

Masterseminar zur Betriebswirtschaftslehre

„Methoden, Anwendungen und Softwaresysteme
für Markovsche Entscheidungsprobleme“

im Sommersemester 2023

Themen und Termine

Markovsche Entscheidungsprobleme modellieren Entscheidungsprozesse, bei denen im Zeitablauf Entscheidungen unter Unsicherheit über die Wahl von Aktionen zu treffen sind und die Übergangswahrscheinlichkeiten von einem Zustand s des betrachteten Systems zu einem Folgezustand s' lediglich von s und s' sowie der in Zustand s gewählten Aktion a abhängen. Die Ausführungen von Aktionen a induzieren dabei jeweils Zahlungen $r(s, a, s')$. Die Lösung eines Markovschen Entscheidungsproblems besteht aus einer Politik π , die für jede Entscheidungsstufe t eine zugehörige Entscheidungsregel π_t definiert, mit der in Abhängigkeit der bisher durchlaufenen Zustände und Aktionen $H_t = (s_1, a_1, s_2, a_2, \dots, s_t)$ die Wahlwahrscheinlichkeiten der Aktionen a_t festgelegt werden. Je nach betrachteter Problemstellung wird eine Politik π gesucht, die den erwarteten Kapitalwert der Zahlungen (Gesamtgewinnkriterium) oder die Zahlung pro Entscheidungsstufe t (Durchschnittsgewinnkriterium) maximiert. Weitere Unterscheidungen im Hinblick auf die Arten Markovscher Entscheidungsprobleme beziehen sich auf die Anzahlen der Entscheidungsstufen (endlich, abzählbar unendlich oder überabzählbar) und der Zustände sowie Aktionen (endlich oder überabzählbar). Unter Voraussetzungen, die für viele Anwendungsfälle Markovscher Entscheidungsprobleme erfüllt sind, existiert stets eine optimale stationäre, deterministische und Markovsche Politik, d. h. eine Politik mit einer einzigen Entscheidungsregel π , die unabhängig von der Stufe t und von der Vorgeschichte $H_t \setminus \{s_t\}$ jedem Zustand s eine eindeutige Aktion $a = \pi(s)$ zuordnet.

Aufgrund ihrer Allgemeinheit besitzen Markovsche Entscheidungsprobleme zahlreiche Anwendungen in den Wirtschaftswissenschaften, den Ingenieurwissenschaften und der Informatik. Nachdem bis in die 1990er Jahre der Schwerpunkt der Forschung auf Ansätzen zur exakten Lösung Markovscher Entscheidungsprobleme lag, werden in jüngerer Zeit intensiv Methoden des maschinellen Lernens untersucht, die auch größere Instanzen innerhalb vertretbarer Rechenzeiten näherungsweise lösen können. Dabei haben sich insbesondere simulationsbasierte Verfahren bewährt, die auf den Verfahren der Politik- oder Wertiteration beruhen und Politiken bzw. Wertfunktionen als künstliche neuronale Netze oder Regressionsfunktionen repräsentieren.

Gegenstand des Seminars sind neben den grundlegenden exakten Ansätzen und maschinellen Lernverfahren auch ausgewählte betriebswirtschaftliche Anwendungen der Markovschen Entscheidungsprobleme sowie Softwarepakete, die Methodenbibliotheken zur Verfügung stellen. Die folgende Tabelle listet die Termine, Themen und Gruppengrößen der einzelnen Seminarvorträge auf. Die Vergabe der Themen findet bei der Vorbesprechung zum Seminar am Dienstag, den 7. Februar 2022 ab 15:15 Uhr im Seminarraum II des Instituts für Wirtschaftswissenschaft (Gebäude D4) statt. Die Anmeldung zum Seminar ist verbindlich. Mit der Vergabe des Themas wird Literatur ausgegeben, die die Grundlage für den Vortrag und die Ausarbeitung bildet. Die ausgegebene Literatur dient als Ausgangspunkt zur selbständigen Recherche und Auswertung weiterführender Literatur.

Nr.	Termin	Thema	# P ^a
1	14.04.	Markovsche Entscheidungsprobleme	1
2		Exakte Lösungsansätze	2
3	21.04.	Approximative Politik-Iteration	1
4		Anwendung: Energiespeicher	1
5	28.04.	Reinforcement Learning	2
6		Anwendung: Revenue Management	1
7	12.05.	Deep Reinforcement Learning	1
8		Anwendung: Projektmanagement	1
9	19.05.	Uniformization-Ansatz für zeitstetige Probleme	1
10		Anwendung: Reihenfertigungssysteme	1
11	26.05.	Policy-Gradient-Methoden für überabzählbare Akti- ons- und Zustandsräume	1
12		Anwendung: Instandhaltung	1
13	09.06.	Softwaresysteme	2

^aAnzahl Personen

Neben einem Vortrag (ca. 25 Minuten pro Person) erstellt jede Gruppe eine Seminararbeit (ca. 15 Seiten pro Person, einzeliger Zeilenabstand, Schriftgröße 12 pt). Im Inhaltsverzeichnis der Ausarbeitung ist für jedes Kapitel der jeweilige Verfasser anzugeben. Die Seminararbeiten müssen spätestens eine Woche und die Vortragsfolien spätestens 24 Stunden vor dem Vortrag als pdf-Dokumente an nora.krippendorff@tu-clausthal.de geschickt werden.

