

State value computation for random walk



Lets assume the five-state world shown above, two possible actions,  $\leftarrow$  and  $\rightarrow$ , and stochastic strategy where direction choice is random and the agent always follows the order. Both actions have the same probability. Exit from our world to left gives reward 0, exit to right gives reward 1. Evaluate the optimal strategy, i.e. compute the values the individual states. For simplicity assume  $\gamma = 1$ .