

Policy evaluation quizz

-150	170	-110
-150		-110
-150		-110

Numbers represent rewards. Red states are terminal.

Agent will follow the order with probability of 0.8 and with probability of 0.1 will go in $\pm 90^\circ$ direction. Discount factor is 0.7. For non-terminal states, find values $V(s)$ for policy: “jeď na sever (go north)”. The reward for the nonterminal states is 0, i.e. $r(s) = 0$.