

Dipl.-Inform. Martin Riedmiller, Karlsruhe

Selbständig lernende neuronale Steuerungen

Reihe **8**: Meß-, Steuerungs-
und Regelungstechnik

Nr. **626**

Inhaltsverzeichnis

1	Einführung	1
1.1	Motivation	1
1.2	Neuronal lernende Steuerungen	3
1.2.1	Die Idee selbständig lernender Steuerungen	3
1.3	Zielsetzung und Aufbau der Arbeit	5
1.3.1	Zielsetzung und Abgrenzung	5
1.3.2	Vorgehensweise und Aufbau der Arbeit	6
2	Lernen in neuronalen Netzen	9
2.1	Das mehrschichtige neuronale Netz	9
2.2	Überwachtes Lernen	11
2.3	Der <i>Backpropagation</i> -Algorithmus	12
2.4	Lernen mit adaptiven Schrittweiten: <i>Rprop</i>	13
2.5	Zusammenfassung	15
3	Grundzüge der Regelungstechnik	16
3.1	Einführung	17
3.1.1	Betrachtete Problemstellung	17
3.1.2	Begriffe	17
3.1.3	Ziele der Regelung	18
3.2	Reglerentwurf im Frequenzbereich	19
3.2.1	Lineare Übertragungsglieder	19
3.2.2	Die Laplace-Transformation	20
3.2.3	Beschreibung linearer Systeme im Frequenzbereich	21

3.2.4	Entwurfsverfahren im Frequenzbereich	21
3.3	Reglerentwurf im Zeitbereich	22
3.3.1	Beschreibung im Zustandsraum	22
3.3.2	Reglerentwurf durch Polvorgabe	24
3.3.3	Reglerentwurf durch Optimierung eines Gütekriteriums	27
3.3.4	Entwurf des Vorfilters	27
3.3.5	Der Beobachter	29
3.4	Zeitdiskrete Modellierung	30
3.4.1	Die zeitdiskrete Zustandsgleichung	30
3.4.2	Alternative Zustandsbeschreibung	31
3.5	Zusammenfassung	33
4	Neuronale Steuerungen	35
4.1	Die Rolle des a priori Wissens	35
4.2	Lernen aus Beispielen	37
4.2.1	Imitation eines Experten	37
4.2.2	Neuronale Systemidentifikation	38
4.3	Lernen mit Referenzmodell	39
4.3.1	Beschreibung des Verfahrens	39
4.3.2	Anwendungsbeispiel: Drehmomentregelung eines Ottomotors	40
4.4	Selbständig lernende Steuerungen	43
4.4.1	Der Zeitpunkt der Beurteilung	44
4.4.2	Lernen mit zeitlich verzögerter Beurteilung	44
4.4.3	'Belohnen und Bestrafen' - Der <i>ASE/ACE</i> -Ansatz	45
4.4.4	Weiterführende Arbeiten	46
4.4.5	Reinforcement Learning und dynamisches Programmieren	48
4.4.6	Überblick und Einordnung dieser Arbeit	49
4.5	Zusammenfassung	50

5	Dynamisches Programmieren	51
5.1	Betrachtete Problemstellung	51
5.2	Optimierungsprobleme mit endlichem Horizont	52
5.2.1	Aufgabenbeschreibung	52
5.2.2	Das Bellman'sche Optimalitätsprinzip	53
5.2.3	Backward Dynamic Programming	54
5.3	Optimierungsprobleme mit unendlichem Horizont	55
5.3.1	Kürzester Pfad (shortest path)-Probleme	55
5.3.2	Das Value Iteration-Verfahren	56
5.3.3	Konvergenz des Value Iteration-Verfahrens	57
5.3.4	Reihenfolge der Anwendung	58
5.4	Zusammenhang zu lernenden Systemen	58
5.4.1	Real Time Dynamic Programming (RTDP)	59
5.4.2	Q-Learning	60
5.5	Zusammenfassung	61
6	Eine selbständig lernende neuronale Steuerung	63
6.1	Formale Aufgabenbeschreibung	64
6.2	Basisalgorithmus	65
6.2.1	Die Prozedur Aktionswahl	66
6.2.2	Die Prozedur Anwendung	67
6.2.3	Die Prozedur Anpassung	67
6.2.4	Zusammenfassung	67
6.3	Der ereignisbasierte Ansatz	68
6.3.1	Modellierung der Steuerungsaufgabe als kürzester Pfad-Problem	68
6.3.2	Der ereignisbasierte Algorithmus	70
6.3.3	Zusammenfassung	71
6.4	Der zeitbasierte Ansatz	72
6.4.1	Nachweis der Konvergenz	73
6.4.2	Terminierung der inneren Schleife	76
6.4.3	Der zeitbasierte Algorithmus	77
6.4.4	Zusammenfassung	78

