

Uczenie maszynowe: *ćwiczenia do wykładu 14*

Paweł Cichosz

	B		
		B	
			1

1. Dla środowiska siatki 4×4 z deterministycznymi efektami akcji prześledzić zmiany funkcji Q wykonane przez algorytm Q-learning przyjmując, że zadanie jest epizodyczne i epizod kończy wejście do komórki w prawym dolnym rogu, przy czym wówczas przyznawana jest nagroda 1, a we wszystkich pozostałych sytuacjach nagroda 0, oraz że początkowe wartości Q wynoszą 0 i stosowany jest rozmiar kroku $\beta = 0.1$, w trakcie trzech następujących po sobie epizodów:
epizod 1: początek w komórce w lewym dolnym rogu, ciąg akcji $\uparrow, \rightarrow, \rightarrow, \uparrow, \downarrow, \rightarrow, \uparrow, \rightarrow$,
epizod 2: początek w w komórce w lewym dolnym rogu, ciąg akcji $\rightarrow, \rightarrow, \uparrow, \rightarrow$,
epizod 3: początek w w komórce w lewym dolnym rogu, ciąg akcji $\rightarrow, \rightarrow, \uparrow, \rightarrow$.
2. Przeprowadzić symulację identyczną jak w poprzednim ćwiczeniu dla algorytmu Sarsa.