

Spis treści

I	Preliminaria	9
1.	Wprowadzenie	11
1.1.	Przykłady zagadnień, w których pojawia się potrzeba adaptacji i uczenia się	13
1.2.	Cel podręcznika i poruszane zagadnienia	14
1.3.	Dziedziny składające się na zawartość podręcznika	15
1.4.	Adaptacja i uczenie się jako droga do inteligentnego zachowania się maszyn i programów	16
1.5.	Organizacja podręcznika	18
1.6.	Potrzebna wiedza wstępna	19
2.	Optymalizacja stochastyczna	21
2.1.	Algorytm gradientu prostego	21
2.2.	Procedura Stochastycznego Najszybszego Spadku, SGD	23
3.	Aproksymacja funkcji i podstawowe mechanizmy adaptacji	31
3.1.	Aproksymatory i zagadnienie aproksymacji	31
3.2.	Zagadnienie uczenia na zbiorze nieskończonym, on-line	35
3.3.	Perceptron wielowarstwowy	37
3.4.	Uczenie się przy użyciu aproksymacji stochastycznej	43
3.5.	Zagadnienia praktyczne związane z używaniem sieci neuronowych w systemach uczących się	46
4.	Proces Decyzyjny Markowa i programowanie dynamiczne	49
4.1.	Proces Decyzyjny Markowa	49
4.2.	Funkcja wartości-akcji i indukowanie strategii	52
4.3.	Klasyczny algorytm programowania dynamicznego	53
4.4.	Programowanie dynamiczne w algorytmice	56
4.5.	Algorytm Iteracji Strategii	58
4.6.	Algorytm Iteracji Funkcji Wartości	59
II	Uczenie się ze wzmocnieniem	61
5.	Podstawy	63
5.1.	Nieznany z góry Proces Decyzyjny Markowa	63
5.2.	Algorytmy <i>Q-Learning</i> i <i>SARSA</i>	66

5.3.	Rozszerzenie algorytmów <i>Q-Learning</i> i <i>SARSA</i> do ciągłych przestrzeni stanów i akcji	68
6.	Optymalizacja stochastycznego wyboru	75
6.1.	Parametryzowane rozkłady prawdopodobieństwa	76
6.2.	Algorytm REINFORCE punktowy	83
6.3.	Stacjonarna strategia decyzyjna	85
6.4.	Algorytm REINFORCE statyczny	87
6.5.	Algorytm REINFORCE epizodyczny	89
7.	Algorytm Aktor-Krytyk	95
7.1.	Idea algorytmu Aktor-Krytyk	95
7.2.	Klasyczny Aktor-Krytyk	96
7.3.	Aktor-Krytyk(λ)	99
8.	Aktor-Krytyk z kompatybilną aproksymacją	105
8.1.	Optymalizacja średniej nagrody	105
8.2.	Gradient strategii	107
8.3.	Aktor-Krytyk z kompatybilną aproksymacją	110
8.4.	Naturalny Aktor-Krytyk	112
8.5.	Dyskonto — ograniczenie wariancji estymatora gradientu	115
9.	Wielokrotne przetwarzanie obserwacji	117
9.1.	Algorytm <i>Q-Learning</i> z powtarzaniem doświadczenia	119
9.2.	Próbkowanie istotnościowe	120
9.3.	Algorytm Aktor-Krytyk z powtarzaniem doświadczenia	122
9.4.	Optymalizacja estymatora wskaźnika jakości	131
10.	Algorytmy wzbogacone	133
10.1.	<i>Asynchronous Advantage Actor-Critic</i>	133
10.2.	<i>Proximal Policy Optimization</i>	134
10.3.	<i>Deep Deterministic Policy Gradient</i>	137
10.4.	<i>Soft Actor-Critic</i>	138
11.	Gęsta dyskretyzacja czasu	143
11.1.	Strategia z autokorelowanymi akcjami	144
11.2.	<i>Actor-Critic with Experience Replay and Autocorrelated aCtions</i>	147
12.	Uczenie się w warunkach częściowo obserwowalnego stanu	151
12.1.	Rekurencyjne sieci neuronowe	152
12.2.	<i>Deep Recurrent Q-Learning</i>	153
13.	Wieloagentowe uczenie się ze wzmocnieniem	155
13.1.	Model synchroniczny niekooperacyjny z częściowo obserwowanym stanem	156
13.2.	<i>Multi-Agent Deep Deterministic Policy Gradient</i>	157

III Sterowanie adaptacyjne	161
14. Obiekty dynamiczne	163
14.1. Wstęp	163
14.2. Liniowe obiekty SISO	165
14.3. Dyskretna aproksymacja obiektów o ciągłej dynamice	175
15. Stabilność i funkcja Lapunowa	179
15.1. Ogólna postać typowego schematu adaptacji	179
15.2. Stabilność	181
15.3. Funkcja Lapunowa	184
15.4. Stabilność w kontekście funkcji Lapunowa	185
16. Sterowanie adaptacyjne z modelem referencyjnym	191
16.1. Liniowy obiekt SISO pierwszego rzędu	191
16.2. Uogólnienie	195
16.3. Obiekty liniowe wyższych rzędów	198
17. Zaawansowane schematy adaptacji	205
17.1. Obiekty o nieliniowej dynamice	205
17.2. Obiekty z nieobserwowalnymi pochodnymi stanu	208
18. Samostrojące się regulatory	215
18.1. Dynamika liniowo parametryzowalna	216
18.2. Liniowe najmniejsze kwadraty	218
18.3. Najmniejsze kwadraty z wykładniczym zapominaniem	220
18.4. Adaptacyjny dobór współczynnika zapominania	223
IV Inne podejścia do adaptacji	225
19. Aproksymowane programowanie dynamiczne	227
20. Stochastyczne sterowanie adaptacyjne	231
21. Sterowanie z iteracyjnym uczeniem się	233
22. Filtr Kalmana	237
22.1. Model	237
22.2. Algorytm	239
22.3. Wyprowadzenia	241
22.4. Rozszerzony Filtr Kalmana	243
Literatura	245