CSSE 313: Worksheet for policy iteration

Consider the policy and associated utility:

→	→	→	+1
→	W	↑	-1
→	→	↓	↓

-0.056	0.562	0.825	+1
-0.120	W	0.451	-1
-0.120	-0.120	-0.120	-0.120

Now, determine the new policy for the state indicated. Here is the expression for determining a policy. Please show your work.

$$\text{if } \max_{a \in A(s)} \sum_{s'} P(s'|s,a) U[s'] > \sum_{s'} P(s'|s, \pi[s]) U[s']$$

$$\text{then } \pi[s] := \underset{a \in A(s)}{\text{argmax}} \sum_{s'} P(s'|s,a) U[s']$$