

Sprachverarbeitung: Musterlösung zur Übung 21

Trainieren diskreter Hidden-Markov-Modelle

Aufgabe 1: Baum-Welch-Algorithmus für eine Beobachtungssequenz

Die Musterlösung der Funktion `[ns,nt,nb] = discr_baum_welch_alg(a,b,X)` ist im Directory `Uebung21/Loesung/` zu finden.

Aufgabe 2: Baum-Welch-Training mit mehreren Beobachtungssequenzen

Die Musterlösung der Funktion `[a,b] = discr_baum_welch_training(Xset,N,M,K)` ist im Directory `Uebung21/Loesung/` zu finden.

Aufgabe 3: Trainieren von Wort-DDHMM

Idealerweise wird die Anzahl Beobachtungssequenzen möglichst gross gewählt, ebenso die Anzahl Iterationen. Je mehr Trainingsmaterial vorhanden ist, desto genauer können daraus die HMM-Parameter ermittelt werden. Insbesondere wenn die Aufenthaltswahrscheinlichkeit in einem Zustand sehr klein ist, braucht es entsprechend mehr Trainingsmaterial um dessen Parameter möglichst korrekt zu schätzen (kleine Aufenthaltsdauer bedeutet wenig Beobachtungen innerhalb einer Sequenz). Es lässt sich leicht erkennen, dass für `ddhmm(1)` im Allgemeinen ein grösseres Trainingsset erforderlich ist als für `ddhmm(2)`, weil die Beobachtungssequenzen von `ddhmm(1)` kürzer sind.

Sehr schön lässt sich dieser Effekt auch an `ddhmm(3)` beobachten: Weil die Aufenthaltswahrscheinlichkeit im Zustand 3 viel grösser ist als in den andern emittierenden Zuständen (also 2 und 4), wird der Zustand 3 auch genauer trainiert.

Aufgabe 4: Testen der trainierten Wort-DDHMM

Die Summe der Erkennungsraten ist dann nicht 100%, wenn Beobachtungssequenzen vorkommen, für welche die Forward- und/oder die Viterbi-Wahrscheinlichkeiten aller HMM null sind. Dies ist in den Übungen 19 und 20 nicht vorgekommen, weil zur Erkennung das generierende Modell verwendet worden ist, und folglich sowohl die Forward- als auch die Viterbi-Wahrscheinlichkeit für diese Modelle grösser null sein müssen. Beim trainierten Modell kann es hingegen vorkommen, dass Elemente der b-Matrix null sind, die beim generierenden grösser null sind. Um solche Fälle in der Praxis der Spracherkennung zu vermeiden, wird verhindert, dass beim Training Elemente der b-Matrix exakt null werden.

Dass für die trainierten Modelle `ddhmm(1)`, `ddhmm(3)` und `ddhmm(4)` die Forward- und die Viterbi-Wahrscheinlichkeit unterschiedlich sein können ist darauf zurückzuführen, dass bei den trainierten Modellen Laute verwechselbar sind, die bei den generierenden Modellen nicht verwechselt werden können (keine gemeinsamen Beobachtungen). Dies ist an den b-Matrizen ersichtlich.