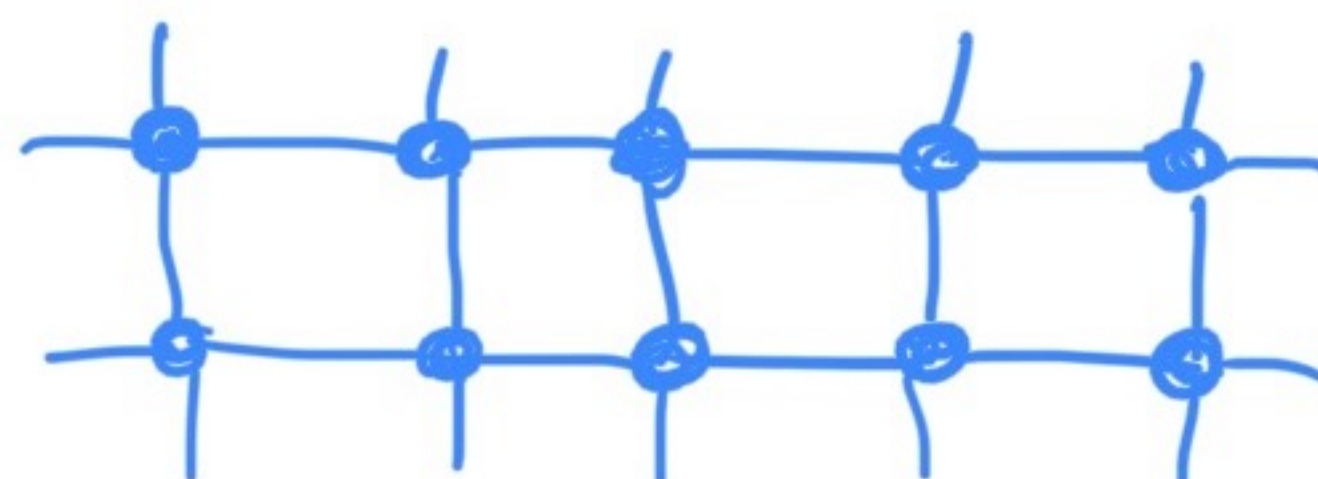


# LSTM

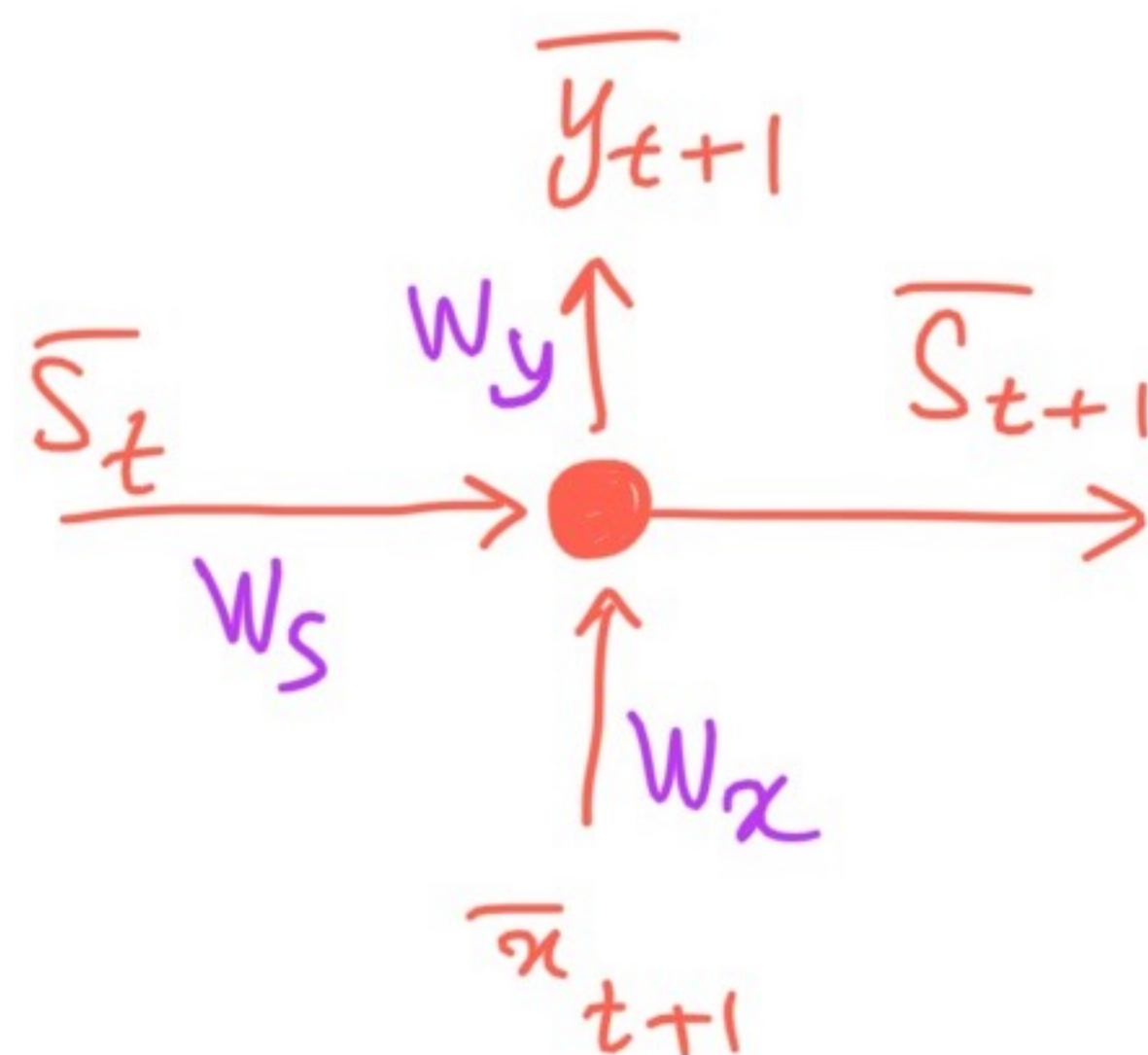
