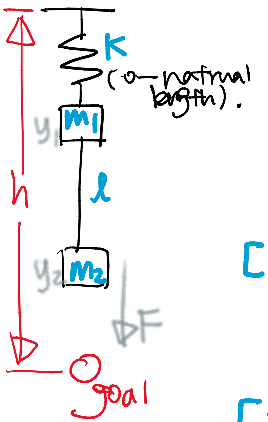
Toy Problem — Mass Spring

Sunday, January 5, 2020 10:28 PM



obs: y_1 (y_2.).

goal: put m_2 to the goal point with min. effort

reward: $-\alpha(y_2-h)^2-\beta F^2$.

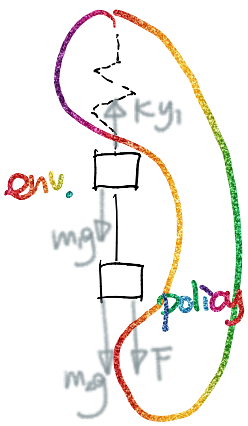
[If only optimize comp. policy].

$F=F(y_1,y_2)$

F do not converge to 0.

So reward is always negative

[If opt. comp. policy & k].



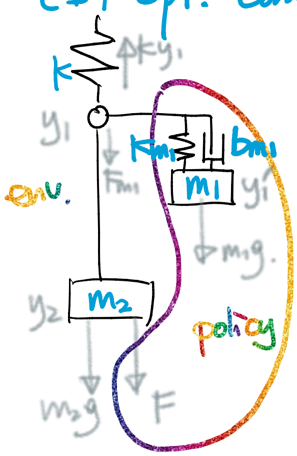
policy: $\pi = F - \underbrace{k}_{\text{param.}}\underbrace{y_1}_{\text{obs.}}$

env. dynamics: $\pi + (m_1+m_2)g = (m_1+m_2)\ddot{y}_1$

expected optimal soln:

$$\frac{(m_1+m_2)g}{k^*}+l=h \Rightarrow k^*=\frac{(m_1+m_2)g}{h-l}.$$

[If opt. comp policy & m_1].



Move m_1 out of env. and into policy.

Add spring & damper in the interface

Add an obs: y_1'

Policy: $\pi = F + F_{m_1}$

$F_{m_1} = -k_{m_1}(y_1'-y_1) - b_{m_1}(\dot{y}_1'-\dot{y}_1)$

policy dynamics:

$m_2g - F_{m_1} = m_1\ddot{y}_1'$ $\quad y_1', \dot{y}_1', \ddot{y}_1'$ are kept track of

(with recurrent nodes in the comp. graph) env. dynamics:

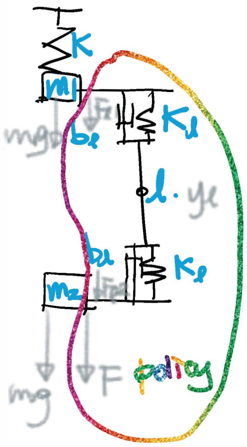
$$m_2 g + F + F_{m_1} - K y_1 = m_2 \ddot{y}_2$$

expected optimal soln:

$$m_1^* = \frac{(h-l)K}{g} - m_2.$$

then $y - h \to 0$, $F \to 0$.

[If opt. comp. policy & l].



Move the bar into the policy.
add springs & dampers.
*Add an obs: y_l in policy sys.
Policy: $\pi = F + F_{l_1} + F_{l_2}$

$$F_{l_1} = -K_l\left(y_l - \frac{l}{2} - y_1\right) - b_l(\dot{y}_l - \dot{y}_1)$$
$$F_{l_2} = -K_l\left(y_l + \frac{l}{2} - y_2\right) - b_l(\dot{y}_l - \dot{y}_1)$$

Policy dynamics: infinitely fast (no mass)

$$y_l = \frac{1}{2}(y_1 + y_2)$$

need more discussion
this is actually ill-defined dynamics
we can also give it some mass

env. dynamics:

$$m_1 g + F_{l_1} - K y_1 = m_1 \ddot{y}_1$$
$$m_2 g + F + F_{l_2} = m_2 \ddot{y}_2$$

expected optimal soln: $l^* + \frac{(m_1+m_2)g}{K} = h$

$$\Rightarrow l^* = h - \frac{(m_1+m_2)g}{K}$$

$y - h \to 0$, $F \to 0$