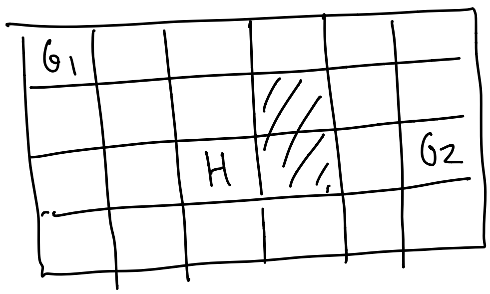

Infinite time horizon

$\gamma = 0.99$ Perfect actuator

Start H

Cost action $-c$ every state

"Sticky" goals

Reward $G_1 = 0.1$ Reward $G_2 = 0.2$

$\boxed{Q1}$ Does c affect opt policy?

A: No. $\sum_t \gamma^t r_t = \sum_t \gamma^t (r_t - c) = \sum_t \gamma^t r_t - \underbrace{\sum_t \gamma^t c}_{\text{constant}}$

[assuming incur cost c in goal states]

$\boxed{Q2}$ Does scaling rewards by $\alpha > 0$ affect opt policy?

A: No. $\sum_t \gamma^t r_t = \sum_t \gamma^t \alpha r_t = \alpha \sum_t \gamma^t r_t$

$\boxed{Q3}$ Does changing discount factor to $\gamma = 0.1$ affect opt policy?

A: Yes $\gamma^4 \frac{0.1}{1-\gamma} < \gamma^5 \frac{0.2}{1-\gamma}$ for $\gamma = 0.99$

$\gamma^4 \frac{0.1}{1-\gamma} > \gamma^5 \frac{0.2}{1-\gamma}$ for $\gamma = 0.1$

Q4 Suppose:

$$r(s, a, s') = r(s,a) + \gamma \Phi(s') - \Phi(s)$$

Does this change optimal policy?

A: No.

$$\sum_t \gamma^t \left(r_t + \gamma \Phi(s_{t+1}) - \Phi(s_t)\right)$$

telescopes,

$$\gamma^0 r_0 + \gamma \Phi(s_1) - \Phi(s_0)$$
$$+ \gamma r_1 + \gamma^2 \Phi(s_2) - \gamma \Phi(s_1)$$
$$+ \gamma^2 r_2 + \gamma^3 \Phi(s_3) - \gamma^2 \Phi(s_2)$$
$$\vdots$$

$$= r_0 + \gamma r_1 + \gamma^2 r_2 + \cdots$$
$$\underbrace{- \Phi(s_0)}_{\text{constant}}$$