

4. Aufgabenblatt zum Grundstudiums-Softwarepraktikum „Muster- und Bilderkennung“

Maja Popović, Björn Hoffmeister
{popovic, hoffmeister}@informatik.rwth-aachen.de

Generierung von Präfixbäumen

Das Ziel der Spracherkennung besteht darin, aus dem akustischen Signal der gesprochenen Sprache die Wortfolge zu ermitteln, die mit der größten Wahrscheinlichkeit gesprochen worden ist.

Hierbei werden die einzelnen Wörter in der Regel nicht als geschlossene Einheiten modelliert, sondern als Folge von unterschiedlichen Lauten, sogenannten Phonemen, betrachtet. Ein Phonem ist definiert als die kleinste bedeutungsunterscheidende Lauteinheit einer Sprache. Mögliche Kriterien der Phonemeinteilung:

- Art der Artikulation (Frikative, Plosive);
- Ort der Artikulation (labial, dental);
- Konsonanten und Vokale;
- Stimmlose und stimmhafte Laute;
- Stationäre und nichtstationäre Laute.

In der englischen Sprache kann man ca. 40-50 unterschiedliche Phoneme unterscheiden. Jedem Wort wird nun über das sogenannte *Aussprachelexikon* eine Phonemfolge zugewiesen, die die Lautfolge, die bei der Aussprache dieses Wortes produziert wird, möglichst gut beschreibt.

Beispiel:

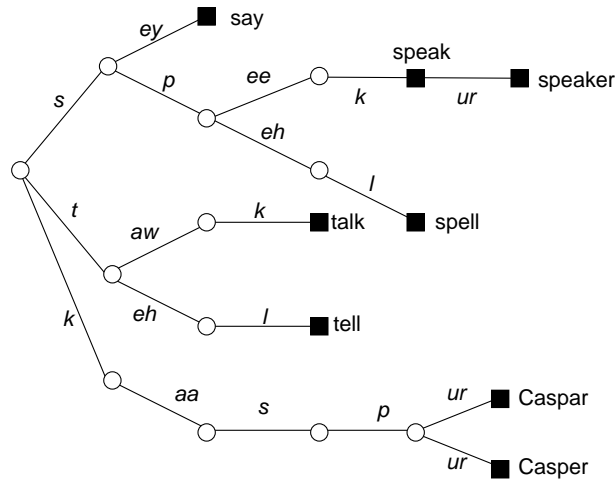
say	s	ey			
speak	s	p	ee	k	
speaker	s	p	ee	k	ur
spell	s	p	eh	l	
talk	t	aw	k		
tell	t	eh	l		
Caspar	k	aa	s	p	ur
Casper	k	aa	s	p	ur

lineares Aussprachelexikon

Jedes dieser Phoneme wird nun durch ein eigenes statistisches Modell beschrieben.

Das Modell eines Wortes erhält man nun durch die Aneinanderkettung der Modelle der einzelnen Phoneme, aus denen dieses Wort laut Aussprachelexikon besteht.

Bei zunehmendem Umfang des Aussprachelexikons nimmt die Anzahl der Wörter mit gleichem *Präfix*, d.h. gleicher Startfolge von Phonemen zu. Es ist daher möglich, das Aussprachelexikon durch eine Baumstruktur, den sogenannten *lexikalischen Präfixbaum* darzustellen, die sehr viel kompakter ist als die oben erläuterte Darstellung als *lineares* Lexikon.



Lexikalischer Präfixbaum

Im Verlauf der Spracherkennung müssen nun die Modelle aller Worte, die im Aussprachelexikon enthalten sind, parallel dahingehend bewertet werden, wie gut sie mit dem vorliegenden akustischen Signal korrelieren, um dann eine Entscheidung über den Satz treffen zu können, der mit der größten Wahrscheinlichkeit gesprochen worden ist. Im Falle einer linearen Darstellung des Aussprachelexikons müssen gleiche Präfixe mehrfach ausgewertet werden. Beim Baumlexikon dagegen werden gleiche Präfixe zusammengefaßt und müssen daher nur einmal bewertet werden. Somit ist mit Hilfe eines lexikalischen Präfix-Baumes eine erhebliche Reduktion des Rechen- und Speicheraufwandes bei der automatischen Spracherkennung möglich.

1. Implementieren Sie ein Programm, das das vorliegende lineare Lexikon in ein Baumlexikon überführt. Achten Sie dabei darauf, das Homophone (Wörter mit gleicher Phonemtransskription) so behandelt werden, daß sie innerhalb des Baumlexikons eindeutig dargestellt werden können.
2. Lesen Sie mit Ihrem Programm das Phoneminventar aus <http://www-i6.informatik.rwth-aachen.de/HTML/Lehre/PraGru/aufgabe4/PhonemeList.txt> und das lineare Lexikon aus <http://www-i6.informatik.rwth-aachen.de/HTML/Lehre/PraGru/aufgabe4/LinLex.txt> und berechnen Sie das entsprechende Baumlexikon. Die Ausgabe Ihres Programms sollte *genau* so aussehen, wie in der Datei <http://www-i6.informatik.rwth-aachen.de/HTML/Lehre/PraGru/aufgabe4/TreeLex.txt> dargestellt. Achten Sie insbesondere darauf, daß die Statistik über die Anzahl der Äste in den Generationen des Baumes mit den Angaben in der genannten Datei übereinstimmen.
3. Berechnen Sie den Kompressionsfaktor, der mit Hilfe des Baumlexikons erreicht wird:

$$\text{Kompressionsfaktor} = \frac{\text{Anzahl Kanten lin. Lex.}}{\text{Anzahl Kanten Baum-Lex.}}$$

Abgabe:

Abgabe (per email): bis Mittwoch, 17. Mai, 18h
 Abnahme (im Rechnerpool): Donnerstag, 18. Mai