Die Seminarvorträge finden jeweils freitags in der Zeit von 10:15 bis 11:45 Uhr im Seminarraum II statt.

Literatur

- 1a Puterman ML (2005) Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley, Hoboken, NJ, Chapter 2
- 1b Powell WB (2007) Approximate Dynamic Programming. Solving the Curses of Dimensionality. John Wiley, Hoboken, NJ, Sections 3.2, 3.3
- 1c Waldmann K-H, Stocker, UM (2013) Stochastische Modelle: Eine anwendungsorientierte Einführung. Springer, Berlin, Abschnitt 6.1
- 2a Waldmann K-H, Stocker, UM (2013) Stochastische Modelle: Eine anwendungsorientierte Einführung. Springer, Berlin, Abschnitte 6.2–6.4
- 2b Brandimarte P (2021) From Shortest Path to Reinforcement Learning. Springer, Cham, Sections 4.4, 4.5
- 2c Powell WB (2007) Approximate Dynamic Programming. Solving the Curses of Dimensionality. John Wiley, Hoboken, NJ, Sections 3.4–3.7
- 3 Powell WB (2007) Approximate Dynamic Programming. Solving the Curses of Dimensionality. John Wiley, Hoboken, NJ, Section 8.4
- 4 Gönsch J, Hassler, M (2016): Sell or store? An ADP approach to marketing renewable energy. OR Spectrum 38:633–660
- 5a Mitchell TM (1997) Machine Learning. McGraw-Hill Series in Computer Science. McGraw-Hill Education, Boston, Chapter 13
- 5b Sutton RS, Barto AG (2018): Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, Chapter 6
- 5c Mahadevan S (1996) Average reward reinforcement learning: Foundations, algorithms, and empirical results. Machine Learning 22:159–195

- 5d Brandimarte P (2021) From Shortest Path to Reinforcement Learning. Springer, Cham, Chapter 5
- 6 Bondoux N, Nguyen AQ, Fiig T, Acuna-Agost R (2020) Reinforcement learning applied to airline revenue management. *Journal of Revenue and Pricing Management* 19:332–348
- 7a Winder P (2020) Reinforcement Learning: Industrial Applications of Intelligent Agents. O’Reilly Media, Sebastopol, CA, Chapter 4
- 7b Graesser L, Keng WK (2020) Foundations of Deep Reinforcement Learning: Theory and Practice in Python. Addison Wesley, Boston, Chapters 4, 5
- 7c Zai A, Brown B (2020) Einstieg in Deep Reinforcement Learning. KI-Agenten mit Python und PyTorch programmieren. Hanser, München, Kapitel 3
- 8 Sung I, Choi B, Nielsen P (2020) Reinforcement Learning for resource constrained project scheduling problem with activity iterations and crashing. *IFAC PapersOnLine* 53-2:10493–10497
- 9a McMahon JJ (2008) Time Dependence in Markovian Decision Processes. Doctoral Dissertation, Faculty of Mathematical and Computer Science, University of Adelaide, Section 2.3.4
- 9b Puterman ML (2005) Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley, Hoboken, NJ, Section 11.5
- 10 Krippendorff N, Schwindt C (2021) Control of shared production buffers: A reinforcement learning approach. In Chai KH, Moon SK, Jiao R, Xie M (eds) *Proceedings of the 2021 IEEE International Conference on Industrial Engineering and Engineering Management*, Singapore, pp. 703–707
- 11a Sutton RS, Barto AG (2018): Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, Chapter 13
- 11b Sewak M (2019) Deep Reinforcement Learning; Frontiers of Artificial Intelligence. Springer Singapore, Chapter 13
- 11c Zai A, Brown B (2020) Einstieg in Deep Reinforcement Learning. KI-Agenten mit Python und PyTorch programmieren. Hanser, München, Kapitel 4
- 12 Krause M (2021) Periodenübergreifende Budgetallokation in der Instandhaltung modularer Systeme: Ein Anwendungsfall von Deep Reinforcement Learning. Dissertation, Institut für Wirtschaftswissenschaft, Technische Universität Clausthal, Kapitel 6
- 13a Nandy A, Biswas M (2018) Reinforcement Learning with Open AI, TensorFlow and Keras Using Python. Apress Media by Springer, New York, Chapters 3, 4, 5
- 13b Lapan M (2020) Deep Reinforcement Learning Hands-on. Packt, Birmingham, Chapters 2, 3, 7
- 13c Sewak M (2019) Deep Reinforcement Learning; Frontiers of Artificial Intelligence. Springer Singapore, Chapter 7