STRUKTURELLE MODELLE IN DER BILDVERARBEITUNG
2. ÜBUNG – MARKOVSCHE KETTEN

Aufgabe 1. In einem Markovschen Modell für Zustandsfolgen y sei einer der Zustände $k^* \in K$ ausgezeichnet. Es interessiert, wieviel mal eine mit diesem Modell generierte Folge y diesen Zustand im Mittel enthält. Geben Sie einen effizientes Verfahren zur Berechnung dieses Mittelwerts an.
Hinweis: Nutzen Sie dabei den Fakt, dass der Erwartungswert einer Summe von zufälligen Größen gleich der Summe ihrer Erwartungswerte ist! Die Anzahl der Vorkommen von k^* in einer Folge y lässt sich offensichtlich in der Form

$$\delta_{y_1 k^*} + \delta_{y_2 k^*} + \ldots + \delta_{y_n k^*}$$

schreiben (δ ist das Kronecker-Symbol, d.h. $\delta_{ab} = 1$, falls $a = b$ und 0 sonst).

Aufgabe 2. Sei P die 2×2 Matrix der Übergangswahrscheinlichkeiten einer homogenen Markovschen Kette mit 2 Zuständen. Geben Sie hinreichende Bedingungen dafür an, dass die Abbildung P (d.h. $p_i = P \cdot p_{i-1}$) der Wahrscheinlichkeitsvektoren p kontrahierend ist.

Aufgabe 3. Ein Land besteht aus Städten, wobei manche Paare davon mit Straßen verbunden sind. Dabei gibt es verschiedene Typen von Straßen – Autobahnen, Landstraßen usw. Ein Fahrzeug fährt durch das Land. Am Anfang befindet es sich gleichwahrscheinlich in einer der Städte. In einem „elementaren Schritt“ wählt es zufällig eine Straße, die aus der aktuellen Stadt ausgeht, und folgt dieser. Dabei meldet das Fahrzeug, welchen Typ der Straße es gerade befährt. Seine Meldungen sind allerdings nicht exakt. So kann zum Beispiel das Fahrzeug mit einer bestimmten Wahrscheinlichkeit $p(\text{Autobahn}|\text{Landstraße})$ eine Autobahn melden, obwohl es gerade eine Landstraße benutzt.
a) Modellieren Sie die Fahrt des Fahrzeugs sowie die Folge der dabei entstehenden Meldungen mittels eines Markovschen Modells.
b) Gegeben sei die Folge der Meldungen des Fahrzeugs während einer Fahrt. Man finde die Folge der Städte, die das Fahrzeug dabei besucht hat.

Aufgabe 4. Gegeben sei ein Markovsches Modell der Sprache. Das a-priori Wissen des Modells besteht aus der Menge K der Phoneme, der Matrix der Übergangswahrscheinlichkeiten $p(k|k')$ und der Wahrscheinlichkeitsverteilung der Phoneme im ersten Zeitpunkt $p(y_1)$. Jedem Paar (y,x) (x ist die Beobachtung) ordnet das Modell die Wahrscheinlichkeit

$$p(y,x) = p(y_1)\prod_{i=2}^{n} p(y_i|y_{i-1}) \cdot \prod_{i=1}^{n} p(x_i|y_i)$$

zu. Weiterhin sei das folgende Wortmodell gegeben. Zu einem Wort A gehören alle Phonemenfolgen, die nur Phoneme aus der Menge $K_A \subset K$ enthalten.
a) Wie ergibt sich die a-priori Wahrscheinlichkeit des Wortes $p(A)$ in diesem Wahrscheinlichkeitsmodell?
b) Gegeben sei eine Beobachtung x. Wie ergibt sich die Wahrscheinlichkeit $p(A,x)$?
c) Sei für ein Wort A nicht nur die Menge der Phoneme K_A bekannt, sondern auch die Reihenfolge, in der die Phoneme vorkommen müssen. Wie ergeben sich in dem Fall die Wahrscheinlichkeiten $p(A)$ und $p(A,x)$?

Aufgabe 5. Man betrachte das Markovsche Modell der Sprache aus der vorigen Aufgabe. In der Menge der Phoneme gibt es dabei ein besonderes Phonem, welches als „Pause“ bezeichnet wird. Es wird für die Trennung von Worten verwendet.
Gegeben sei eine Beobachtung x. Zusätzlich sei bekannt, dass dabei genau zwei Worte ausgesprochen wurden. Also besitzt die „wahre“ Folge der Phoneme genau eine Pause. Wie lässt sich die Position der Pause bestimmen?