Grundlagen adaptiver Wissenssysteme

Übungen zur Vorlesung

Prof. Dr. Thomas Gabel
Frankfurt University of Applied Sciences
Faculty of Computer Science and Engineering
tgabel@fb2.fra-uas.de

Aufgabenblatt 1

1. **Aufgabenblatt 1 – Übung 1**
2. Aufgabenblatt 1 – Übung 2
3. Aufgabenblatt 1 – Übung 3
4. Aufgabenblatt 1 – Übung 4

Aufgabe 1: Problemmodellierung als MDP

Modellieren Sie die folgenden Problemstellungen als Markov'sche Entscheidungsprozesse:

(a) autonomes Autofahren auf einer Autobahn
(b) Staubsaugeroboter
(c) Agent in einem Labyrinth
(d) Heizungsregler
(e) Apres-Ski-Hütten-Betreiber

Aufgabe 1: Problemmodellierung als MDP

(a) autonomes Autofahren auf einer Autobahn

- S -
- A -
- p -
- c -
- T -

Aufgabe 1: Problemmodellierung als MDP

(b) Staubsaugeroboter

- S -
- A -
- p -
- c -
- T -

FRANKFURT UNIVERSITY OF APPLIED SCIENCES

Aufgabe 1: Problemmodellierung als MDP

(c) Agent in einem Labyrinth

- S -
- A -
- p -
- c -
- T -

Aufgabe 1: Problemmodellierung als MDP

(d) Heizungsregler

- S -
- A -
- p -
- c -
- T -

Aufgabe 1: Problemmodellierung als MDP

(e) Apres-Ski-Hütten-Betreiber

- S -
- A -
- p -
- c -
- T -

Aufgabenblatt 1

1. Aufgabenblatt 1 – Übung 1
2. **Aufgabenblatt 1 – Übung 2**
3. Aufgabenblatt 1 – Übung 3
4. Aufgabenblatt 1 – Übung 4

Aufgabe 2: Markov-Eigenschaft

Betrachten Sie ein intelligentes kamerabasiertes System. Nach dem Einschalten nimmt die Kamera ein erstes Bild der Szene auf. Verschiedene Objekte können gesehen werden, nicht jedoch solche, die beispielsweise verdeckt sind oder die sich hinter der Kamera befinden.

Aufgabe 2: Markov-Eigenschaft

Betrachten Sie ein intelligentes kamerabasiertes System. Nach dem Einschalten nimmt die Kamera ein erstes Bild der Szene auf. Verschiedene Objekte können gesehen werden, nicht jedoch solche, die beispielsweise verdeckt sind oder die sich hinter der Kamera befinden.

(a) Ist für die aufgenommene Umgebung die Markov-Eigenschaft erfüllt, sobald die Kamera ihr erstes Bild aufgenommen hat?

Aufgabe 2: Markov-Eigenschaft

Betrachten Sie ein intelligentes kamerabasiertes System. Nach dem Einschalten nimmt die Kamera ein erstes Bild der Szene auf. Verschiedene Objekte können gesehen werden, nicht jedoch solche, die beispielsweise verdeckt sind oder die sich hinter der Kamera befinden.

(a) Ist für die aufgenommene Umgebung die Markov-Eigenschaft erfüllt, sobald die Kamera ihr erstes Bild aufgenommen hat?

- In a) hat das System keine Informationen über frühere Systemzustände erhalten und solche nicht in die Beschreibung des aktuellen Zustandes integriert (z.B. ein Mensch hat sich hinter einem Objekt versteckt). Folgezustand kann auch bei bekannten Aktionen nicht prediziert werden; die ME ist nicht gegeben.

Aufgabe 2: Markov-Eigenschaft

Betrachten Sie ein intelligentes kamerabasiertes System. Nach dem Einschalten nimmt die Kamera ein erstes Bild der Szene auf. Verschiedene Objekte können gesehen werden, nicht jedoch solche, die beispielsweise verdeckt sind oder die sich hinter der Kamera befinden.

(b) Ist die Markov-Eigenschaft gegeben, wenn die Kamera defekt ist und daher permanent das gleiche (Stand-)Bild liefert?

Aufgabe 2: Markov-Eigenschaft

Betrachten Sie ein intelligentes kamerabasiertes System. Nach dem Einschalten nimmt die Kamera ein erstes Bild der Szene auf. Verschiedene Objekte können gesehen werden, nicht jedoch solche, die beispielsweise verdeckt sind oder die sich hinter der Kamera befinden.

(b) Ist die Markov-Eigenschaft gegeben, wenn die Kamera defekt ist und daher permanent das gleiche (Stand-)Bild liefert?

- Der Zustand des Systems (so wie er durch die defekte Kamera geliefert wird) ändert sich zu keinem Zeitpunkt (durch keine Aktionen), es gilt stets $s_{t+1} = s_t$. Damit ist die Markov-Eigenschaft erfüllt, auch wenn solch ein “defektes” System für praktische Überlegungen ohne Bedeutung ist.

