LAS Thesis project kickoff meeting

Moritz Wolter

September 23, 2016

Outline

Project Overview

- Transcribe speech utterances to characters.
- Use a listen attend and spell (LAS) model to do this.
- Train model components jointly.

The LAS-Architecture

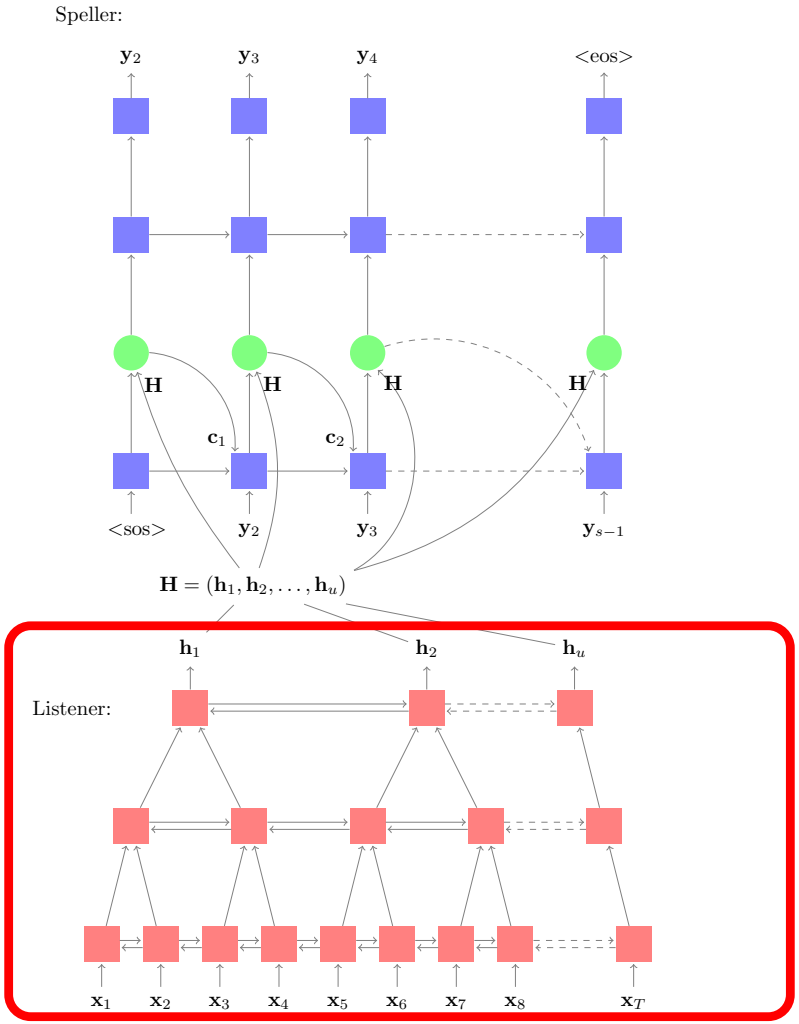

Figure 1 : The LAS architecture

Debugging The Listener

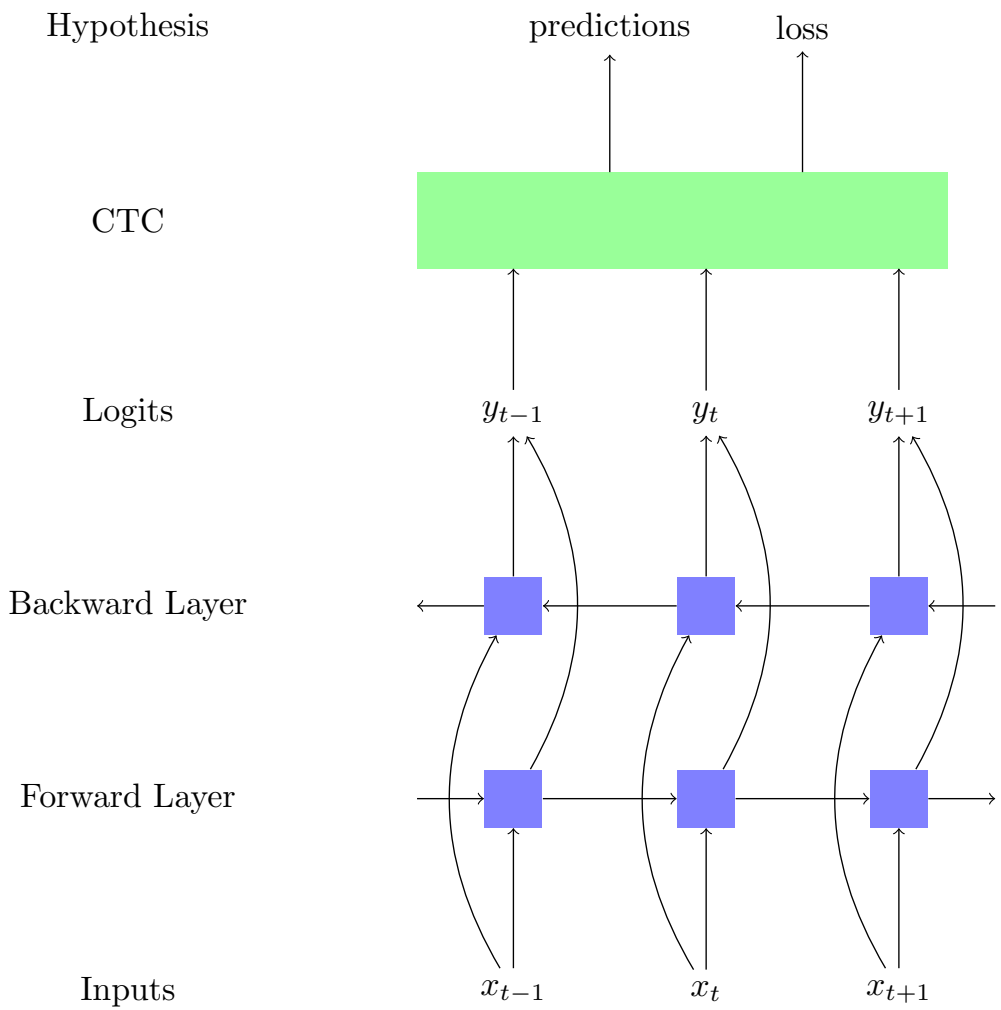

Figure 2 : BLSTM CTC schematic

First Results on Timit

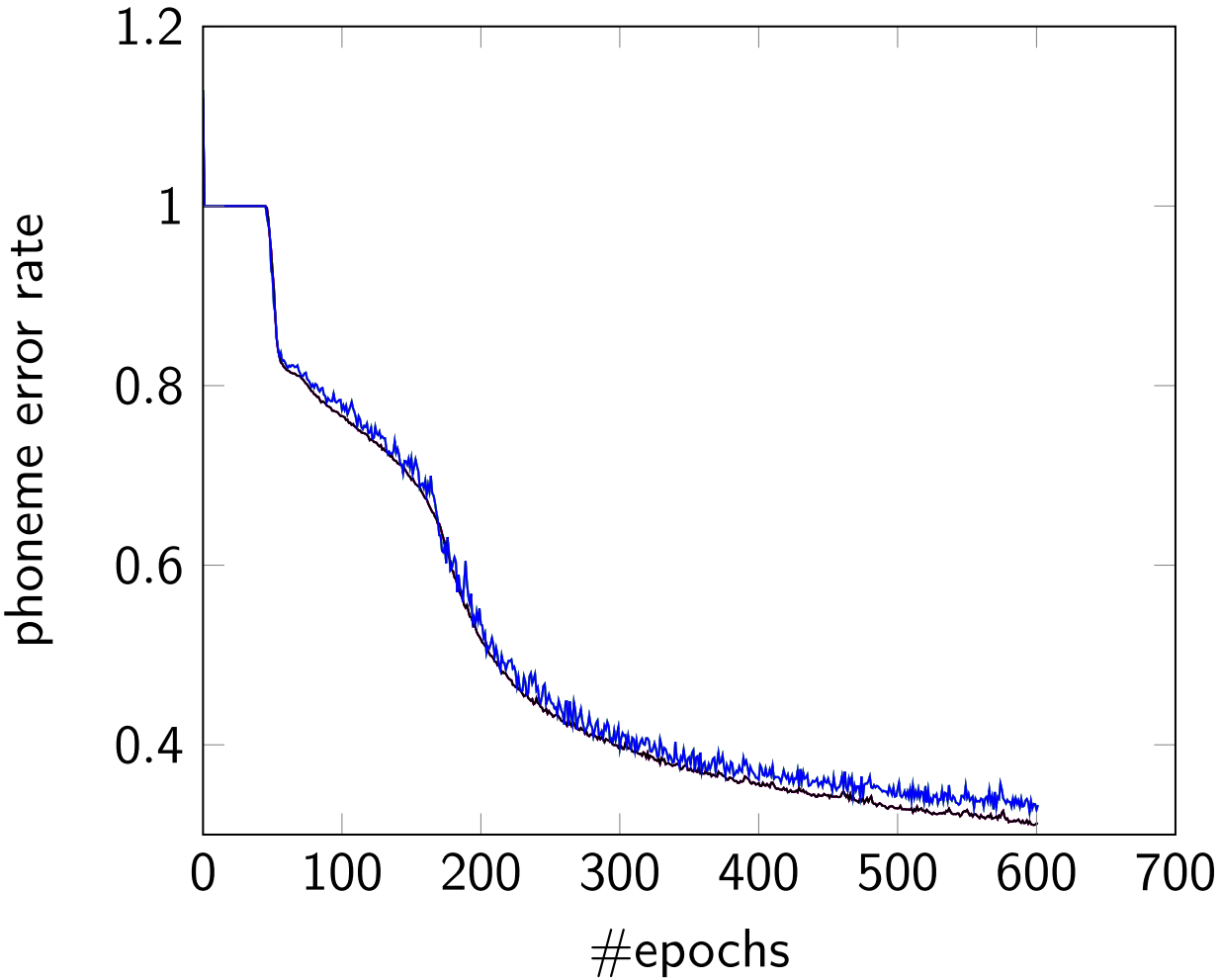

Figure 3 : BLSTM-CTC training and validation error on timit.

What happend so far? What will happen next?

- What happend so far?
  1. Listener with CTC on Timit .
- What will happen next?
  1. Switch to Aurora4.
  2. Test the CTC-listener on Aurora4.
  3. Add attention based spelling to the listener.
  4. Decoding with beam search.
- After that you take over!
  1. Test the las skeleton on LibriSpeech.

Questions

Thank you for your attention. Questions?
moritzalexander.wolter@student.kuleuven.be
or come and meet me in our LAS-room (02.88).