of control:

Splitting up the Action-Space & Rewards

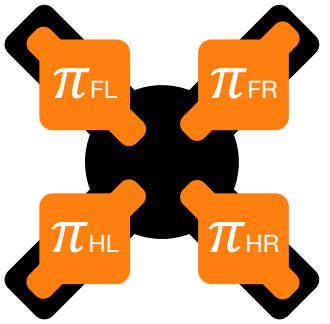
Centralized Controller ⟷ Fully Decentralized Contr.



1 policy
for 8 joints



2 policies,
each for 2 legs, 4 joints



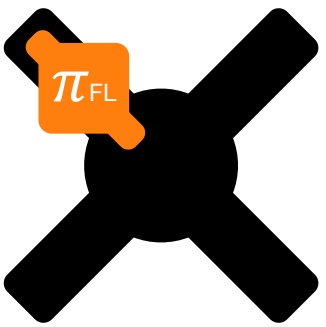
4 policies,
each for 1 leg, 2 joints

$$\text{Rew} = x_{vel} - 0.5 \sum_{i=1..8} \|a_i\|^2 \quad \text{Rew}_{\text{L}} = \frac{1}{2} x_{vel} - 0.5 \sum_{\text{FL,HL}} \|u\|^2 \quad \text{Rew}_{\text{FL}} = \frac{1}{4} x_{vel} - 0.5 \sum_{\text{FL}} \|u\|^2$$

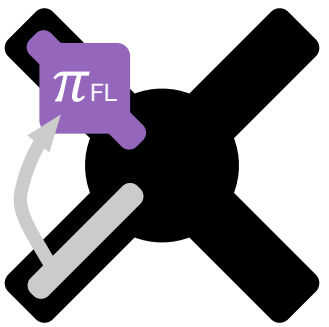
b) Input scope (for decentralized arch.):

Differentiation of Observation Space

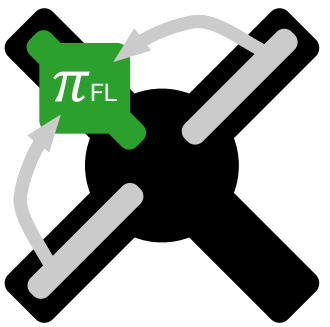
Fully Decentralized | Single Neighb. Input | Local Input (both neighb.)



Each Policy: only input from
the controlled leg



Input: from controlled leg +
single neighboring leg



Local Input: from controlled
leg + both neighb. legs