Aufgabenblatt 1

1. Aufgabenblatt 1 – Übung 1
2. Aufgabenblatt 1 – Übung 2
3. **Aufgabenblatt 1 – Übung 3**
4. Aufgabenblatt 1 – Übung 4

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

In dieser Aufgabe setzen wir uns mit Problemen mit einem endlichen Horizont der Länge N auseinander.
Betrachten wir einen Studenten, der an einer schriftlichen Prüfung teilnimmt. Der Aufgabenzettel umfasst k zu lösende verschiedene Aufgaben, die jeweils r_i ($i \in \{1, \dots, k\}$, sh. Tabelle auf folgender Folie) Punkte erbringen. Um die Problemstellung einfach zu halten, nehmen wir an, dass jede Aufgabe nur entweder komplett falsch oder komplett korrekt gelöst werden kann (im ersten Fall gibt es 0 Punkte, in letzterem r_i) und dass der Student, nachdem er sich mit einer Aufgabe auseinandergesetzt hat, mit Sicherheit weiß, ob er sie richtig gelöst hat oder nicht.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Der Student kann jede Aufgabe beliebig oft zu lösen versuchen, wobei jeder Lösungsversuch mit einer Wahrscheinlichkeit p_i^S gelingt, die vom Schwierigkeitsgrad und vom Kenntnisstand des Studenten abhängt (diese Wahrscheinlichkeiten sind in der folgenden Tabelle aufgelistet).

Aufgabe	Punktzahl	Lösungswahrscheinlichkeit
1	4	0.1
2	1	0.8
3	3	0.3
4	2	0.5

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Der Student kann jede Aufgabe beliebig oft zu lösen versuchen, wobei jeder Lösungsversuch mit einer Wahrscheinlichkeit p_i^S gelingt, die vom Schwierigkeitsgrad und vom Kenntnisstand des Studenten abhängt (diese Wahrscheinlichkeiten sind in der folgenden Tabelle aufgelistet).

Aufgabe	Punktzahl	Lösungswahrscheinlichkeit
1	4	0.1
2	1	0.8
3	3	0.3
4	2	0.5

Selbstverständlich ist die Bearbeitungszeit in Prüfungen begrenzt. Wir nehmen im Folgenden an, dass die Dauer der Prüfung es dem Studenten ermöglicht, insgesamt bis zu $N = 5$ Aufgabenlösungsversuche zu unternehmen (vereinfachend dauere jeder Versuch gleich lang). Ferner ist bekannt, dass zum Bestehen der Prüfung mindestens 40% der maximal erzielbaren Punkte erreicht werden müssen.

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

(a) Formalisieren Sie das beschriebene Problem als Markov'schen Entscheidungsprozess.

FRANKFURT UNIVERSITY OF APPLIED SCIENCES

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(a) Formalisieren Sie das beschriebene Problem als Markov'schen Entscheidungsprozess.

- MDP $M = [T, S, A, p, c]$
- Entscheidungszeitpunkte

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(a) Formalisieren Sie das beschriebene Problem als Markov'schen Entscheidungsprozess.

- MDP $M = [T, S, A, p, c]$
- Entscheidungszeitpunkte $T = \{0, \ldots, N\}$
- Zustände:

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(a) Formalisieren Sie das beschriebene Problem als Markov'schen Entscheidungsprozess.

- MDP $M = [T, S, A, p, c]$
- Entscheidungszeitpunkte $T = \{0, \dots, N\}$
- Zustände: Jede Aufgabe kann gelöst oder ungelöst sein. Damit ergeben sich 2^m mögliche Zustände, wenn m die Anzahl der Aufgaben in der Prüfung bezeichnet.
- Also:

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(a) Formalisieren Sie das beschriebene Problem als Markov'schen Entscheidungsprozess.

- MDP $M = [T, S, A, p, c]$
- Entscheidungszeitpunkte $T = \{0, \ldots, N\}$
- Zustände: Jede Aufgabe kann gelöst oder ungelöst sein. Damit ergeben sich 2^m mögliche Zustände, wenn m die Anzahl der Aufgaben in der Prüfung bezeichnet.
- Also: $S = \{i = (i_1, \ldots, i_m) | i_v \in 0, 1\}$.
- Da in der gegebenen Aufgabe $m = 4$ gilt, ergibt sich $|S| = 2^4 = 16$.
- Der Student kann jeweils eine der noch nicht gelösten Aufgaben mit einem Lösungsversuch zu lösen versuchen, also Aktionen

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(a) Formalisieren Sie das beschriebene Problem als Markov'schen Entscheidungsprozess.

- MDP $M = [T, S, A, p, c]$
- Entscheidungszeitpunkte $T = \{0, \ldots, N\}$
- Zustände: Jede Aufgabe kann gelöst oder ungelöst sein. Damit ergeben sich 2^m mögliche Zustände, wenn m die Anzahl der Aufgaben in der Prüfung bezeichnet.
- Also: $S = \{i = (i_1, \ldots, i_m) | i_v \in 0, 1\}$.
- Da in der gegebenen Aufgabe $m = 4$ gilt, ergibt sich $|S| = 2^4 = 16$.
- Der Student kann jeweils eine der noch nicht gelösten Aufgaben mit einem Lösungsversuch zu lösen versuchen, also Aktionen $A = \{a_1, a_2, a_3, a_4\}$ ($A(i)$ variiert mit der Zeit).

