

1	Einleitung	1
1.1	Motivation und Problemstellung	2
1.2	Zielstellung der Arbeit und Forschungsfragen	8
1.3	Forschungsmethodik und Aufbau der Arbeit	9
2	Grundlagen der Produktionsablaufplanung	13
2.1	Begriffsbestimmung und thematische Abgrenzung	13
2.2	Prozess der Produktionsablaufplanung	16
2.3	Mathematische Optimierung der Produktionsablaufplanung	19
2.3.1	Mathematische Formalisierung	19
2.3.1.1	Probleme	20
2.3.1.2	Nebenbedingungen	22
2.3.1.3	Zielfunktionen	24
2.3.2	Modellbildung	26
2.3.2.1	Ganzzahlige und gemischt-ganzzahlige lineare Programmierung	27
2.3.2.2	Ereignisdiskrete Simulation	29
2.3.3	Konventionelle Lösungsverfahren	31
2.3.3.1	Ganzzahlige lineare Optimierung	31
2.3.3.2	Heuristiken	35
2.3.3.3	Metaheuristiken	37
3	Grundlagen des Bestärkenden Lernens	41
3.1	Einordnung in die künstliche Intelligenz und in das maschinelle Lernen	42
3.1.1	Überwachtes Lernen als angrenzendes Paradigma	43

3.1.2	Unüberwachten Lernens als angrenzendes Paradigma	44
3.2	Grundprinzip und Taxonomie des bestärkenden Lernens	46
3.3	Gradientenabhängiges bestärkendes Lernen	48
3.3.1	Markov-Entscheidungsproblem	49
3.3.2	Nutzenfunktion	50
3.3.3	Aktionsnutzen-bewertende Verfahren	53
3.3.4	Entscheidungspolitik-approximierende Verfahren	57
3.3.5	Actor-Critic-Verfahren	61
3.4	Gradientenfreies bestärkendes Lernen	64
3.4.1	Modellsuchende und parameteroptimierende Verfahren	66
3.4.2	Hybride Verfahren – NeuroEvolution of Augmenting Topologies	68
4	Stand der Wissenschaft und Technik: Bestärkendes Lernen in der Produktionsablaufplanung	77
4.1	Gradientenabhängige Verfahren für die Produktionsablaufplanung	77
4.1.1	Agentenbasierte Auswahl von Prioritätsregeln	78
4.1.2	Agentenbasierte Ressourcenbelegungsplanung	81
4.1.3	Agentenbasierte Reihenfolgeplanung	83
4.1.4	Agentenbasierte Losbildung	87
4.1.5	Agentenbasiertes Reparieren von ungültigen Ablaufplänen	88
4.2	Gradientenfreie Verfahren für die Ablaufplanung im Allgemeinen	90
4.2.1	Einsatz der Kreuzentropie-Methode in der Ablaufplanung	91
4.2.2	Einsatz von Bayes'scher Optimierung in der Ablaufplanung	93
4.2.3	Einsatz von Neuro-Evolution in der Ablaufplanung	96
4.3	Zusammenfassung und Diskussion der Forschungslücke	100
5	Eine Methode zum Einsatz von bestärkenden Lernverfahren für die Produktionsablaufplanung	105
5.1	Ausgangssituation, Problemstellung und Anforderungsdefinition	106

5.2	Von der Produktionsablaufplanung zur agentenbasierten Produktionsablaufsteuerung – Prozessmodell und Funktionsprinzip	111
5.2.1	Agentenbasierte Ressourcenbelegungsplanung	115
5.2.2	Agentenbasierte Reihenfolgeplanung und Losbildung	118
5.3	Projektierung und Entwicklung von agentenbasierten Produktionsablaufsteuerungen	123
5.3.1	Entwurf von Agentenumgebungen	124
5.3.1.1	Modellierung der Erzeugung von Aufträgen	126
5.3.1.2	Modellierung von Aufträgen	127
5.3.1.3	Modellierung von Produktionsressourcen	132
5.3.2	Definition von maschinellen Lernaufgaben und Gestaltung von Agenten	138
5.3.2.1	Formulierung von Allokationsproblemen für die Ressourcenbelegungsplanung als maschinelle Lernaufgabe	139
5.3.2.2	Formulierung von Sequenzierungsproblemen für die Reihenfolgeplanung und Losbildung als maschinelle Lernaufgabe	142
5.3.3	Integration und Inbetriebnahme von Agenten und Agentenumgebungen	148
5.3.3.1	Integrationskonzept für gradientenabhängiges bestärkendes Lernen	149
5.3.3.2	Integrationskonzept für gradientenfreies bestärkendes Lernen	159
5.3.4	Auswahl und Implementierung von bestärkenden Lernverfahren	161
5.3.5	Gestaltung von Belohnungsfunktionen	171
5.3.5.1	Zeitbasierte Belohnungsfunktionen	172
5.3.5.2	Belastungsorientierte Belohnungsfunktionen	180
5.3.6	Training von Agenten	182
5.4	Zusammenfassung der Methode	193

6	Evaluation der entwickelten Methode	197
6.1	Flexible-Job-Shop-Problem mit flexibler Operationsplanung	198
6.1.1	Problembeschreibung	199
6.1.2	Anwendung des DQN-Algorithmus zur Lösung des Problems	200
6.1.3	Diskussion der Ergebnisse	206
6.1.4	Erweiterung des Problems um einen dynamischen Auftragshorizont	209
6.2	Dynamisches Parallel-Maschinen-Problem mit familienabhängigen Rüstzeiten und ressourcenabhängigen Bearbeitungsgeschwindigkeiten	212
6.2.1	Problembeschreibung	212
6.2.2	Anwendung des PPO-Algorithmus zur Lösung des Problems	214
6.2.3	Diskussion der Ergebnisse	220
6.3	Zweistufiges Hybrid-Flow-Shop-Problem mit familienabhängigen Rüstzeiten	223
6.3.1	Problembeschreibung	224
6.3.2	Anwendung des A2C-Algorithmus zur Lösung des Problems	226
6.3.3	Anwendung des NEAT-Algorithmus zur Lösung des Problems	238
6.3.4	Vergleich mit anderen Lösungsverfahren	251
7	Schlussbetrachtung	259
7.1	Zusammenfassung und Diskussion	259
7.2	Ausblick	264
	Literaturverzeichnis	267