Homework 6

Zhiyuan Hu

June 20, 2018

1 Task Introduction

1.1 The data:

- Input is a sequece of binary number, time t, input is 0 or 1, both are 50% probability. Like $X = [1, 1, 0, 0, 1, ...]$.

- Output Y at time t is also 0 or 1 with 50% probability. But if $X_{t-3} = 1$, probability of $Y_t = 1$ increase 50%. if X_{t-8} is 1, probability of $Y_t = 1$ decreases 25%.

1.2 Loss:

Use cross entropy to train the function.

1.3 Reference:

Code Link is here. Github Link
Paper Link here here here

2 Network

Like as is shown in the figure1 and figure 2:

3 Implementation in Tensorflow

Justlike as is shown in figure 3:

$$S_t = tanh(W(X_t \bigoplus S_{t-1})) + b_s$$

$$P_t = softmax(US_t + b_p)$$

$$W \in R^{d \times (2+d)}, b_s \in R^d, U \in R^{2 \times d}, b_p \in R^2$$
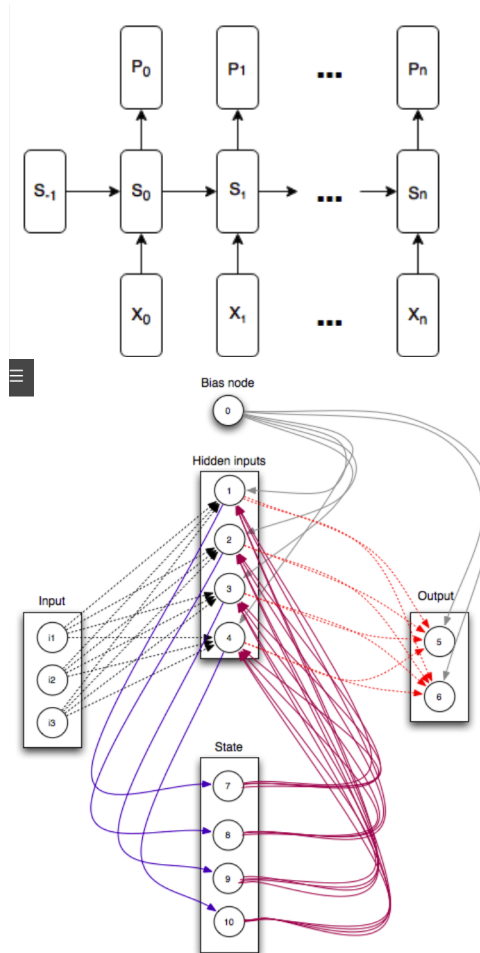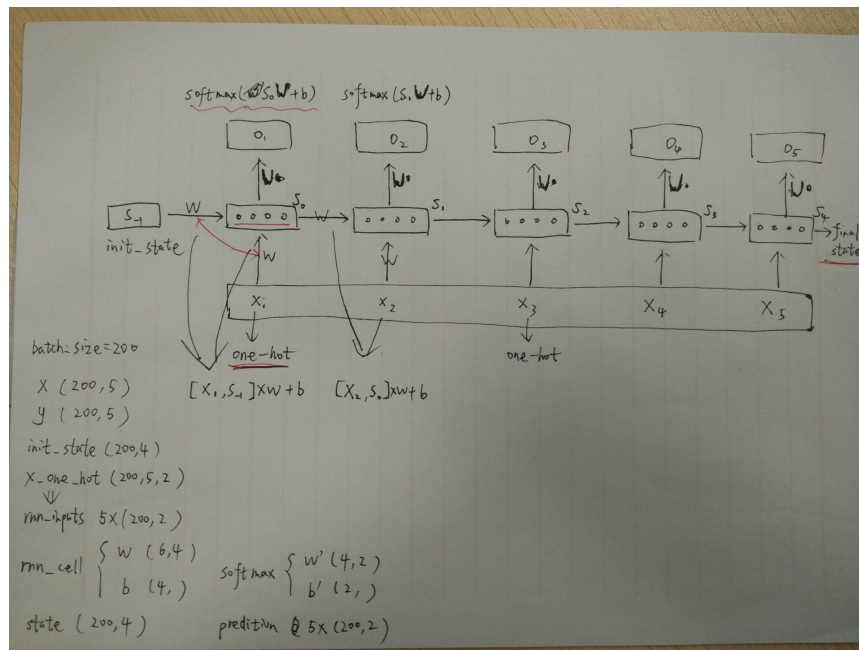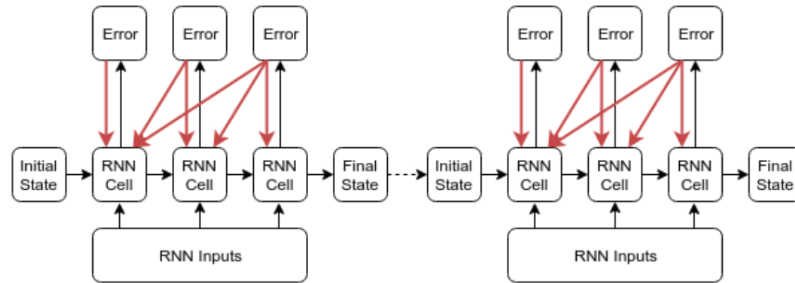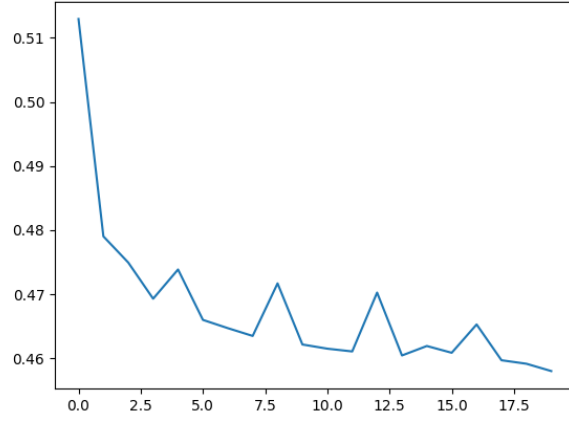
Figure 1:

Figure 2:

Figure 3:

Figure 4:

4 Result

Like as is shown in figure 4. with num_step 10 and state 16, finally the
network has learned the dependencies and loss is nearby 0.45.