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

- Die Übergangsfunktion

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

- Die Übergangsfunktion ist stochastisch und hängt von den in der zuvor angegebenen Tabelle genannten Wahrscheinlichkeiten ab.
- Beispielsweise gilt $p_{ij}(a) = 0.1$ für $i = (0, 0, 1, 0)$ und $j = (1, 0, 1, 0)$ unter $a = a_1$.
- Die Kostenfunktion $c : S \times A \to \mathbb{R}$ muss

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

- Die Übergangsfunktion ist stochastisch und hängt von den in der zuvor angegebenen Tabelle genannten Wahrscheinlichkeiten ab.
- Beispielsweise gilt $p_{ij}(a) = 0.1$ für $i = (0, 0, 1, 0)$ und $j = (1, 0, 1, 0)$ unter $a = a_1$.
- Die Kostenfunktion $c : S \times A \to \mathbb{R}$ muss zum Optimierungsziel korrespondieren, was im konkreten Fall bedeutet, so viele Punkte wie möglich zu erzielen.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

- Die Übergangsfunktion ist stochastisch und hängt von den in der zuvor angegebenen Tabelle genannten Wahrscheinlichkeiten ab.
- Beispielsweise gilt $p_{ij}(a) = 0.1$ für $i = (0, 0, 1, 0)$ und $j = (1, 0, 1, 0)$ unter $a = a_1$.
- Die Kostenfunktion $c : S \times A \to \mathbb{R}$ muss zum Optimierungsziel korrespondieren, was im konkreten Fall bedeutet, so viele Punkte wie möglich zu erzielen.
- In der hier gegebenen Aufgabenstellung hängen die Kosten (oder analog: negierte Belohnungen) nicht nur vom aktuellen Zustand und der Aktion ab, sondern auch vom eingetretenen Folgezustand. Daher ist die in der Literatur gebräuchliche Modellierung als $c : S \times A \times S \to \mathbb{R}$ angebracht.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

- Die Übergangsfunktion ist stochastisch und hängt von den in der zuvor angegebenen Tabelle genannten Wahrscheinlichkeiten ab.
- Beispielsweise gilt $p_{ij}(a) = 0.1$ für $i = (0, 0, 1, 0)$ und $j = (1, 0, 1, 0)$ unter $a = a_1$.
- Die Kostenfunktion $c : S \times A \to \mathbb{R}$ muss zum Optimierungsziel korrespondieren, was im konkreten Fall bedeutet, so viele Punkte wie möglich zu erzielen.
- In der hier gegebenen Aufgabenstellung hängen die Kosten (oder analog: negierte Belohnungen) nicht nur vom aktuellen Zustand und der Aktion ab, sondern auch vom eingetretenen Folgezustand. Daher ist die in der Literatur gebräuchliche Modellierung als $c : S \times A \times S \to \mathbb{R}$ angebracht.
- Wir setzen
$$c(i, a, j) = \begin{cases} 0 & i = j \\ -z_a & \text{else} \end{cases}$$
wobei mit z_a die Anzahl der Punkte bezeichnet wird, die es für die Aufgabe a gibt (sh. Tabelle).

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

Ferner ist bekannt, dass zum Bestehen der Prüfung mindestens 40% der maximal erzielbaren Punkte erreicht werden müssen.

(b) Wie würden Sie das Risiko des Nicht-Bestehens in diesem Szenario modellieren?

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

Ferner ist bekannt, dass zum Bestehen der Prüfung mindestens 40% der maximal erzielbaren Punkte erreicht werden müssen.

(b) Wie würden Sie das Risiko des Nicht-Bestehens in diesem Szenario modellieren?

- Im konkreten Fall muss der Student also mindestens 4 von 10 Punkten erzielen, um zu bestehen.
- In N-stufigen Entscheidungsproblemen hat jeder Zustand Terminalkosten $g(i)$, welche anfallen, wenn das System nach N Stufen im Zustand i terminiert.
- Die Anzahl der erreichten Punkte (und damit den Zusammenhang zur Note: höhere Punktzahlen entsprechen niedrigeren Kosten und einer besseren Note) haben wir bereits durch die direkten Kosten modelliert.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(b) Wie würden Sie das Risiko des Nicht-Bestehens in diesem Szenario modellieren?

- Für Terminalzustände i, in denen die Prüfung bestanden wurde, sind also nicht notwendigerweise Terminalkosten zu definieren, d.h. $g(i) = 0$.
- Für Terminalzustände i, bei denen insgesamt nur 3 oder weniger Punkte erzielt wurden, definieren wir $g(i) = D$.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(b) Wie würden Sie das Risiko des Nicht-Bestehens in diesem Szenario modellieren?

