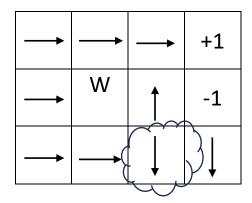
CSSE 313: Worksheet for policy iteration

Consider the policy and associated utility:



-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.

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

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