Hermann Ney,
RWTH Aachen University

On Architectural Issues of Neural Networks in Speech Recognition

Recently, artificial neural networks (ANN) were able to improve the performance of speech recognition systems dramatically. There have been more than 25 years of extensive research on neural networks in speech recognition. Despite this huge effort, there are a number of open issues concerning the architecture of ANN based systems for speech recognition. Examples of such issues are: 1) Unlike the hybrid approach of replacing the emission probability function by an ANN, there is the possibility of a direct approach that models the posterior state sequence of (phonetic) labels directly without using the generative concepts of classical hidden Markov models (HMM). 2) In the CTC approach (connectionist temporal classification), the HMM is simplified by using a single label per phoneme (or character in handwriting recognition) only. The CTC training criterion is the sum over all possible posterior distributions of label sequences. 3) Recently there have been so-called attention based approaches that replace the conventional HMM formalism by a recurrent neural network. In these three cases, we are faced with the questions of how these ANN based approaches compare with the conventional discriminative framework of hybrid HMMs. We will discuss the advantages and disadvantages of these approaches in more detail and compare them with conventional hybrid HMMs.

Hermann Ney is a full professor of computer science at *RWTH Aachen University*, Germany. His main research interests lie in the area of statistical classification, machine learning and human language technology and specific applications to speech recognition, machine translation and handwriting recognition. In particular, he has worked on dynamic programming and discriminative training for speech recognition, on language modeling and on phrase-based approaches to machine translation. His work has resulted in more than 700 conference and journal papers (h-index 83, 35000 citations; estimated using Google scholar). He and his team contributed to a large number of European (e.g. TC-STAR, QUAERO, TRANSLECTURES, EU-BRIDGE) and American (e.g. GALE, BOLT, BABEL) joint projects. Hermann Ney is a fellow of both *IEEE* and *ISCA* (*International Speech Communication Association*). In 2005, he was the recipient of the *Technical Achievement Award* of the I*EEE Signal Processing Society*. In 2010, he was awarded a senior DIGITEO chair at *LIMIS/CNRS* in Paris, France. In 2013, he received the award of honor of the *International Association for Machine Translation*. In 2016, he was awarded an ERC advanced grant.