- Für Terminalzustände i, in denen die Prüfung bestanden wurde, sind also nicht notwendigerweise Terminalkosten zu definieren, d.h. $g(i) = 0$.
- Für Terminalzustände i, bei denen insgesamt nur 3 oder weniger Punkte erzielt wurden, definieren wir $g(i) = D$.
- In welcher Relation D zu den direkten Kosten bzw. Pfadkosten steht, ist "Ermessensfrage" des Studenten.
  - Z.B. wenn vom Bestehen der Prüfung die Fortführung des Studiums abhängt, so würde man D sehr groß wählen.
- Im Folgenden setzen wir $D = 10$.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Im Folgenden sollen verschiedene “Prüfungsstrategien” verglichen werden, d.h. die Verfahrensweisen, nach denen verschiedene Studenten die zu bearbeitenden Aufgaben auswählen. Für alle folgenden Berechnungen soll für das Nicht-Bestehen $g(i_N) = 10$ für $i_N \in S_{fail}$ angenommen werden.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Im Folgenden sollen verschiedene "Prüfungsstrategien" verglichen werden, d.h. die Verfahrensweisen, nach denen verschiedene Studenten die zu bearbeitenden Aufgaben auswählen. Für alle folgenden Berechnungen soll für das Nicht-Bestehen $g(i_N) = 10$ für $i_N \in S_{fail}$ angenommen werden.

(c) Student $\mathcal{S}$ schwankt zwischen zwei möglichen Strategien, $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$, gemäß derer er seine Aufgaben auswählen will. Bei der erstgenannnten bearbeitet er die gestellten Aufgaben in der Reihenfolge aufsteigender Schwierigkeit, d.h. sinkender Erfolgswahrscheinlichkeit $p_i^{\mathcal{S}}$. Bei Strategie $\pi_B^{\mathcal{S}}$ würde er die Aufgaben in umgekehrter Reihenfolge bearbeiten. Vergleichen Sie die beiden Vorgehensweisen, indem Sie die Pfadkosten der Strategien $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$ ermitteln.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Im Folgenden sollen verschiedene "Prüfungsstrategien" verglichen werden, d.h. die Verfahrensweisen, nach denen verschiedene Studenten die zu bearbeitenden Aufgaben auswählen. Für alle folgenden Berechnungen soll für das Nicht-Bestehen $g(i_N) = 10$ für $i_N \in S_{fail}$ angenommen werden.

(c) Student $\mathcal{S}$ schwankt zwischen zwei möglichen Strategien, $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$, gemäß derer er seine Aufgaben auswählen will. Bei der erstgenannnten bearbeitet er die gestellten Aufgaben in der Reihenfolge aufsteigender Schwierigkeit, d.h. sinkender Erfolgswahrscheinlichkeit $p_i^{\mathcal{S}}$. Bei Strategie $\pi_B^{\mathcal{S}}$ würde er die Aufgaben in umgekehrter Reihenfolge bearbeiten. Vergleichen Sie die beiden Vorgehensweisen, indem Sie die Pfadkosten der Strategien $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$ ermitteln.

- $\rightarrow$ Tafel

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Im Folgenden sollen verschiedene “Prüfungsstrategien” verglichen werden, d.h. die Verfahrensweisen, nach denen verschiedene Studenten die zu bearbeitenden Aufgaben auswählen. Für alle folgenden Berechnungen soll für das Nicht-Bestehen $g(i_N) = 10$ für $i_N \in S_{fail}$ angenommen werden.

(c) Student $\mathcal{S}$ schwankt zwischen zwei möglichen Strategien, $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$, gemäß derer er seine Aufgaben auswählen will. Bei der erstgenannnten bearbeitet er die gestellten Aufgaben in der Reihenfolge aufsteigender Schwierigkeit, d.h. sinkender Erfolgswahrscheinlichkeit $p_i^{\mathcal{S}}$. Bei Strategie $\pi_B^{\mathcal{S}}$ würde er die Aufgaben in umgekehrter Reihenfolge bearbeiten. Vergleichen Sie die beiden Vorgehensweisen, indem Sie die Pfadkosten der Strategien $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$ ermitteln.

- $\rightarrow$ Tafel
- Resultate:
  $V^{\pi_A^{\mathcal{S}}} = 1.13$ und $V^{\pi_B^{\mathcal{S}}} = 3.45 \Rightarrow$ **Strategie A ist besser.**

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Im Folgenden sollen verschiedene “Prüfungsstrategien” verglichen werden, d.h. die Verfahrensweisen, nach denen verschiedene Studenten die zu bearbeitenden Aufgaben auswählen. Für alle folgenden Berechnungen soll für das Nicht-Bestehen $g(i_N) = 10$ für $i_N \in S_{fail}$ angenommen werden.

(d) Leiten Sie eine stationäre Strategie $\pi_C^{\mathcal{S}}$ her, die erfolgsversprechender als $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$ ist. Welche Probleme ergäben sich für Sie, wenn Sie eine solche Strategie in der Praxis anwenden wollten?