or long short-term memory cell

let's compare LSTM and RNN

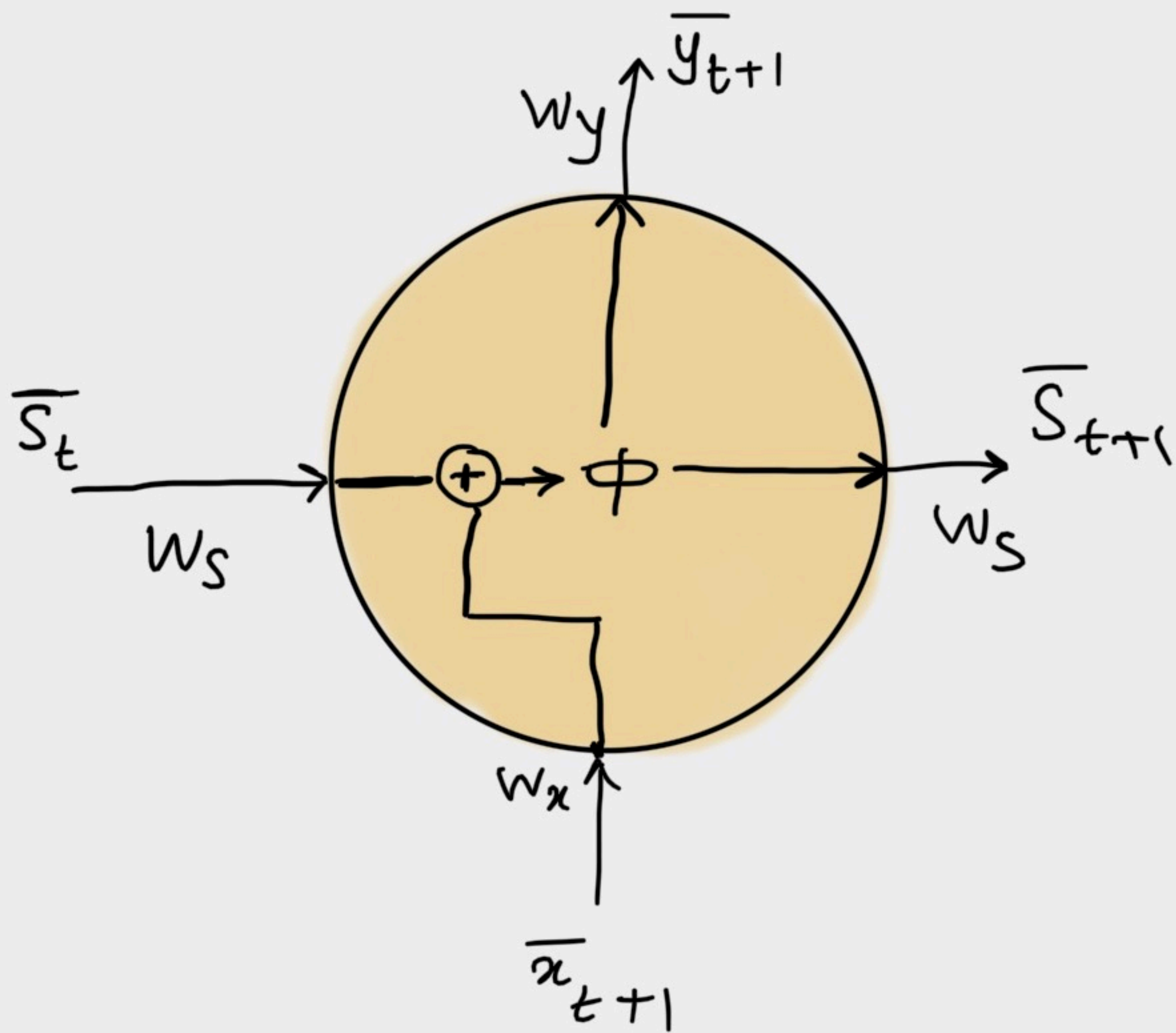
RNN



