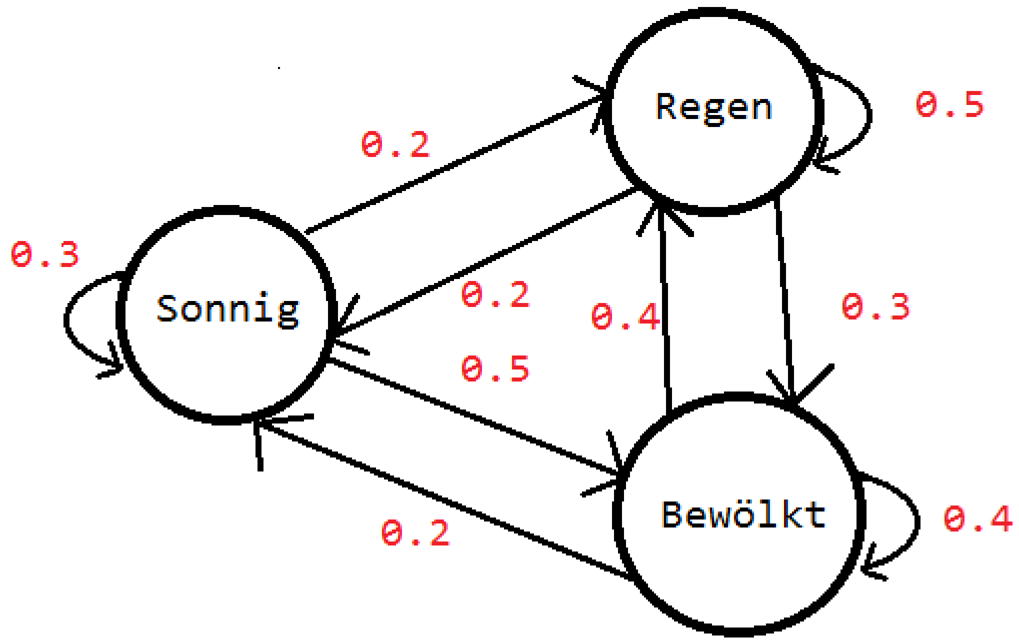
Zusammenfassung Tutorien der Woche 27.-31. 01. 2014 ALDABI

Markov-Ketten:

Viele Ereignisse schon in unserem Alltag beeinflussen sich gegenseitig, können also als „Ablauf" oder Kette von Ereignissen gesehen werden. So kommt zum Beispiel nach einem regnerischen Tag meist wieder nicht so gutes Wetter oder
Sollten diese Abfolgen eine Regelmäßigkeit haben, so kann man damit ein Modell erstellen und somit gewisse Sachverhalte vorhersagen oder prüfen.

Beispiel 1: Wetter:

Nach Beobachtung des Wetters im Herbst sind Sie zu folgendem Modell gekommen:



Offensichtlich kommen zur Zeit zwei sonnige Tage nur mit einer Wahrscheinlichkeit von 1/3 hintereinander, dagegen regnet es sehr häufig oder ist bewölkt.
Mit diesem Modell ergeben sich nun interessante Verwendungsmöglichkeiten:

Sie schauen aus dem Fenster, es ist sonnig. Sie wollen wissen, wie das Wetter wohl morgen am wahrscheinlichsten wird.
Offensichtlich ist es nach Modell morgen mit 50% Wahrscheinlichkeit bewölkt.

Um dem Wetter zu entfliehen, reisen Sie an einem strömenden Regentag für eine Woche ins sonnige Italien. Dort aber sehen Sie die ganze Woche keine Sonne. Zurück im gerade bewölkten Deutschland erzählt Ihnen ein Freund schadenfroh, es sie noch zwei Tage bewölkt gewesen und dann nur Sonne. Dies sei wieder der erste bewölkte Tag.
Seine Glaubwürdigkeit ist bei solchen Sticheleien gering, und nun können sie seiner Geschichte eine Wahrscheinlichkeit zuordnen:
Die angeblich beobachtete Wetterlage war: **RBBSSSSB**
Die Wahrscheinlichkeit für diesen Verlauf ist:
0.3*0.4*0.2*0.3*0.3*0.3*0.5 = 0.108%
ACHTUNG: Dies ist jetzt die Wahrscheinlichkeit dieser Kette unter ALLEN Ketten gleicher Länge.

Eine richtige Referenz fehlt vielleicht. Als Vergleich könnten Sie die Wahrscheinlichkeit bestimmen, dass es nur geregnet hat und dies der erste bewölkte Tag ist:
p(RRRRRRRB)= 0.625%

Auch stärker verkettete Fragen sind möglich:
Zwei Tage vor Ihrem Geburtstag sieht er Himmel sehr bewölkt aus.
Sie hatten eigentlich vorgehabt, zu Grillen. Nun ist die Frage: Wie wahrscheinlich ist übermorgen, dass es nicht regnet?

