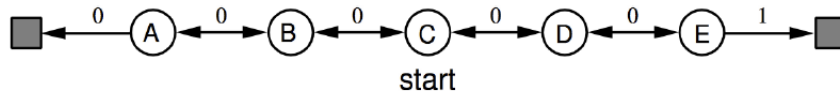


Wprowadzenie do RL 6

Zadanie 1

Dla spaceru przypadkowego (*random walk*) omówionego na wykładzie:



udowodnij, że: $V(A)=1/6$, $V(B)=2/6$, $V(C)=3/6$, $V(D)=4/6$, $V(E)=5/6$.

Wykorzystaj wskazówkę z wykładu.

Zadanie 2

Zapoznaj się z przykładową implementacją algorytmu TD(0) do problemu *random walk* znajdującą się w notatniku [RandomWalk.ipynb](#).

Wykonaj podane tam polecenie.

Zadanie 3

Napisz program implementujący poniższy **algorytm** dla środowiska *FrozenLake* wyliczenia polityki π .

Tabular TD(0) for estimating v_π

Input: the policy π to be evaluated

Algorithm parameter: step size $\alpha \in (0, 1]$

Initialize $V(s)$, for all $s \in \mathcal{S}^+$, arbitrarily except that $V(\text{terminal}) = 0$

Loop for each episode:

 Initialize S

 Loop for each step of episode:

$A \leftarrow$ action given by π for S

 Take action A , observe R, S'

$V(S) \leftarrow V(S) + \alpha[R + \gamma V(S') - V(S)]$

$S \leftarrow S'$

 until S is terminal

Rozwiązanie w formacie **py** (czysty Python) lub **ipynb** (notatnik Colab).