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Im Folgenden sollen verschiedene “Prüfungsstrategien” verglichen werden, d.h. die Verfahrensweisen, nach denen verschiedene Studenten die zu bearbeitenden Aufgaben auswählen. Für alle folgenden Berechnungen soll für das Nicht-Bestehen $g(i_N) = 10$ für $i_N \in S_{fail}$ angenommen werden.

(d) Leiten Sie eine stationäre Strategie $\pi_C^{\mathcal{S}}$ her, die erfolgsversprechender als $\pi_A^{\mathcal{S}}$ und $\pi_B^{\mathcal{S}}$ ist. Welche Probleme ergäben sich für Sie, wenn Sie eine solche Strategie in der Praxis anwenden wollten?

- Die Strategie $\pi_C^{\mathcal{S}}$ ist nicht eindeutig bestimmt.
- Eine mögliche Wahl berücksichtigt die Erfolgswahrscheinlichkeiten der Lösungsversuche und die erzielbaren Punkte gleichzeitig.

$$\pi_C^{\mathcal{S}}(i) \in \arg\max_{a \in A(i)} z_a \cdot p_{ij}(a).$$

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Aufgabe	Punktzahl	Lösungswahrscheinlichkeit
1	4	0.1
2	1	0.8
3	3	0.3
4	2	0.5

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

Aufgabe	Punktzahl	Lösungswahrscheinlichkeit
1	4	0.1
2	1	0.8
3	3	0.3
4	2	0.5

- Eine mögliche Wahl berücksichtigt die Erfolgswahrscheinlichkeiten der Lösungsversuche und die erzielbaren Punkte gleichzeitig.

$$\pi_C^{\mathcal{S}}(i) \in \arg\max_{a \in A(i)} z_a \cdot p_{ij}(a).$$

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

Aufgabe	Punktzahl	Lösungswahrscheinlichkeit
1	4	0.1
2	1	0.8
3	3	0.3
4	2	0.5

- Eine mögliche Wahl berücksichtigt die Erfolgswahrscheinlichkeiten der Lösungsversuche und die erzielbaren Punkte gleichzeitig.

$$\pi_C^{\mathcal{S}}(i) \in \arg\max_{a \in A(i)} z_a \cdot p_{ij}(a).$$

- Diese Strategie wählt die Aufgaben daher in der Reihenfolge 4, 3, 1, 2. Der Ausdruck $z_a \cdot p_{ij}(a)$ nimmt hier die Were 1.0, 0.9, 0.8, 0.4 an.
- Für die Pfadkosten ergibt sich: $V^{\pi_C^{\mathcal{S}}}(i_0) = -0.70$ $(< V^{\pi_A^{\mathcal{S}}} < V^{\pi_B^{\mathcal{S}}})$.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(e) Schlagen Sie eine Methode vor, durch die jede der soweit betrachteten Strategien im Hinblick auf die Reduzierung der Wahrscheinlichkeit des Nicht-Bestehens verbessert werden kann. *Tipp: Verwenden Sie nicht stationäre Strategien.* Inwieweit werden durch Ihre Methode auch die Pfadkosten verbessert?

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(e) Schlagen Sie eine Methode vor, durch die jede der soweit betrachteten Strategien im Hinblick auf die Reduzierung der Wahrscheinlichkeit des Nicht-Bestehens verbessert werden kann. *Tipp: Verwenden Sie nicht stationäre Strategien.* Inwieweit werden durch Ihre Methode auch die Pfadkosten verbessert?

- Alle bisher betrachteten Strategien waren stationär, d.h. sie berücksichtigten bei der Aktionswahl nicht die aktuelle Stufe des Entscheidungsproblems.
- Wenn eine Aufgabe nicht gelöst werden konnte (und sich somit der Zustand nicht änderte), wurde stets die gleiche Aktion wieder ausgewählt.

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(e) Schlagen Sie eine Methode vor, durch die jede der soweit betrachteten Strategien im Hinblick auf die Reduzierung der Wahrscheinlichkeit des Nicht-Bestehens verbessert werden kann. *Tipp: Verwenden Sie nicht stationäre Strategien.* Inwieweit werden durch Ihre Methode auch die Pfadkosten verbessert?

- Alle bisher betrachteten Strategien waren stationär, d.h. sie berücksichtigten bei der Aktionswahl nicht die aktuelle Stufe des Entscheidungsproblems.
- Wenn eine Aufgabe nicht gelöst werden konnte (und sich somit der Zustand nicht änderte), wurde stets die gleiche Aktion wieder ausgewählt.
- Im Hinblick auf die Gefahr, bei der Prüfung durchzufallen, ist solch ein Vorgehen alles andere als optimal.
  - Z.B. macht es ganz offensichtlich wenig Sinn, kurz vor Abgabe noch eine Aufgabe lösen zu wollen, die nur p Punkte bringt, wenn klar ist, dass ganz sicher noch q (mit $p < q$) weitere Punkte zum Bestehen notwendig wären.