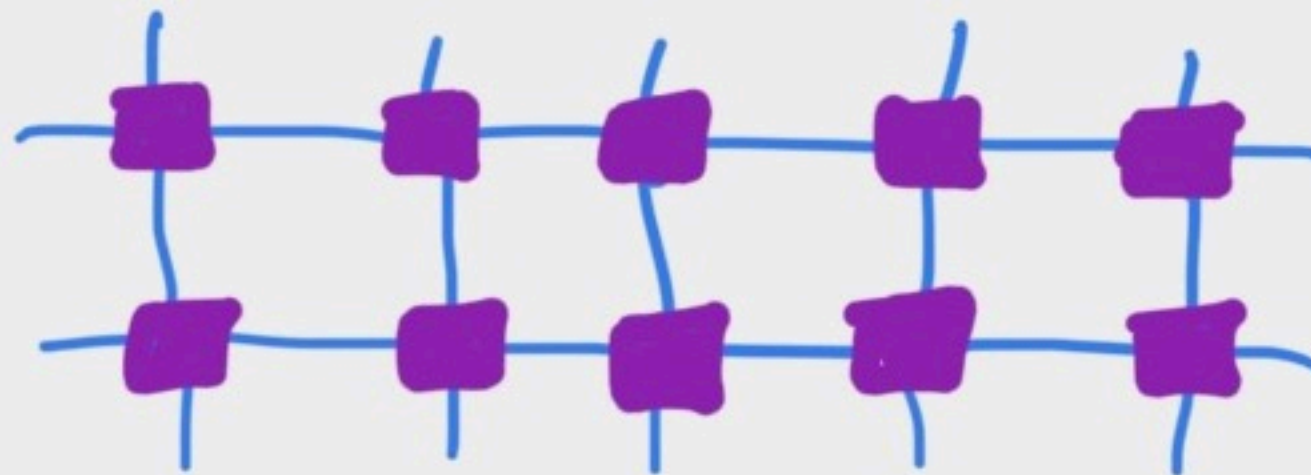
Zooming In



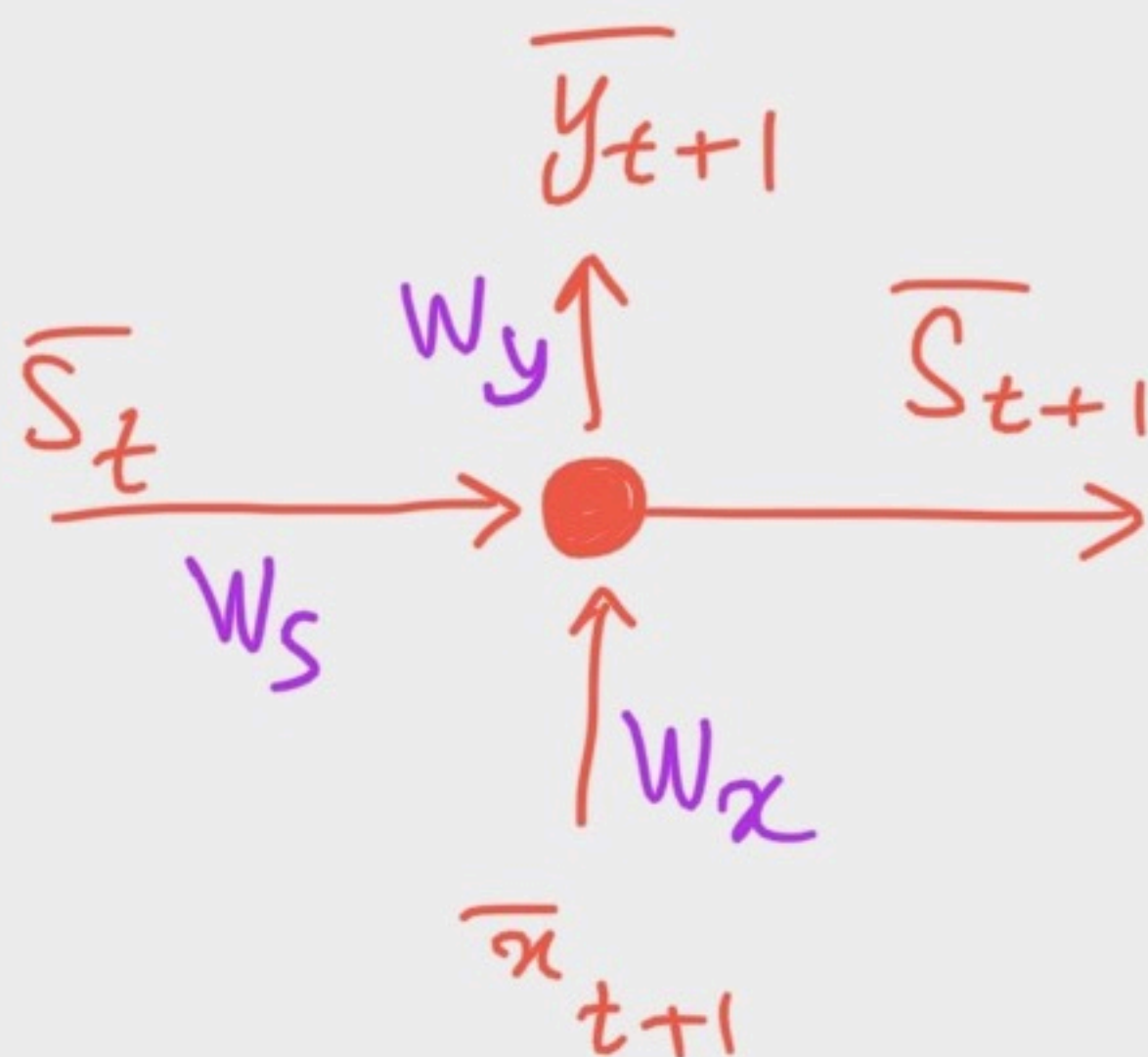
