$V(S) \leftarrow \sum_a \pi(a|s) \sum_{s',r} p(s',r|s,a)[r+\gamma V(s')]$

$\pi(a|s)$ dla kazdego stanu s i akcji a wynosi tyle samo czyli = 0.5

$p(s',r \mid s,a)$ dla kazdgo s' i r wynosi 1, poniewaz w naszym srodowisku kazde wykonanie akcji prowadzi do przejscia do kolejnego stanu s

$\gamma = 1$, poniewaz rozpatrujemy problem jako niezdyskontowany

Początkowe V(S) dla każdego stanu s wynosi 0.

Sl → stan terminalny po lewo
Sp → stan terminalny po prawo

$V(A) \leftarrow \sum_a \pi(a|s) \sum_{s',r} p(s',r|s,a)[r+\gamma V(s')] = 0.5 * 1 * [0+\gamma * V(Sl)]$

$V(A) = 0.5[V(Sl) + V(B)] + 0.5[1(0+\gamma * V(B))] = 0.5V(B)$

$V(B) = 0.5[V(A) + V(C)] = > 2V(B) = V(A) + V(C)$

$V(C)=0.5 [V(B)+V(D)] = > 2V(C) = V(B) + V(D)$

$V(D)=0.5 [V(C)+V(E)] = > 2V(D) = V(C) + V(E)$

$V(E)=0.5 [V(D)+V(Sp)] = > 2V(B) = V(D) + 1 + V(Sp) = V(D) + 1$

Stad:
$2V(B) = 0.5V(B) + V(C) => 1.5V(B) = V(C)$
$3V(A) = V(C)$
$4V(A) = V(D)$

$6V(A) = 1 => V(A) = 1/6$
$V(B) = 2/6$
$V(C) = 3/6$
$V(D) = 4/6$
$V(E) = 5/6$