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

(e) Schlagen Sie eine Methode vor, durch die jede der soweit betrachteten Strategien im Hinblick auf die Reduzierung der Wahrscheinlichkeit des Nicht-Bestehens verbessert werden kann. *Tipp: Verwenden Sie nicht stationäre Strategien.* Inwieweit werden durch Ihre Methode auch die Pfadkosten verbessert?

- Naheliegende Möglichkeit zur Strategieverbesserung: Wenn nur noch ein Lösungsversuch unternommen werden kann und noch q Punkte zum Bestehen benötigt werden, wird eine Aufgabe gewählt, die q Punkte erbringt.

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen

(e) Schlagen Sie eine Methode vor, durch die jede der soweit betrachteten Strategien im Hinblick auf die Reduzierung der Wahrscheinlichkeit des Nicht-Bestehens verbessert werden kann. *Tipp: Verwenden Sie nicht stationäre Strategien.* Inwieweit werden durch Ihre Methode auch die Pfadkosten verbessert?

- Naheliegende Möglichkeit zur Strategieverbesserung: Wenn nur noch ein Lösungsversuch unternommen werden kann und noch q Punkte zum Bestehen benötigt werden, wird eine Aufgabe gewählt, die q Punkte erbringt.
- Erweiterung der Strategie $\pi_A^{\mathcal{S}}$ des Studenten $\mathcal{S}$ (leichteste Aufgaben zuerst) zur Strategie $\pi_D^{\mathcal{S}}$ durch Hinzufügung dieses nicht-stationären Zusatzes in der letzten Entscheidungsstufe, bringt für unser Szenario:

$$V^{\pi_D^{\mathcal{S}}} = 0.65 \; (< V^{\pi_A^{\mathcal{S}}} = 1.13).$$

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(f) Student $\mathcal{T}$ hat selektiv gelernt und sich gar nicht mit den Themen befasst, die in den Aufgaben 2-4 abgefragt werden ($p_i^{\mathcal{T}} = 0.0$ für $i \in \{2,3,4\}$). Wie gut muss $\mathcal{T}$ für das in Aufgabe 1 behandelte Themengebiet vorbereitet sein (d.h. wie groß muss $p_1^{\mathcal{T}}$ mindestens sein), damit $\mathcal{T}$ mindestens so erfolgreich in der Prüfung ist wie Student $\mathcal{S}$, wenn dieser nach Strategie $\pi_A^{\mathcal{S}}$ verfährt?

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

(f) Student $\mathcal{T}$ hat selektiv gelernt und sich gar nicht mit den Themen befasst, die in den Aufgaben 2-4 abgefragt werden ($p_i^{\mathcal{T}} = 0.0$ für $i \in \{2, 3, 4\}$). Wie gut muss $\mathcal{T}$ für das in Aufgabe 1 behandelte Themengebiet vorbereitet sein (d.h. wie groß muss $p_1^{\mathcal{T}}$ mindestens sein), damit $\mathcal{T}$ mindestens so erfolgreich in der Prüfung ist wie Student $\mathcal{S}$, wenn dieser nach Strategie $\pi_A^{\mathcal{S}}$ verfährt?

- Es ist klar, das für Student $\mathcal{T}$ die Strategie $\pi^{\mathcal{T}}(i) = a_1$ optimal ist, wenn wir von $p_1^{\mathcal{T}} > 0$ ausgehen.
- Es ergibt sich die Frage, für welche Werte $p_1^{\mathcal{T}} > p_{1,min}^{\mathcal{T}}$ gilt:

$$V^{\pi^{\mathcal{T}}} < V^{\pi_A^{\mathcal{S}}} = 1.13$$

Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

- Für $V^{\pi^{\mathcal{T}}}$ erhalten wird

$$
\begin{aligned}
V^{\pi^{\mathcal{T}}}(p_1^{\mathcal{T}}) \quad = \quad & (1 - p_1^{\mathcal{T}})^5 \cdot 10 \\
& + (1 - p_1^{\mathcal{T}})^4 \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& + (1 - p_1^{\mathcal{T}})^3 \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& + (1 - p_1^{\mathcal{T}})^2 \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& + (1 - p_1^{\mathcal{T}}) \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& + p_1^{\mathcal{T}} \cdot (-4) \\
= \quad & 10(1 - p_1^{\mathcal{T}})^5 - 4p_1^{\mathcal{T}}((1 - p_1^{\mathcal{T}})^4 + (1 - p_1^{\mathcal{T}})^3 \\
& + (1 - p_1^{\mathcal{T}})^2 + (1 - p_1^{\mathcal{T}}) + 1)
\end{aligned}
$$