Zooming In



LSTM

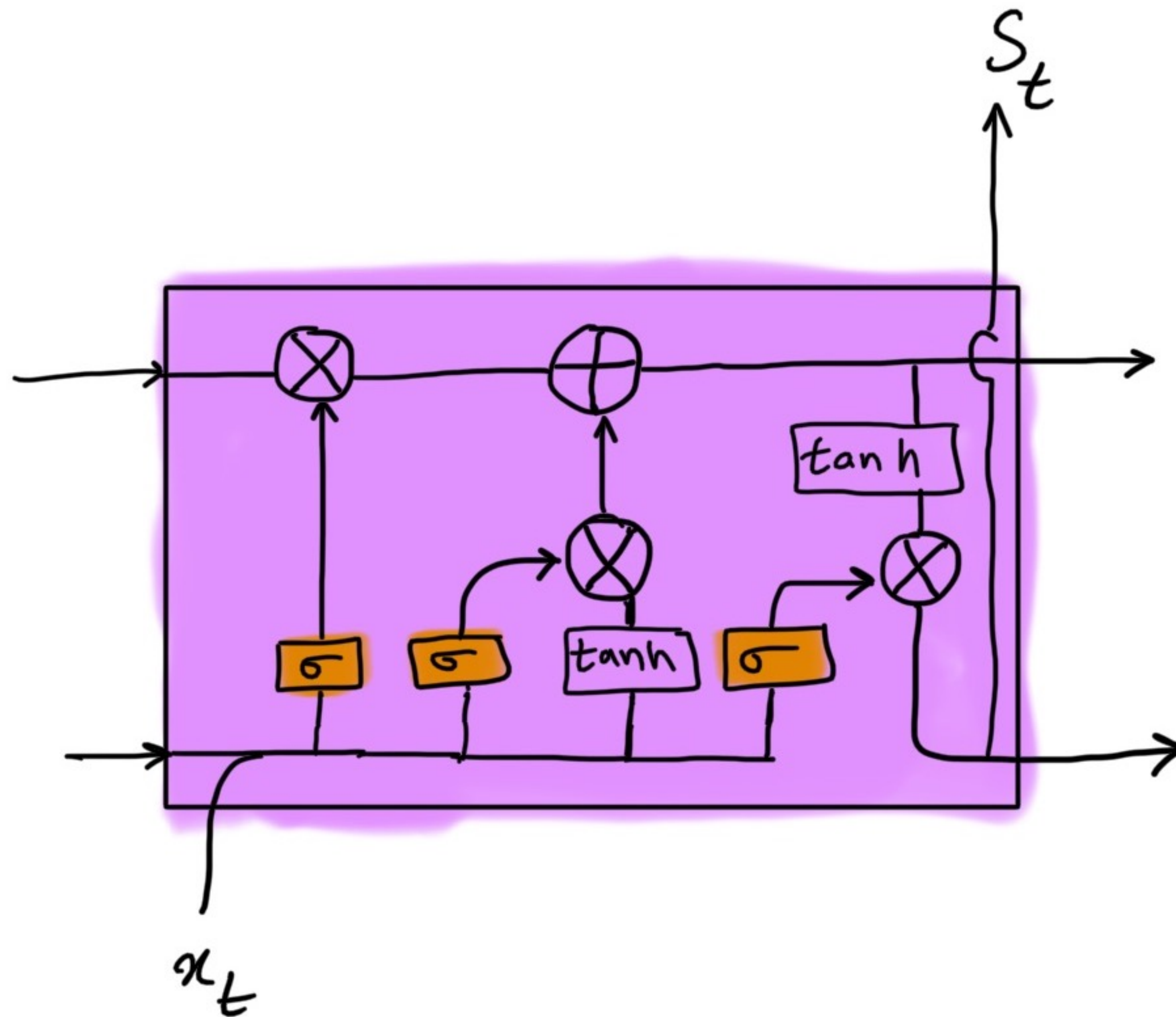


