Algorithm 1: Value Iteration

Input: c, K, n, p, u, α
Output: policy convergeed J^*

Initialize $\theta = 0.001, \delta = \text{inf}$;
while $\delta > \theta$ **do**
  for $i = 0$ **to** n **do**

$$J_{K+1}(i) = min(k+\alpha(1-p)J_K(0)+\alpha pJ_K(1), ci+\alpha(1-p)J_k(i)+\alpha pJ_k(i+1)$$

$$= |J_{K+1}(i) - J_K(i)|$$

for $i = 0$ **to** n **do**

$$policy_i = argmin_{u \in U}(i + \alpha(1-p)J_K(i) + \alpha pJ_k(i+1))$$

Algorithm 2: Policy Iteration

Input: c, K, n, p, u, α, s
Output: policy converged J^*

Initialize $\theta = 0.001, \delta = \text{inf}, s = 0$;
while $\delta > \theta$ **do**
  for $i = 0$ **to** n **do**

$$value_{expected} = 0$$

$$value_{expercted} = \sum_{u \in U} p_u(K + \alpha \sum_{s \in S} J_k(i)$$

for $i = 0$ **to** n **do**

$$policy_i = argmin_{uU}(i + \alpha(1-p)J_k(i) + \alpha pJ_k(i+1))$$