

## Lecture 8

GRU

update gate  $z_t = \sigma(\omega_z \underline{h_{t-1}} + u_z \underline{x_t} + b_z)$

reset gate  $r_t = \sigma(\omega_r \underline{h_{t-1}} + u_r \underline{x_t} + b_r)$

$$h'_t = \tanh(u_h x_t + r_t \odot \omega_h \underline{h_{t-1}} + b_h)$$

$$\underline{h_t} = z_t \odot \underline{h_{t-1}} + (1 - z_t) \odot h'_t$$

$x_t$  - input,  $h_t$  - hidden state, related to output  
 $1 \rightarrow D$   $2 \rightarrow D$

$$\begin{array}{c}
 z_t \odot h_{t-1} \\
 \begin{array}{cc}
 \downarrow & \downarrow \\
 20 \times 1 & 20 \times 1
 \end{array}
 \end{array}
 \quad
 \begin{array}{c}
 z_t = \sigma(\omega_z \underline{h_{t-1}} + u_z \underline{x_t} + b_z) \\
 \begin{array}{ccccc}
 2 \times 1 & 20 \times 20 & 20 \times 1 & 20 \times 1 (1 \times 1) & 1 \\
 \underbrace{\hspace{1.5cm}} & \underbrace{\hspace{1.5cm}} & & & \\
 20 \times 1 & 20 \times 1 & & & 20 \times 1
 \end{array}
 \end{array}$$

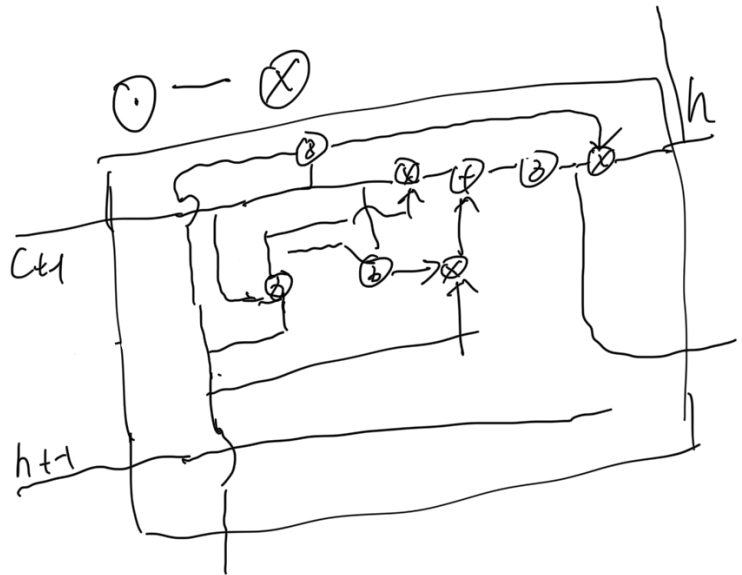
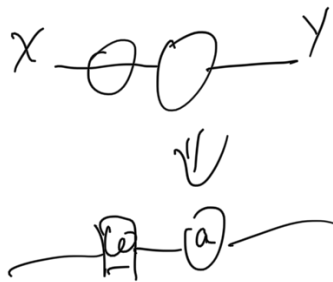
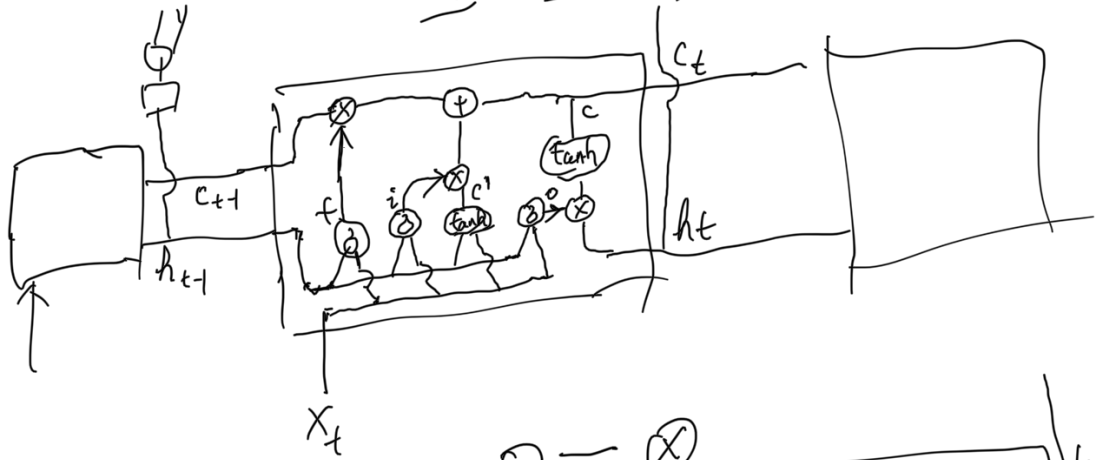
LSTM

$$\begin{array}{l}
 \text{forget gate} \\
 \text{input/update gate} \\
 \text{output gate} \\
 \text{cell input/state}
 \end{array}
 \left\{
 \begin{array}{l}
 f_t = \sigma(\omega_f h_{t-1} + u_f x_t + b_f) \\
 i_t = \sigma(\omega_i h_{t-1} + u_i x_t + b_i) \\
 o_t = \sigma(\omega_o h_{t-1} + u_o x_t + b_o) \\
 c'_t = \tanh(\omega_c h_{t-1} + u_c x_t + b_c)
 \end{array}
 \right.$$

$$C_t = f_t \odot C_{t-1} + i_t \odot C'_t$$

hidden state

$$h_t = O_t \odot \tanh(C_t)$$

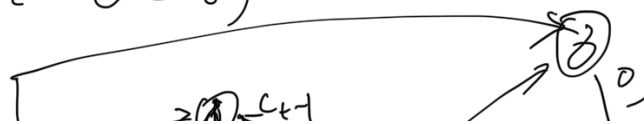


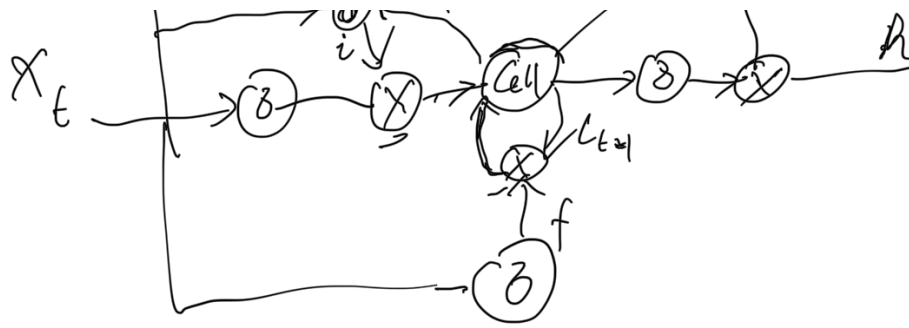
peephole LSTM

$$\begin{cases} f_t = \sigma(W_f C_{t-1} + U_f x_t + b_f) \\ i_t = \sigma(W_i C_{t-1} + U_i x_t + b_i) \\ O_t = \sigma(W_o C_{t-1} + U_o x_t + b_o) \end{cases}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot (U_c x_t + b_c)$$

$$h_t = O_t \odot \tanh(C_t)$$





peephole convolutional LSTM