

Policy evaluation quizz

-30	160	-60
-30		-60
-30		-60

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 0.5. For non-terminal states, find values $V(s)$ for policy: “ jeđ na západ (go west) ”