

Aufgabenserie L3 zur Vorlesung "Maschinelles Lernen"

1. Eine Befragung zum Musikkonsum ergab folgende Ergebnisse in Form von Häufigkeiten zu Präferenzen  $i \succ j$ :

$i \succ j$	Rock/Pop	Schlager	Klassik	Country&Folk
Rock/Pop	-	12	17	22
Schlager	8	-	4	13
Klassik	5	11	-	18
Country&Folk	7	10	9	-

Finden Sie das Ranking heraus, für das die Scorefunktion des Rankings maximal wird. Führen Sie außerdem den Cohen-Algorithmus durch. Unterscheiden sich die Ergebnisse?

2. Ein Gefäß, das 8 Liter fasst, ist voll bis zum Rand mit Wasser gefüllt. Außerdem stehen noch zwei weitere Gefäße zur Verfügung, die 3 bzw. 5 Liter fassen. Diese sind zunächst beide leer. Durch wiederholtes Umfüllen sollen genau 4 Liter abgemessen werden. Formulieren Sie dieses Problem als Reinforcement Learning Problem (Zustände, Start, Zielzustand, Belohnung). Wie kann man ohne Wissen der optimalen Strategie vorgehen, um diese mit Hilfe der bekannten Algorithmen herauszubekommen?

3. Wir betrachten ein einfaches Labyrinth. Waagrecht von links nach rechts wird die  $x$ -Koordinate angegeben, senkrecht von oben nach unten die  $y$ - Koordinate. (1,1) ist der Startpunkt und (3,3) der Zielpunkt, in dem es den Wert 100 als Belohnung gibt.

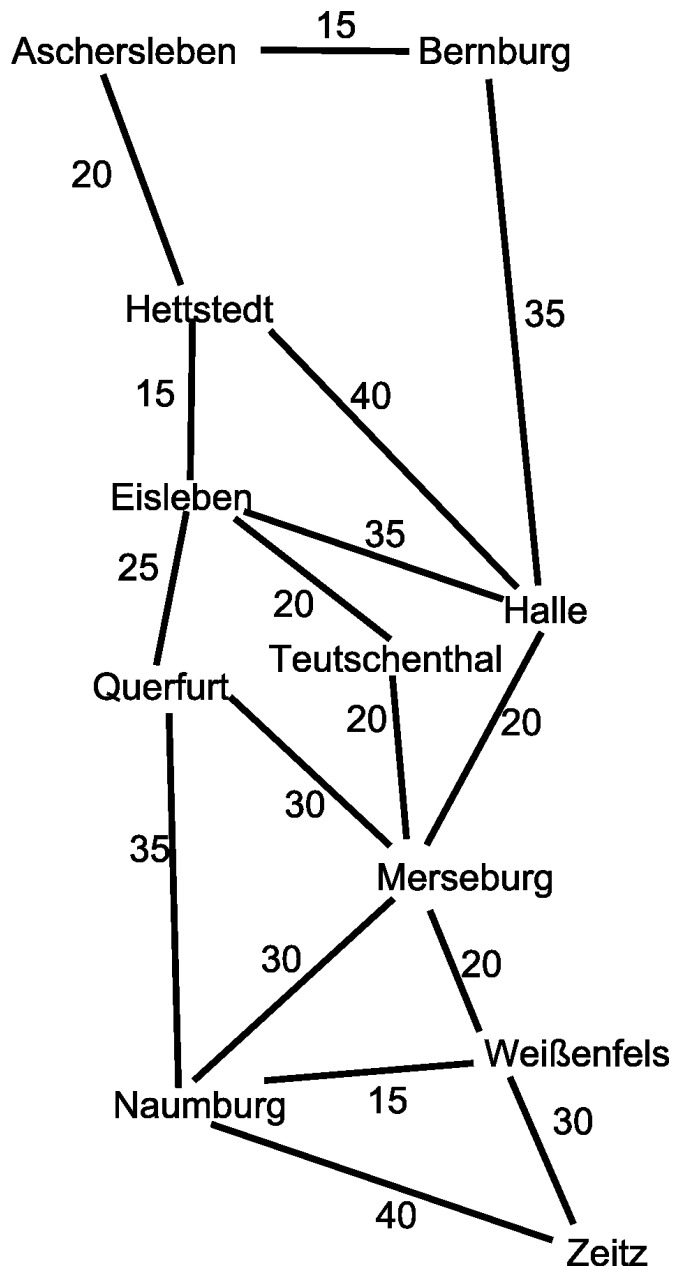
	■	■
	■	■
		Z
		■

Mit Hilfe des Differenz-Lernalgorithmus ist der kürzeste Weg zu finden. Folgende Episoden (Folge von Feldern) wurden durchgeführt:

- [1] (1,1),(1,2),(1,3),(1,4),(2,4),(2,3),(3,3)
- [2] (1,1),(1,2),(1,3),(2,3),(3,3)
- [3] (1,1),(1,2),(1,3),(1,4),(1,3),(2,3),(3,3)
- [4] (1,1),(1,2),(1,3),(1,2),(1,3),(2,3),(3,3)
- [5] (1,1),(1,2),(1,3),(1,4),(2,4),(2,3),(3,3)

Verwenden Sie einen Lernfaktor von 0.3 und  $\gamma = 0.7$ .

4. Es besteht das Problem, den schnellsten Weg von einem Ausgangspunkt zu einem Endpunkt zu finden. Dabei verändert sich die Fahrzeit von Mal zu Mal durch stochastische Einflüsse wie Ampeln, Staus u.ä. Simulieren Sie die Wahl des Weges und die Fahrzeiten. Bestimmen Sie dann mit Hilfe eines Algorithmus des verstärkten Lernens den optimalen Weg. Sie nutzen dazu das unten stehende Netz und suchen die kürzeste Verbindung von Aschersleben nach Zeitz. Die angegebenen Zeiten in Minuten sind die kürzesten Zeiten. Durch zufällige Störeinflüsse können die Zeiten bis zu 30 % länger sein (evtl. diesen Wert verändern).



Wege

ASL-HETT-EISL-QFT-NMB-ZZ [135,165]

ASL-HETT-EISL-QFT-MERS-WSF-ZZ [140,170]

ASL-HETT-EISL-TEUTSCH-MERS-WSF-ZZ [125,155]

ASL-BBG-HAL-MERS-WSF-ZZ [120,150]

zufällige Versionsnummer:  $=\text{Ceiling}(\text{Rnd}(4))$

**5.** Simulieren Sie den Ablauf des Q-Lern-Algorithmus im Rahmen des verstärkten Lernens. Dazu wird ein Raster mit Feldern vorgegeben, auf die der Agent ziehen kann. Einzelne Felder sind als Hindernisse zu markieren und dürfen nicht betreten werden. Der Agent zieht vom Startfeld links oben zum Zielfeld, welches definiert werden kann. In einem Zug wird auf ein Nachbarfeld links/rechts/oben/unten gerückt. Für die Felder können Direktbelohnungen festgelegt werden.

Diskutieren Sie die Ergebnisse für verschiedene Aufgabenstellungen hinsichtlich der Belohnungen, des Zielfeldes und des Lernfaktors. Das Ziel sollte sein, den kürzesten Weg bzw. den Weg mit der höchsten Belohnung zu finden.