

## Deep Learning Assignment 3 10 points

In this assignment you will be answering some questions about recurrent network architectures and word embeddings, and running a simple LSTM implemented in TensorFlow.

### 1. Reber Grammar and Extended Reber Grammar

Read the description of the Reber Grammar and Extended Reber Grammar on Christian Herta's page:

<http://christianherta.de/lehre/dataScience/machineLearning/neuralNetworks/reberGrammar.php>

a. [1 mark]

Which of the following strings are in the language specified by the Reber Grammar (i.e. they can be generated by the automaton)? Which are not?

- (i) BPVPSE
- (ii) BTSXXTPVXXVPSE
- (iii) BTSSXXTPVXTTPXVVE

b. [1 mark]

For each of these strings of the Extended Reber Grammar, what would be the correct prediction for the next letter?

- (i) BPBPVVE
- (ii) BTBTSSXXTTVPSE

2. [3 marks] Briefly describe the problem of long range dependencies, and comment on the ability of each of the following architectures to learn long range dependencies:

NETtalk, Simple Recurrent Networks (SRN), Long Short Term Memory (LSTM), Gated Recurrent Units (GRU).

3. [3 marks] Briefly explain why full softmax may not be computationally feasible for language processing tasks. Name, and briefly describe, two alternatives to full softmax which are computationally feasible.

4. [2 marks] Understanding LSTM in TensorFlow using MNIST

Read through these two blog posts:

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

<https://jasdeep06.github.io/posts/Understanding-LSTM-in-Tensorflow-MNIST/>

The code from the jasdeep06 blog is copied in tf\_lstm.py

Download this code and run it by typing

```
python tf_lstm.py
```

Copy the final Training Accuracy, Loss and Testing Accuracy into your answers.

Increase the number of iterations from 800 to 2000, by changing this line

```
while iter<800:
```

to this:

```
while iter<2000:
```

Run the code again,  
and copy the final Training Accuracy, Loss and Testing Accuracy into your answers.

--

Optional Supplementary Exercise (not for credit):

Work through this tutorial on LSTM Networks for Word Prediction:

<https://www.tensorflow.org/tutorials/recurrent>