

Policy evaluation quizz

-70	190	-120
-70		-120
-70		-120

Numbers represent rewards. Red states are terminal.

Agent will follow the order with probability of 0.9 and with probability of 0.05 will go in $\pm 90^\circ$ direction. Discount factor is 0.6. For non-terminal states, find values $V(s)$ for policy: “ jeđ na sever (go north) ” . The reward for the nonterminal states is 0, i.e. $r(s) = 0$.