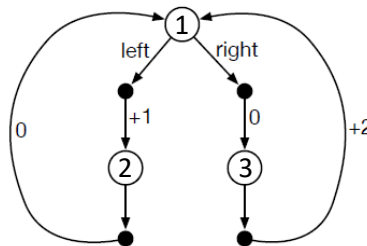


## Wprowadzenie do RL 4

### Zadanie 1

Rozważmy ciągły **MDP** przedstawiony na poniższym rysunku:



Rozważmy dwie deterministyczne polityki  $\pi_{\text{RIGHT}}$  i  $\pi_{\text{LEFT}}$  określające wybór w stanie 1. Jaka polityka jest optymalna dla:  $\gamma = 0$ ,  $\gamma = 0.9$ ,  $\gamma = 0.5$  ?

**UWAGA:** Przed przystąpieniem do zadań 2 i 3 zapoznaj się z zawartością pliku [COLAB\\_instrukcja.pdf](#).

### Zadanie 2

Zapoznaj się z notatnikiem [FrozenLake\\_1.ipynb](#) (otwórz go w Colab) i wykonaj **polecenia 1, 2 i 3** widoczne w spisie treści notatnika:

#### Table of contents

##### FrozenLake 1

Wprowadzenie

Dynamika

Polecenie 1 (do uzupełnienia)

Poruszanie i wizualizacja

Polecenie 2 (do uzupełnienia)

Ruch agenta w pętli

Polecenie 3 (do uzupełnienia)

### Zadanie 3

Zapoznaj się dokładnie z notatnikiem [FrozenLake\\_2.ipynb](#). Jeżeli to konieczne wróć do notatnika [FrozenLake\\_1.ipynb](#). **Jeżeli czegoś nie będziesz rozumiał – zapytaj w czasie zajęć!!!**

Wykonaj 2 polecenia z notatnika [FrozenLake\\_2.ipynb](#) oraz napisz program implementujący poniższy **algorytm iteracyjnego obliczenia polityki** dla środowiska *FrozenLake* w celu znalezienia wartości oczekiwanych zwrotu  $V(s)$  dla każdego stanu  $s$  przy zadanej polityce  $\pi$ .

### Iterative Policy Evaluation, for estimating $V \approx v_\pi$

Input  $\pi$ , the policy to be evaluated

Algorithm parameter: a small threshold  $\theta > 0$  determining accuracy of estimation

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

Loop:

$\Delta \leftarrow 0$

Loop for each  $s \in \mathcal{S}$ :

$v \leftarrow V(s)$

$V(s) \leftarrow \sum_a \pi(a|s) \sum_{s',r} p(s',r|s,a) [r + \gamma V(s')]$

$\Delta \leftarrow \max(\Delta, |v - V(s)|)$

until  $\Delta < \theta$

Przyjmij, że **polityka stochastyczna**  $\pi$  określa jednakowe prawdopodobieństwo dla każdej akcji:  
 $\forall a \pi(a|s) = 1/4$ . Przetestuj program dla różnych wartości parametru  $\gamma$  np. 1, 0.9, 0.5.

Wskazówki do programu znajdziesz w notatniku [FrozenLake\\_2.ipynb](#).