Für den Geburtstag kommen also **S** und **B** in Frage. Heute ist es **B**.
Die zu betrachtenden Pfade sind also:

p(BSS) = 0.06 p(BBS) = 0.08 p(BRS) = 0.08 p(BSB) = 0.1 p(BBB) = 0.16 p(BRB) = 0.12 p(Geb. Kein Regen) = 0.6	p(BSR) = 0.04 p(BBR) = 0.16 p(BRR) = 0.2 p(Geb. Regen) = 0.4

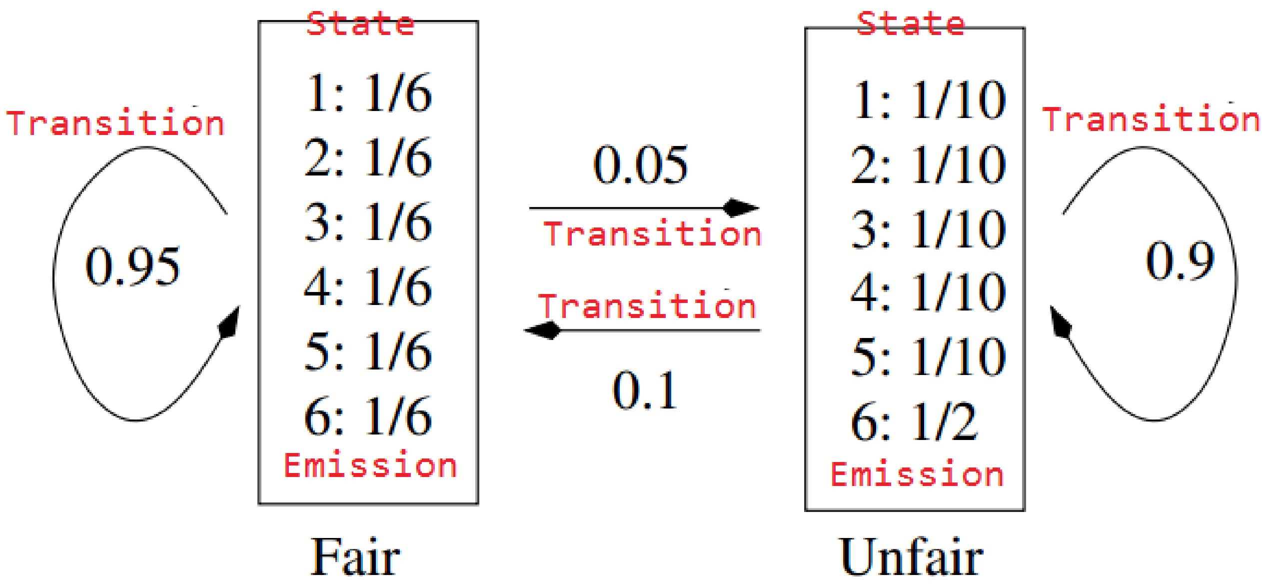
Wie es sein sollte, ergänzen sich die Wahrscheinlichkeiten von Ereignis und Gegenereignis zu 1.

Sie sollten nun verstanden haben, wie man ein Markov-Modell durläuft und eine Wahrscheinlichkeit bestimmt.

Hidden-Markov-Modells:

In manchen Fällen aber ist das, was man beobachten kann, nur ein „Indiz“ für den Zustand des Modells, der dahinter steht.
Als Beispiel hierfür wurde in der Vorlesung der gezinkte Würfel genannt. Diesen kann man nicht an seiner Form oder ähnlichem erkennen. Alles, was wir beobachten können, ist ein Würfel, der eine Zahl von 1 bis 6 anzeigt.
Dennoch verändert ein gezinkter Würfel natürlich das Würfelergebnis. Idealerweise kann man also nur durch die Würfelergebnisse bestimmen, ob und wann der gezinkte Würfel zum Einsatz kam.

Wie also sieht nun ein Hidden-Markov-Modell aus:



State (Hidden): Der Zustand, den zu erfahren meist das Ziel ist. In diesem Beispiel: Der Gegner kann einen Wurf mit dem **fairen** oder dem **gezinkten** Würfel machen.

