

# Optimization of Large Scale Neural Networks for Speech Recognition

Michiel Bacchiani<sup>1,a)</sup>

***Abstract:***

Recent years have shown a large scale adoption of speech recognition by the public, in particular around mobile devices. Simultaneously, the state of the art of speech recognition has improved significantly with the adoption of deep neural networks as the key technology for acoustic modeling. This approach is particularly effective when applied at scale. This talk will first provide an overview of how the Google speech group has grown over the years. Starting from a research effort, it has developed into an integral input modality of the Android operating system and is launched in more than 50 languages worldwide. The presentation will provide some historical perspective of that growth, discussing earlier products and outline the current mobile products and future directions of those. The success and adoption of the Google speech technology is evident by the decade of speech our systems receives each day. The talk will then focus on recent advances in training algorithms and infrastructure to train large neural networks on large data sets. Although not exclusively, the talk will provide a Google perspective to approaching this problem. It will describe our asynchronous parallel training infrastructure and how key algorithmic improvements fit within that framework. Particular attention will be paid to optimization using a sequence objective and recurrent network architectures (using Long Short Term Memory models).

---

<sup>1</sup> Google Inc., 76 Ninth Ave, New York, NY, 10011, USA

<sup>a)</sup> michiel@google.com