

Policy evaluation quizz

-150	160	-40
-150		-40
-150		-40

Numbers represent rewards. Red states are terminal.

Agent will follow the order with probability of 0.6 and with probability of 0.2 will go in $\pm 90^\circ$ direction. Discount factor is 0.8. For non-terminal states, find values $V(s)$ for policy: “ jeď na zapad (go west) ” . The reward for the nonterminal states is 0, i.e. $r(s) = 0$.