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

Aachen, 22. 06. 2018

R. Schlüter,
K. Irie,
A. Zeyer

Automatic Speech Recognition

5. Exercise

Submission Deadline: 05. 07. 2018 at the beginning of the exercise session. It is a Thursday.
IMPORTANT: Any electronic submission should be sent to:
`zeyer@i6.informatik.rwth-aachen.de`
with a copy (cc.) to `irie@i6.informatik.rwth-aachen.de` .

Task 5.1 Derivation of minimum classification error (MCE) criterion

Consider the minimum classification error (MCE) criterion given in the lecture notes pp. 484:

$$F(\theta) = \frac{1}{R} \sum_{r=1}^{R} \frac{1}{1 + \Big[\frac{p_\theta^\alpha(X_r|W_r) \cdot p^\alpha(W_r)}{\sum_{W \neq W_r} p_\theta^\alpha(X_r|W) \cdot p^\alpha(W)} \Big]^{2\varrho}}$$

where α and ϱ are smoothing hyper-parameters (that we assume constant).

Derive the expression of the derivative of this function with respect to the parameters θ as a function of $\nabla_\theta \log p_\theta(x_{rt}|s)$ (therefore, there is no need for replacing $p_\theta(x_{rt}|s)$ by any explicit expression). The examples of such a derivation can be found on pp. 498-503 of the lecture notes for the maximum mutual information (MMI) and on pp. 504-508 for the minimum phone error (MPE). Explain the efficient computation of statistics (*hint*: the sum in the denominator can be rewritten as the full sum minus the contribution from W_r). (10 P)

Task 5.2 Equivalence between Log-linear and Gaussian Models

Prove the equivalence between the Gaussian and Log-linear model-based class posteriors using the log-linear to Gaussian parameter transformation introduced on Slide 521, i.e. derive explicitly, that the Gaussian-based class posterior probability with the shown parameter setting equals the log-linear model-based class posterior probability. (10 P)