HOMEWORK 4

ALEC PORTELLI

(1)

(A)

$$E(A) = (0.1 \cdot 20) + (.1 \cdot 0) = 2 + 0 = 2$$
$$E(B) = (0.4 \cdot 30) + (0.6 \cdot 0) = 12 + 0 = 12$$

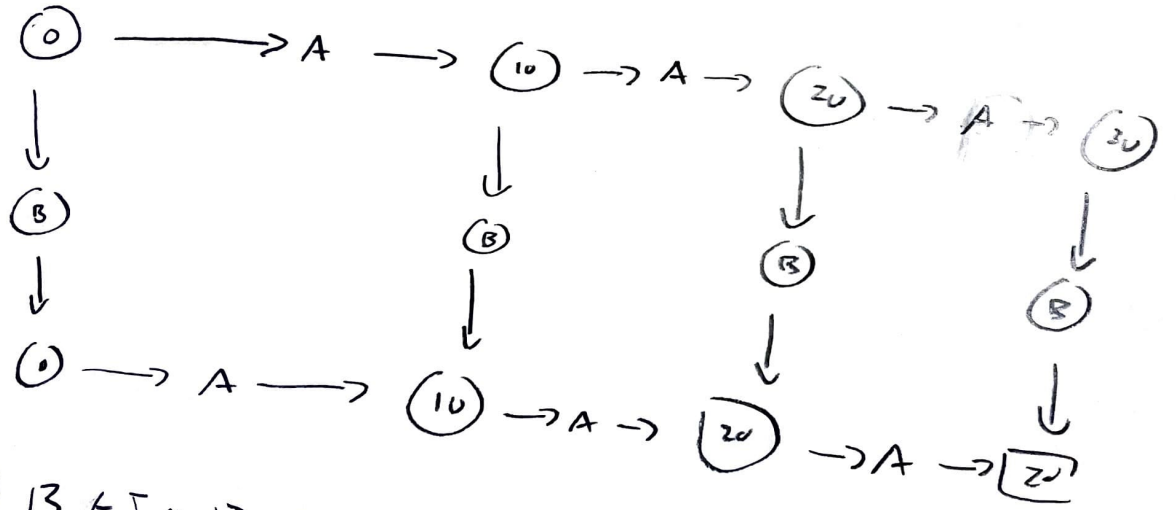
(B)

State space is all possible combinations of money in pocket

$\{0, 10, 20, 30\} = S$

Action space is either playing machine A or machine B

$\{A, B\} = A$



(C) $\beta \in [0, 1]$ If β is closer to 1, choose A, if not, choose B

$\pi(\beta) = V_{\pi\beta}$

$\pi(A|s) = \beta$ if $A = a$ and $S = 20$

$\pi(A|s) = 1$ if $a = B$ or $S \neq 20$

$V_{\pi\beta}(s) = \sum_{a \in A} \pi_\beta(a|s) \sum_{s', r} p(s', r|s, a)[r + \gamma V_{\pi_{13}}(s')]$

(D)

$E(A)$ = cost + chance of winning · reward = \$12

$E(B)$ = cost + probability of winning

Optimal policy: $E(B) \leq E(A)$ for all η, so it tends to:

$\$20 + \eta \cdot \$30 \leq 12$

Solve for η

$\eta \cdot \$30 \leq -\$8$, no value of η works for B, so always choose machine A