Transition: Produktionen von einem Zustand zu einem neuen, wieder versehen mit Wahrscheinlichkeiten. „Wenn der Gegner den gezinkten Würfel benutzt, was benutzt er dann beim nächsten Wurf?“

Sigma: Alphabet (für Emissionen)

Emission: Auch auf deutsch „Emissionen“. Dies sind die Möglichkeiten, einen Buchstaben aus Sigma in einem bestimmten State zu emittieren. „Wie wahrscheinlich ist eine 3 mit gezinkten Würfel?“

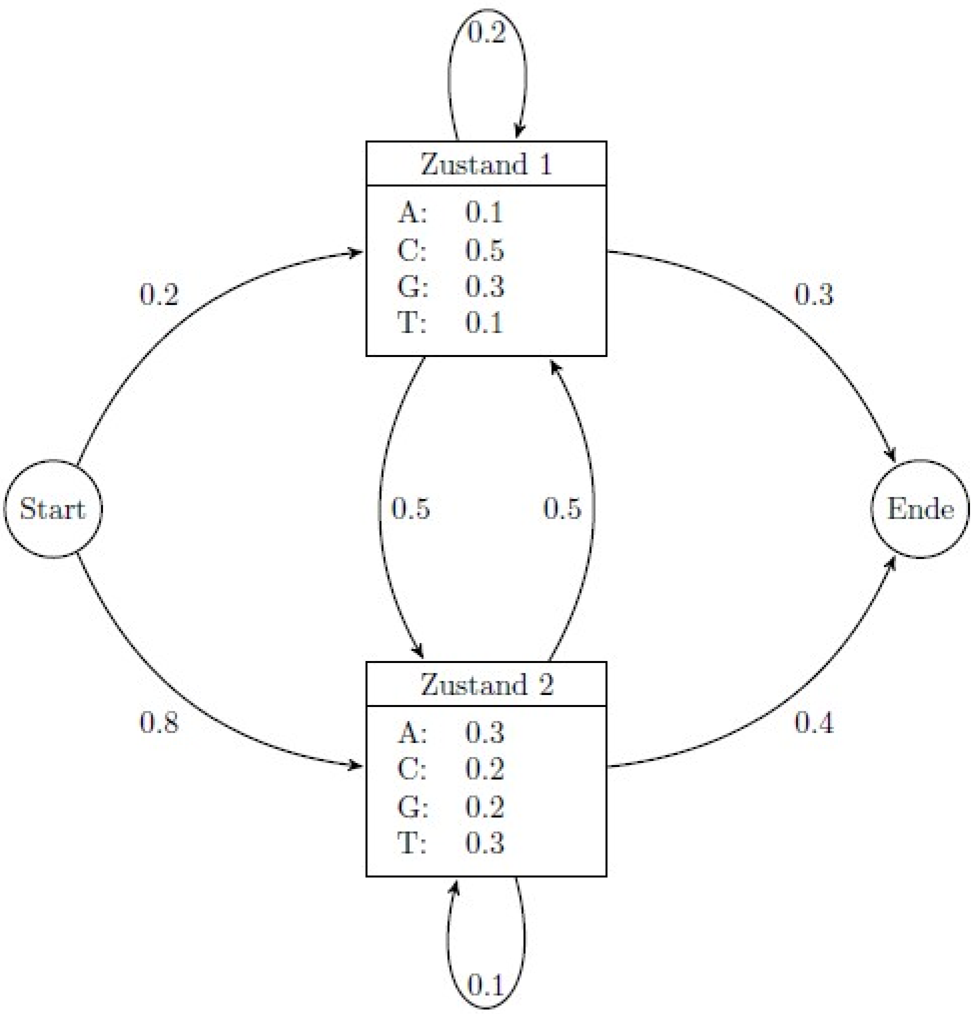
Selbst mit so einem gegebenen Modell steht man beim Bewerten einer Wahrscheinlichkeit vor dem Decodierungsproblem. In welchem State befinden man sich gerade?

Wenn man dies auch nicht mit Sicherheit sagen kann, so kann für jede Ausgabesequenz (nur die emittierten Zeichen) bestimmen, in welchem State sich das am wahrscheinlichsten befindet, gegeben ein Startzustand, der klärt, mit welchem Zustand es los geht.

Die Idee einer Rekursion und damit einer DP-Matrix ist geboren.
Um den wahrscheinlichsten Pfad für eine Sequenz x über ein HMM zu bestimmen, muss ich mir nur an jeder Stelle für jeden möglichen State anschauen, aus welchen States der HMM gekommen sein kann. Man nimmt dann den besten Weg und multipliziert dies noch mit der Wahrscheinlichkeit der Emission des gelesenen Zeichens.

Der Viterbi-Algorithmus an einem Beispiel:

Gegeben sie folgendes HMM für DNA, das nach zwei Zuständen unterschieden werden soll:



Der Viterbi-Algorithmus nach Vorlesung wir dann in folgener Tabelle ausgeführt:

		G	G	
Start/Ende				
Zustand 1				
Zustand 2				

siehe Bilderordner Viterbi.