

Uczenie maszynowe: *ćwiczenia do wykładu 13*

Paweł Cichosz

1. Wyznaczyć wartości funkcji V dla stanów w ostatniej kolumnie przykładowego środowiska siatki 4×4 względem strategii, która zawsze wybiera akcje prowadzące w kierunku komórki z wypłatą nagrody, zakładając, że efekty akcji są deterministyczne.
2. Wyznaczyć wartości funkcji Q dla stanów w ostatniej kolumnie przykładowego środowiska siatki 4×4 oraz akcji \uparrow i \downarrow względem strategii, która zawsze wybiera akcje prowadzące w kierunku komórki z wypłatą nagrody, zakładając, że efekty akcji są deterministyczne.
3. Wyznaczyć wartość funkcji V^* dla stanu x_0 przykładowego środowiska „ $x_0 \rightarrow \dots \rightarrow x_n$ ”, zakładając, że efekty akcji są niedeterministyczne i prawdopodobieństwo przejścia w wybranym kierunku wynosi 0.8, a prawdopodobieństwo pozostania w dotychczasowym stanie wynosi 0.2.
4. Wyznaczyć wartość funkcji V^* dla stanu x_0 przykładowego środowiska „ $x_0 \rightarrow \dots \rightarrow x_n$ ”, zakładając, że efekty akcji są niedeterministyczne i prawdopodobieństwo przejścia w wybranym kierunku wynosi 0.8, a prawdopodobieństwo przejścia w przeciwnym do wybranego kierunku wynosi 0.2.
5. Wyznaczyć wartości funkcji V dla stanów w ostatniej kolumnie przykładowego środowiska siatki 4×4 względem strategii, która zawsze wybiera akcje prowadzące w kierunku komórki z wypłatą nagrody, zakładając, że efekty akcji są niedeterministyczne i prawdopodobieństwo przejścia w wybranym kierunku wynosi 0.8, a prawdopodobieństwo pozostania w dotychczasowym stanie wynosi 0.2.
6. Wyznaczyć wartości funkcji Q dla stanów w ostatniej kolumnie przykładowego środowiska siatki 4×4 oraz akcji \uparrow i \downarrow względem strategii, która zawsze wybiera akcje prowadzące w kierunku komórki z wypłatą nagrody, zakładając, że efekty akcji są niedeterministyczne i prawdopodobieństwo przejścia w wybranym kierunku wynosi 0.8, a prawdopodobieństwo pozostania w dotychczasowym stanie wynosi 0.2.