Lehrstuhl Informatik 6
RWTH Aachen University
Prof. Dr.-Ing. H. Ney

Aachen, 01. 06. 2018

R. Schlüter,
K. Irie

# Automatic Speech Recognition

## 2. Exercise

---

In this exercise, we will review word conditioned lexical prefix tree search using across-word modeling, language model look-ahead, as well as time conditioned search. The solutions for this exercise sheet are to be submitted on 08. 06. 2016 at the beginning of the exercise session.

**Task 2.1** Tree Search, Across-Word Modeling and Time conditioned search

We assume a pronunciation lexicon with 4 words {*da, D, DA, DE*} and a set of phonemes {d, a, e} with the pronunciation dictionary as follows:

| word | pronunciation |
|---|---|
| *da* | d a |
| *D* | d e |
| *DA* | d e a |
| *DE* | d e e |

In this exercise, you are asked to draw search networks for this lexicon under different model assumptions and search approaches. In order to clarify the form of solutions which is expected for the following questions, examples for questions (a), (b), and (c) are provided at the end of this sheet. Please use the same convention for nodes and arcs to answer the following questions.

For questions (a), (b), and (c) we assume a unigram language model.

(a) Draw a lexical prefix tree search network using monophones. (1)

(b) Repeat (a) using *within-word triphones*. Use symbol '#' for word boundary context. (1)

(c) Repeat (a) using *across-word triphones*. Use symbol '$' for silence context. (2)

(d) Repeat (b) using a bigram language model. (3)

(e) Repeat (c) using a bigram langugage model, but draw the network only for the bigram context *da*. (3)

(f) We now assume the following bigram distribution in the context of word *da*:

$$p(w|da) = \begin{cases} 0.1 & w = da \\ 0.2 & w = D \\ 0.3 & w = DA \\ 0.4 & w = DE \end{cases}$$

Push/distribute the language model probabilities through the network of (d) and (e). This means that a language model look-ahead factored score has to be computed and written on each acoustic arc in the network, like in the left figure on p. 202 of the lecture notes. You are not asked to draw the compressed language model look-ahead tree, though. Again, for (e) please do this only for the bigram context *da*. (4)

(g) Consider time conditioned search with and without across-word context. Adapt the network of (b) and (c) to time conditioned search. For this, please confirm p. 278 of the lecture notes. Please note that time conditioned search assumes unique tree roots and make sure that the word conditioned body of the search network in (c) has multiple subtrees, each with a unique root state. (6)

Examples with a simple lexicon:

| word | pronunciation |
|---|---|
| X | a |
| Y | a a |