State values for a random walk
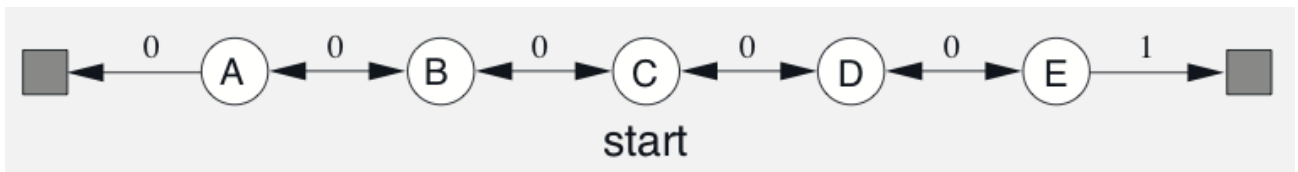

Consider the world with 5 states as above. There are two possible actions, ← and →, and a stochastic policy, in which the choice of action is completely random. The agent always follows the command. Both actions are thus equally likely. Exiting the world on the left gives a reward of 0; exiting on the right a reward of 1. Evaluate the policy, i.e. calculate the values for every state.