

Policy evaluation quizz

-50	100	-20
-50		-20
-50		-20

Numbers represent rewards. Red states are terminal.

Agent will follow the order with probability of 0.65 and with probability of 0.175 will go in $\pm 90^\circ$ direction. Discount factor is 0.7. 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. $r(s) = 0$.