L 14: 27.06.2022

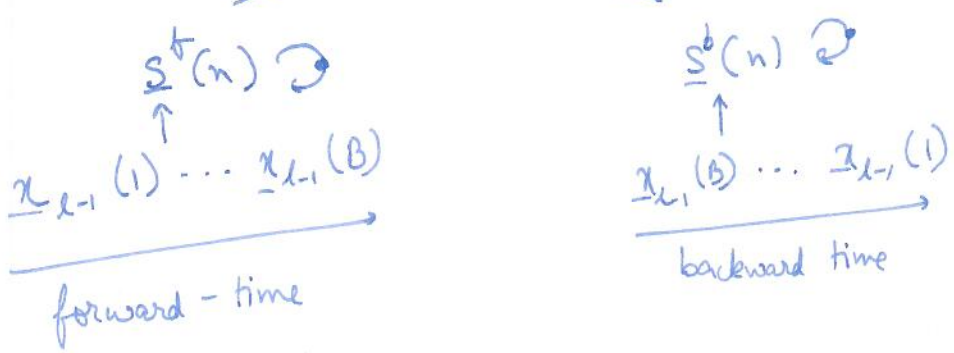
8.2 Bidirectional Recurrent Neural Network

Causal and non-causal systems

RNN is a causal system as it is dependent on past and present.

8-12

$$\underline{x}_l(n) = \begin{bmatrix} \underline{s}_l^f(n) \\ \underline{s}_l^b(B-n) \end{bmatrix} \qquad 1 \le n \le B$$



~~Q How is the above eqn possible due to~~

8-13

Training: backpropagation of derivatives through the unfolded bidirectional graph.

8.3 Long Short Term Memory (LSTM)

RNN is more difficult to train than CNN due to the exploding/vanishing gradient

DNN in Ch 5 and 7: large L = # layers
→ long chain rule over layers

RNN : large B = # time recursions
→ long chain rule over time in addition

Difficult choice of B :

B ↓ : noisy gradient, short memory

B ↑ : vanishing / exploding gradient

Solution : LSTM - special RNN

8-14

forget gate - clear the memory for next time
input gate - to overwrite the input or not
output gate - to control if the output is to be obtained or not

~~8-15~~

→ calculated by machine

8-15

why sigmoid $\sigma \in (0,1)$ → so 0 → gate closed
1 → gate opened

8-16

There are in total 7 operations in each LSTM cell/layer.

Effects of gating

$0 <$ gate signal < 1

≈ 0 close gate clear memory

≈ 1 open gate keep memory

input/output gate forgetgate

Slide 8-17

8-18

8-19

8-20

⋮

8-23

9 Autoencoders and generative models

9.1 Autoencoder

9-1
9-2
9-3
9-4

Autoencoder (AE):

an unsupervised learned NN to learn an efficient representation / code of the input.