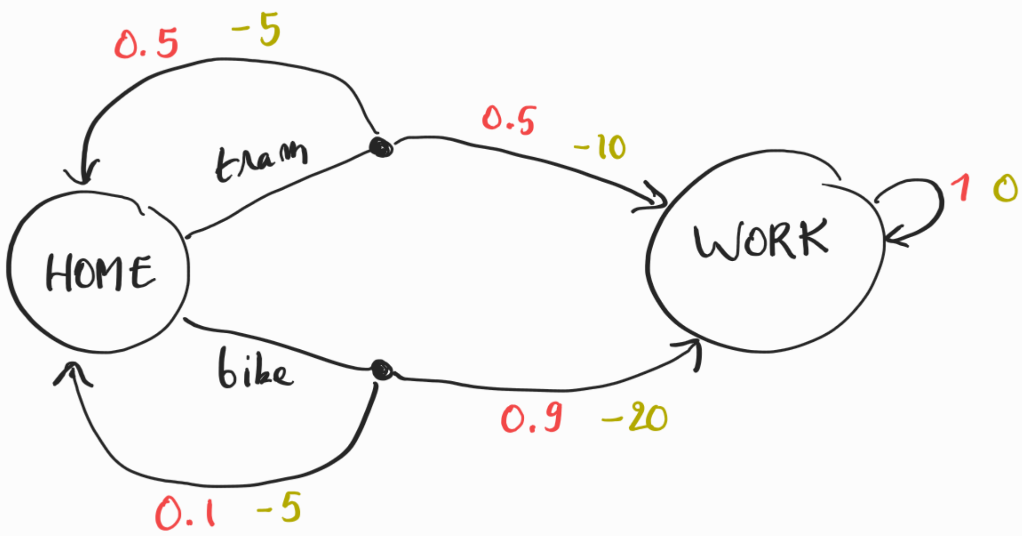
MARKOV DECISION PROCESSES

states : S actions : A

$$\Delta : S \times A \longrightarrow \mathcal{D}ist(S \times \mathbb{R})$$

$\Delta(s', r \mid s, a)$ probability to go to state s' with reward r if playing action a from state s



PATH $\equiv$ PLAY $\equiv$ TRAJECTORY

$s_0 \; a_0 \; r_0 \; s_1 \; a_1 \; r_1 \; s_2 \; a_2 \; r_2 \; \dots$

TOTAL (DISCOUNTED) PAYOFF

$r_0 + r_1 + r_2 + \dots$ may not be defined!

case 1: reachability: $r \in \{0,1\}$ 1 is a sink

case 2: discounted : $\gamma \in (0,1)$

$$\sum_{i \geq 0} \gamma^i r_i = r_0 + \gamma r_1 + \gamma^2 r_2 + \dots$$

STRATEGIES ≡ POLICIES

games $\nearrow$ STRATEGIES   RL $\nearrow$ POLICIES

$S^* \to \mathcal{D}(A)$ general

$S^* \to A$ pure

$S \to \mathcal{D}(A)$ positional

$\sigma : S \to A$ pure & positional

VALUE FUNCTION

$$v : S \to \mathbb{R}$$

$$val_\sigma(s) = \mathbb{E}_\sigma\left[\sum_{i \geq 0} \gamma^i r_i\right] \quad \text{on-policy}$$

$$val_*(s) = \sup_\sigma val_\sigma(s)$$

$$\sigma_* = \arg\max_\sigma val_\sigma$$ not well defined a priori!

Theorem:
there exists σ pure and positional which is optimal