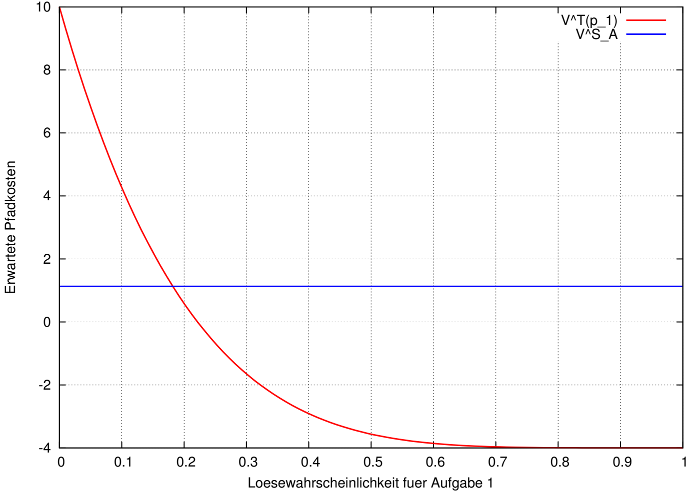
Aufgabe 3: Pfadkosten in N-stufigen Entscheidungsproblemen

- Für $V^{\pi^{\mathcal{T}}}$ erhalten wird

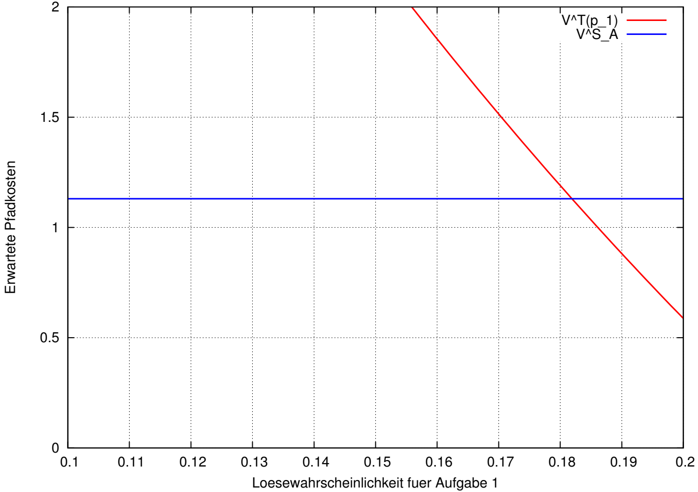
$$
\begin{aligned}
V^{\pi^{\mathcal{T}}}(p_1^{\mathcal{T}}) \quad = \quad & (1-p_1^{\mathcal{T}})^5 \cdot 10 \\
& +(1-p_1^{\mathcal{T}})^4 \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& +(1-p_1^{\mathcal{T}})^3 \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& +(1-p_1^{\mathcal{T}})^2 \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& +(1-p_1^{\mathcal{T}}) \cdot p_1^{\mathcal{T}} \cdot (-4) \\
& +p_1^{\mathcal{T}} \cdot (-4) \\
= \quad & 10(1-p_1^{\mathcal{T}})^5 - 4p_1^{\mathcal{T}}((1-p_1^{\mathcal{T}})^4 + (1-p_1^{\mathcal{T}})^3 \\
& +(1-p_1^{\mathcal{T}})^2 + (1-p_1^{\mathcal{T}}) + 1)
\end{aligned}
$$

- Damit gilt $V^{\pi^{\mathcal{T}}} < V^{\pi_A^{\mathcal{S}}} = 1.13$ für $p_1^{\mathcal{T}} > p_{1,min}^{\mathcal{T}} \approx 0.182$

Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen



Aufgabe 3: Pfadkosten in *N*-stufigen Entscheidungsproblemen



Aufgabenblatt 1

1. Aufgabenblatt 1 – Übung 1
2. Aufgabenblatt 1 – Übung 2
3. Aufgabenblatt 1 – Übung 3
4. **Aufgabenblatt 1 – Übung 4**

Aufgabe 4: Backward-DP-Algorithmus

(a) Schlagen Sie eine Modifikation des in der Vorlesung vorgestellen Backward-DP-Algorithmus vor, die diesen für Umgebungen einsetzbar macht, in denen die direkten Kosten vom Folgezustand abhängen können.

Aufgabe 4: Backward-DP-Algorithmus

(a) Schlagen Sie eine Modifikation des in der Vorlesung vorgestellen Backward-DP-Algorithmus vor, die diesen für Umgebungen einsetzbar macht, in denen die direkten Kosten vom Folgezustand abhängen können.

- Backward-DP-Algorithmus

Require: MDP $M = [T, S, A, p, c]$ mit $T = \{0, \ldots, N\}$ und Terminalkosten g

1: $V_0^*(i) = g(i)$ //$k = 0$
2: **for** $k = 1$ **to** N **and** $\forall i \in S$ **do**
3: $\quad V_k^*(i) = \min_{a \in A(i)} \sum_{j=1}^{n} p_{ij}(a) \left(c(i, a) + V_{k-1}^*(j) \right)$
4: **end for**
5: **return** $\{V_k^* | k = 0, \ldots, N\}$

