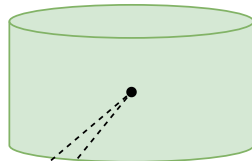


Agent

Goal

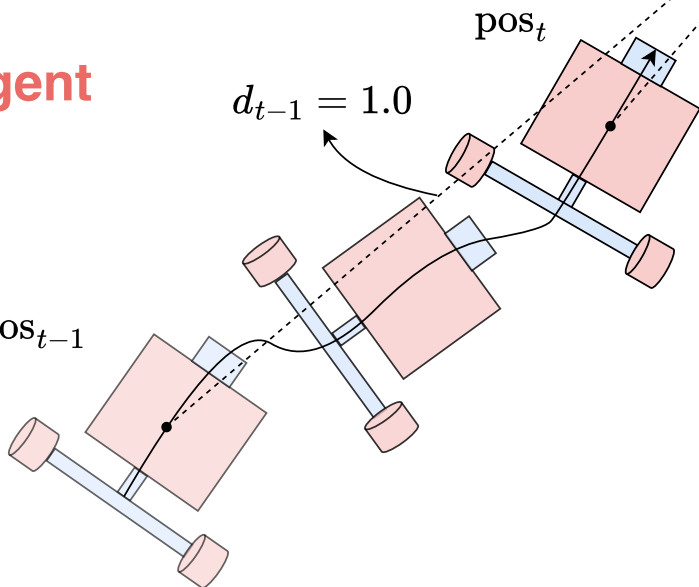


pos_t

$d_{t-1} = 1.0$

$d_t = 0.5$

pos_{t-1}



reward = 0.5