Hinweise zur Bearbeitung von Hausaufgabe Nr. 3

- Die Bearbeitung muss in Gruppen von 3-5 Personen erfolgen
- Abgabe bis Sonntag, 09. Juli 2017, 23:59 CEST per Email an stefan.luedtke2@uni-rostock.de.
- Die Lösungen müssen als **PDF-Datei** abgegeben werden.
- Auf den Lösungsblätternmüssen die **Namen und Matrikelnummern aller Gruppenmitglieder** angegeben werden.
- Die Zahlen im Rand geben die erreichbaren Punkte pro Aufgabe an.

1 Hidden Markov Model 18

Ein Gefängnisinsasse möchte gerne wissen, wie das Wetter draußen ist. Er den Himmel aus dem Fenster seiner Zelle nicht direkt beobachten. Jeden Tag kommt ein Wärter an seiner Zelle vorbei. Dieser hat manchmal einen Regenschirm dabei, und manchmal nicht. Daraus möchte der Gefangene ableiten, wie das Wetter draußen ist.

Der Gefangene geht von folgenden Annahmen aus:

- Wenn es regnet, regnet es am nächsten Tag mit einer Wahrscheinlichkeit von 70 %. Zu 30 % ist es am nächsten Tag bewölkt.
- Wenn es bewölkt ist, gibt es am nächsten Tag zu 30 % Regen, zu 40 % ist es wieder bewölkt, und zu 30 % ist es sonnig.
- Wenn es sonnig ist, ist es am Folgetag wieder sonnig (30 %), oder es ist bewölkt (50 %), oder es regnet (20 %).
- Wenn es regent, hat der Wärter immer einen Schirm dabei.
- Wenn es bewölkt ist, hat der Wärter mit einer Wahrscheinlichkeit von 50 % einen Schirm dabei.
- Wenn es sonnig ist, hat der Wärter mit einer Wahrscheinlichkeit von 10 % einen Schirm dabei.
- Der Gefangene hat initial keine Informationen über das Wetter, d.h. er sieht alle drei Möglichkeiten als gleich wahrscheinlich an.

1. Spezifizieren Sie diesen Prozess als HMM. Notieren Sie die Initialzustand und das Beobachtungsmodell. 5
2. Berechnen Sie, ausgehend von der initialen Wahrscheinlichkeitsverteilung über den Zuständen, ...
   a) die *Vorhersage*, d.h., Wahrscheinlichkeitsverteilung allein auf Basis des Transitionsmodelles! 6
   b) die *Korrektur* nach der Beobachtung „kein Schirm“! 6

2 Bayessche Netze

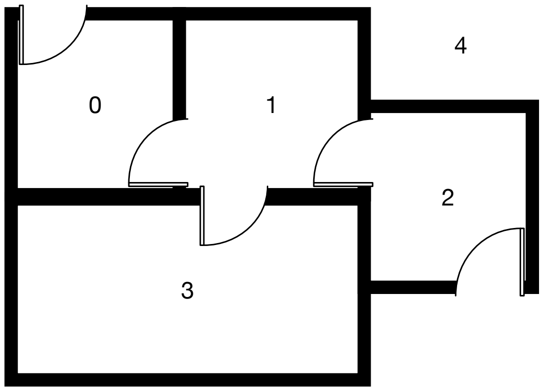
9

Die Todeswahrscheinlichkeit einer Person hängt (unter anderem) von zwei Faktoren ab: Ob die Person raucht (R) und ob die sie einen Sportwagen fährt (S). Die Wahrscheinlichkeit, dass eine Person im nächsten Jahr stirbt, wenn sie weder raucht, noch Sportwagen fährt, sei 10 %. Wenn sie raucht, oder einen Sportwagen fährt, sei die Wahrscheinlichkeit jeweils 20 %. Wenn sie beides macht, sei sie 40 %. Die a-priori-Wahrscheinlichkeit für Rauchen ist 40 %, für Sportwagen fahren 5 %. 4

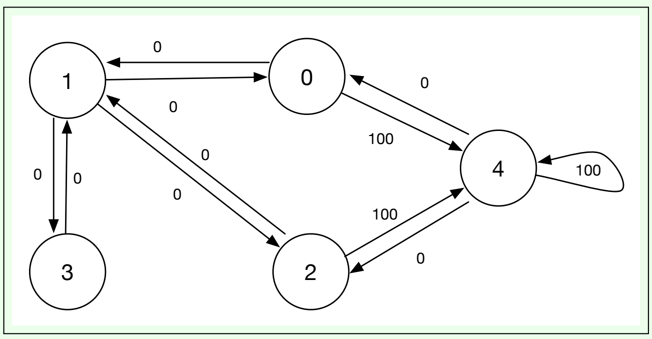
1. Modellieren Sie die beschriebene Situation als Bayessches Netz. 5
2. Bob ist gestorben. Er war starker Raucher. Berechnen Sie die Wahrscheinlichkeit, dass Bob einen Sportwagen gefahren ist.

3 Markov Decision Processes

Gegeben sei die folgende Umgebung mit 4 Räumen, die durch Türen verbunden sind. Die Räume sind von 0 bis 3 nummeriert und 4 bedeutet draußen. Ein Roboter soll, nachdem er in einigem beliebigen Raum positioniert wurde, durch Q-learning den Weg nach draußen finden, d.h., 4 (draußen) ist der Zielzustand. Alle Türen die nach draußen führen haben ein Reward von 100 und alle anderen 0. 15



1. Modellieren Sie die Räume als einen Graph, in dem die Räume als Zustände (Knoten) und die Türverbindungen als Transitionen gegeben sind. Notieren Sie die Rewards für jede Transition. 3



2. Geben Sie die Reward-Matrix R, die die Rewards zwischen den Zustände beschreibt. Unmögliche Transitionen bekommen ein Reward von -1. 3

$R =$

State	*Action* 0	1	2	3	4
0	-1	0	-1	-1	100
1	0	-1	0	0	-1
2	-1	0	-1	-1	100
3	-1	0	-1	-1	-1
4	0	-1	0	-1	100

3. Geben Sie die initiale Q-Matrix an! Nehmen Sie dabei an, dass der Roboter die Anzahl der Räume kennt. 3

$Q =$

State	*Action* 0	1	2	3	4
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0

4. Berechnen Sie die Rewards für eine "Episode", die im Raum 1 anfängt. Benutzen Sie den "discount factor" $\gamma = 0.5$. Benutzen Sie dafür die folgende Formel (diese resultiert aus der Standardregel für die Lernrate $\alpha = 1$).

$$Q(state, action) = R(state, action) + \gamma * Max[Q(next\ \ state, all\ \ actions)].$$

Dabei sollen die Aktionen so gewählt werden, das der Roboter die Räume in der folgenden Reihenfolge besucht: $1 \to 3 \to 1 \to 2 \to 4$. Geben Sie die resultierende Q-Matrix an! 5

*Q(1,3) = 0 + 0.5 * 0 = 0*

*Q(3,1) = 0 + 0.5 * 0 = 0*

*Q(1,2) = 0 + 0.5 * 0 = 0*

*Q(2,4) = 100 + 0.5 * 0 = 100*

$Q =$

State	*Action* 0	1	2	3	4
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	100
3	0	0	0	0	0
4	0	0	0	0	0