Aufgabe 4: Backward-DP-Algorithmus

- Backward-DP-Algorithmus

$V_0^*(i) = g(i)$ //$k = 0$
for $k = 1$ **to** N **and** $\forall i \in S$ **do**
  $V_k^*(i) = \min_{a \in A(i)} \sum_{j=1}^{n} p_{ij}(a) \left(c(i,a) + V_{k-1}^*(j) \right)$
end for
return $\{V_k^* | k = 0, \ldots, N\}$

- Backward-DP-Algorithmus für folgezustandsabhängige direkte Kosten

$V_0^*(i) = g(i)$ //$k = 0$
for $k = 1$ **to** N **and** $\forall i \in S$ **do**
  $V_k^*(i) = \min_{a \in A(i)} \sum_{j=1}^{n} p_{ij}(a) \left(c(i,a,j) + V_{k-1}^*(j) \right)$
end for
return $\{V_k^* | k = 0, \ldots, N\}$

Aufgabe 4: Backward-DP-Algorithmus

(b) Wenden Sie Ihren Backward-DP-Algorithmus auf das Problem aus Aufgabe 3 an (für die Schritte $k = 0$ und $k = 1$).

Aufgabe 4: Backward-DP-Algorithmus

(b) Wenden Sie Ihren Backward-DP-Algorithmus auf das Problem aus Aufgabe 3 an (für die Schritte $k = 0$ und $k = 1$).

- $\rightarrow$ Tafel

Aufgabe 4: Backward-DP-Algorithmus

(c) Gegeben sei die folgende rekursive Definition der Pfadkostenfunktion:

$$V_k(i) = \begin{cases} g(i) & \text{if } k = 0 \\ 1 + V_{k-1}(i) & \text{else} \end{cases}.$$

Geben Sie die Aktionen an, die auf Basis dieser Kostenfunktion im ersten und im zweiten Entscheidungsschritt gewählt würden.

Aufgabe 4: Backward-DP-Algorithmus

(c) Gegeben sei die folgende rekursive Definition der Pfadkostenfunktion:

$$V_k(i) = \begin{cases} g(i) & \text{if } k = 0 \\ 1 + V_{k-1}(i) & \text{else} \end{cases} .$$

Geben Sie die Aktionen an, die auf Basis dieser Kostenfunktion im ersten und im zweiten Entscheidungsschritt gewählt würden.

Vorgehensweise:

- Ermittlung des Kostenfunktionen V_4 und V_3 gemäß obiger Vorschrift
- Bestimmung der ersten Aktion durch "gierige" Auswertung von V_5 nach

$$\pi_5(i) \in \arg\min_{a \in A(i)} \sum_{j=1}^{n} p_{ij}(a) \left(c(i,a,j) + V_4^*(j)\right)$$

für den Startzustand $i = i_0 = (0,0,0,0)$

Aufgabe 4: Backward-DP-Algorithmus

Vorgehensweise: (Forts.)

- Iteration über die Menge der möglichen Folgezustände $i_1 \in \{j \in S | p_{i_0 j}(\pi_5(i_0)) > 0\}$
  - Bestimmung der jeweils zugehörigen zweiten Aktionen durch "gierige" Auswertung von V_4 nach

$$a_1 = \pi_4(i) \in \arg\min_{a \in A(i)} \sum_{j=1}^{n} p_{ij}(a)\,(c(i,a,j) + V_3^*(j))$$

Aufgabe 4: Backward-DP-Algorithmus

Vorgehensweise: (Forts.)

- Iteration über die Menge der möglichen Folgezustände $i_1 \in \{j \in S | p_{i_0 j}(\pi_5(i_0)) > 0\}$
  - Bestimmung der jeweils zugehörigen zweiten Aktionen durch "gierige" Auswertung von V_4 nach

$$a_1 = \pi_4(i) \in \arg\min_{a \in A(i)} \sum_{j=1}^{n} p_{ij}(a) \left(c(i, a, j) + V_3^*(j) \right)$$

- Resultate: $\rightarrow$ Tafel

Projekte in Wissen 1

Zur Erinnerung:

- Parallel zum aktuellen Übungsblatt wird Ihnen eine Sammlung möglicher Vorschläge für ein von Ihnen im Laufe des Semesters umzusetzendes Projekt zur Verfügung gestellt. Im Rahmen dieses Projektes werden Sie einen der in der Vorlesung besprochenen Algorithmen implementieren und praktisch anwenden.
- Außerdem werden Sie ihr Projekt mit einer **Live-Demo** und kurzen **Präsentation** im Rahmen der mündlichen Prüfung zum Ende des Semesters vorstellen.
- Die Projekte werden nach einer FCFS-Strategie (First Come First Serve) vergeben und dürfen von Ihnen in Teams von bis zu zwei Personen bearbeitet werden.

Projekte in Wissen 1

Zur Erinnerung:

- Zusätzlich zu den von mir vorgeschlagenen Projekten dürfen Sie Ihrerseits auch gern eigenen Projektideen vorschlagen und nach Rücksprache mit mir als “Ihr Projekt” verwenden.
- Bitte beachten Sie, dass Sie sich bis spätestens vier Wochen vor Semesterende für ein Projektthema entschieden haben müssen.