6.5	Das Systemmodell	79
6.5.1	Der Satz vom falschen Modell	81
6.5.2	Praktische Bedeutung	84
6.6	Das Lernverfahren	85
6.6.1	Lernen der neuronalen Kostenfunktion	86
6.6.2	Die Familie der TD(λ)-Lernverfahren	88
6.7	Architektur einer neuronalen Steuerung	91
6.7.1	Das Modell	92
6.7.2	Die neuronale Bewertungsfunktion	92
6.7.3	Das Auswahlmodul	93
6.8	Zusammenfassung	93
7	Lernverhalten und Leistungsfähigkeit	95
7.1	Steuerung eines dynamischen Systems 2. Ordnung	95
7.1.1	Analytisch optimaler linearer Reglerentwurf	97
7.1.2	Neuronaler Reglerentwurf	97
7.2	Untersuchung der ereignisbasierten Methode	99
7.2.1	Referenzparameter	99
7.2.2	Lernverhalten	100
7.2.3	Entwicklung der Regelgüte	104
7.2.4	Abhängigkeit von der Gewichtsinitialisierung	106
7.2.5	Abhängigkeit von der Netzgröße	106
7.2.6	Abhängigkeit von der Lernrate	108
7.2.7	Abhängigkeit vom Explorationsfaktor	110
7.2.8	Die Wahl des Zielbereichs \mathcal{X}^+	111
7.3	Lernverhalten der zeitbasierten Methode	113
7.3.1	Referenzparameter	113
7.3.2	Lernverhalten	114
7.3.3	Entwicklung der Regelgüte	115
7.3.4	Abhängigkeit von Gewichtsinitialisierung und Netzgröße	117
7.3.5	Abhängigkeit von der Lernrate	117
7.3.6	Abhängigkeit vom Explorationsfaktor	118

7.3.7	Zusammenfassung ereignisbasierter und zeitbasierter Ansatz	118
7.4	Praktische Aspekte	120
7.4.1	Ein generisches Prozeßmodell	120
7.4.2	Eigenschaften des Prozeßmodells	122
7.4.3	Lernen mit falschem Modell beim P-T ₂ -System	125
7.4.4	Repräsentation des Systemzustands	126
7.4.5	Nichtlineare Eingangskennlinie	128
7.5	Zusammenfassung	129
8	Verbesserung des Lernverhaltens	131
8.1	Prinzipielle Probleme beim Lernen	131
8.2	Die Methode der wachsenden Kompetenz	132
8.2.1	Prinzip	132
8.2.2	Ergebnisse	133
8.3	Die Methode der schrittweisen Verfeinerung	134
8.3.1	Prinzip	134
8.3.2	Ergebnisse	135
8.4	Mehr-Experten Steuerung	137
8.4.1	Prinzip	137
8.4.2	Mehr-Regler Architektur	137
8.4.3	Mehr-Netz Architektur	138
8.4.4	Training der Mehr-Netz Architektur	139
8.4.5	Grob- und Feinbereichsregelung mit variabler Aktionsmenge	140
8.5	Zusammenfassung	141
9	Strukturvarianten	142
9.1	Erzeugung variabler Stellgrößen	142
9.1.1	Nachträgliche Erweiterung der Aktionsmenge	143
9.1.2	Das Konzept der dynamischen Ausgabeelemente (<i>DAE</i>)	144
9.1.3	Verhalten der <i>DAE</i> -Architektur	146
9.1.4	Proportional verzögerndes <i>DAE</i> (P-T _I - <i>DAE</i>)	148
9.1.5	Integrierendes <i>DAE</i> (I- <i>DAE</i>)	148

9.1.6	Verzögernd-Integrierendes DAE (I-T ₁ -DAE)	149
9.2	Realisierung einer Folgeregelung	150
9.2.1	Idee des neuronalen Vorfilters	150
9.2.2	Folgeregelung des P-T ₂ -Systems	152
9.3	Zusammenfassung	153
10	Anwendungen	155
10.1	Steuerung eines Roboterarms	155
10.1.1	Aufgabenstellung	155
10.1.2	Neuronale Steuerung	156
10.2	Regelung eines inversen Pendels	159
10.2.1	Aufgabenstellung	159
10.2.2	Reinforcement Learning: Der AHC/ASE-Regler	160
10.2.3	Analytischer Reglerentwurf durch Polvorgabe	161
10.2.4	Neuronale Steuerung	162
10.3	Regelung eines zweigliedrigen inversen Pendels	166
10.3.1	Aufgabenstellung	166
10.3.2	Neuronale Steuerung	167
10.3.3	Zusammenfassung	168
10.4	Regelung eines Rührkesselreaktors	169
10.4.1	Aufgabenstellung	169
10.4.2	Analytischer Reglerentwurf	170
10.4.3	Neuronale Steuerung	171
10.5	Steuerung des mobilen Roboters Khepera	174
10.5.1	Aufgabenstellung	174
10.5.2	Neuronale Steuerung	176
10.6	Regelung eines realen Stabbalancierers	178
10.6.1	Aufgabenstellung	179
10.6.2	Neuronale Steuerung	179
10.7	Zusammenfassung	180
11	Zusammenfassung und Ausblick	182
11.1	Überblick	182
11.2	Vorgehensweise beim Entwurf	183
11.3	Kritische Würdigung und Ausblick	186