LAS Thesis project kickoff meeting

Moritz Wolter

September 22, 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

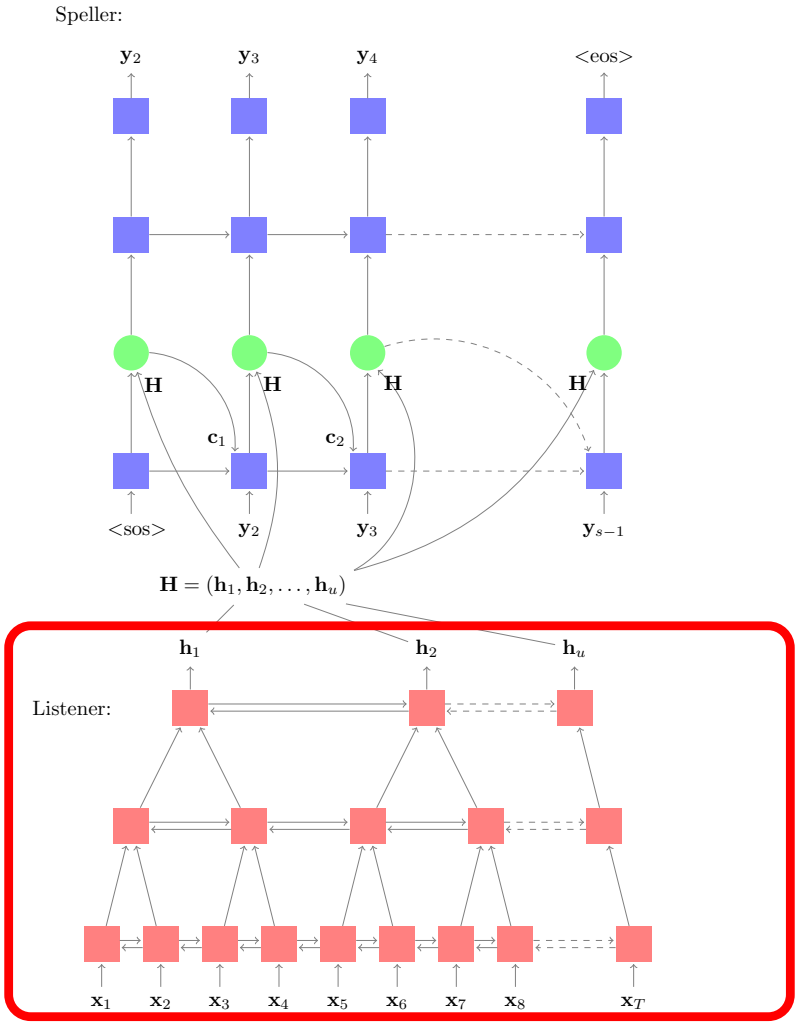

Figure 1 : The LAS architecture

What is Tensorflow?

- "TensorFlow is an interface for expressing machine learning algoritms and an implementation for executing such algorithms. "[1]
- Computations are described by directed graphs.
- Data-Tensors flow along graph edges.
- Graphs are constructed using user specified elementary operations.
- Computations are started by requesting certain values, which leads to (partial) evaluation of the graph.

[1] TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, Abadi et al, Google Research.

Tensorflow

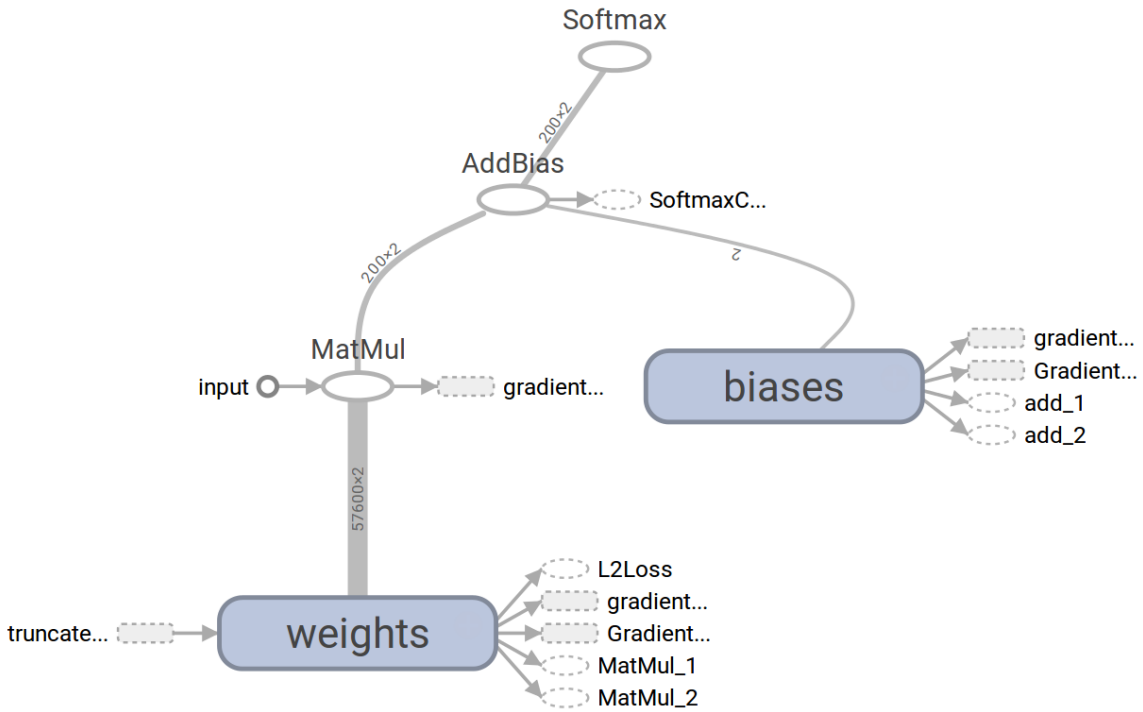

Figure 2 : A simple linear node in tensorboard

What happend so far? What will happen next?

- What happend so far?
  1. Literature overview.
  2. Learned how to use tensorflow.
  3. Looked into timit and aurora4.
- What will happen next?
  1. Finish implementing a skeleton LAS on Timit.
  2. Decoding with beam search.
  3. Port to aurora4.

Summary and Questions

The presentation covered:

- Input feature generation.
- The LSTM building block.
- A LAS-Architecture overview.
- The tensorflow toolbox.
- The plan.

Thank you for your attention. Questions?

moritzalexander.wolter@student.kuleuven.be