Zooming In





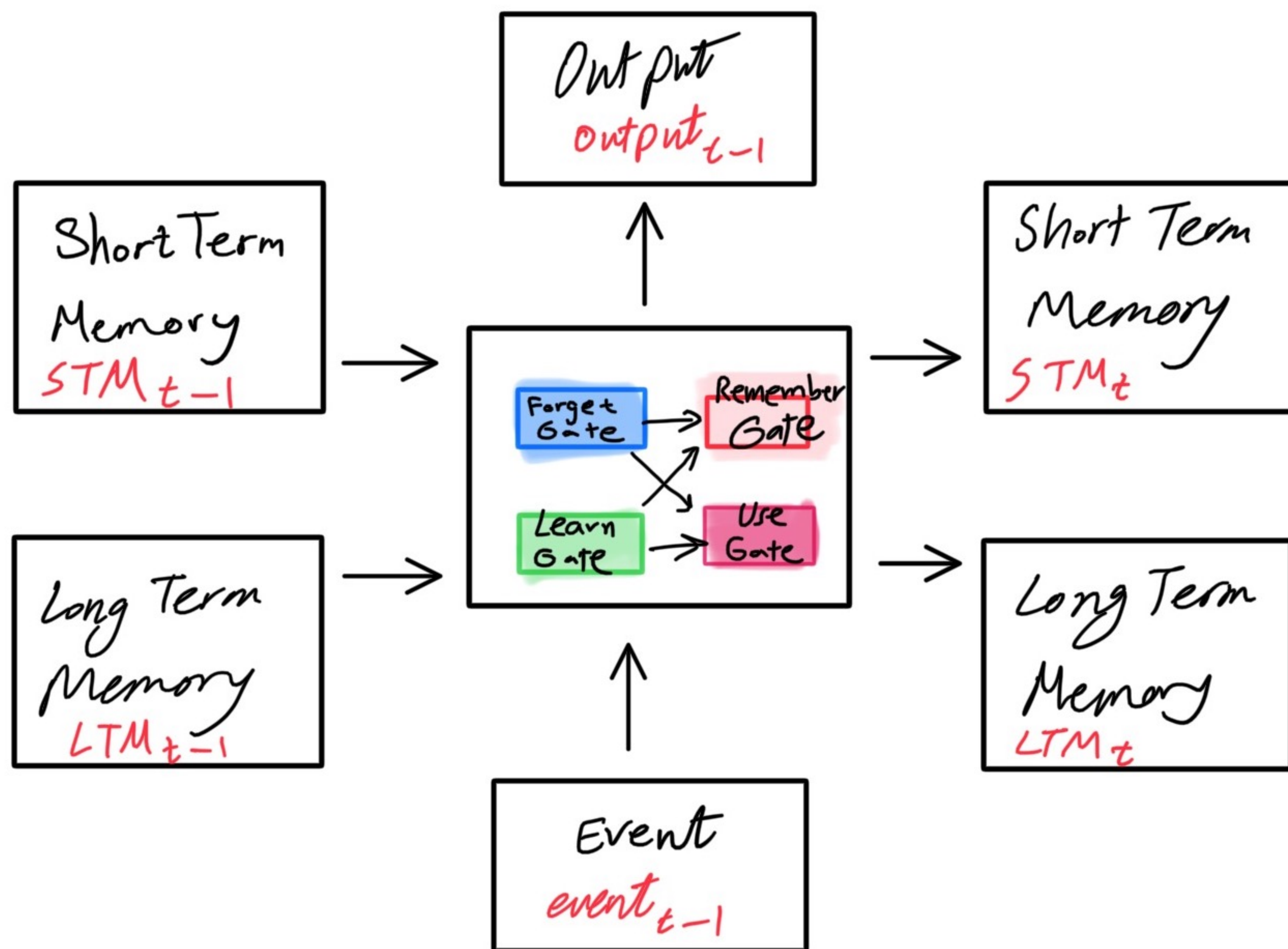
↓ Zooming In



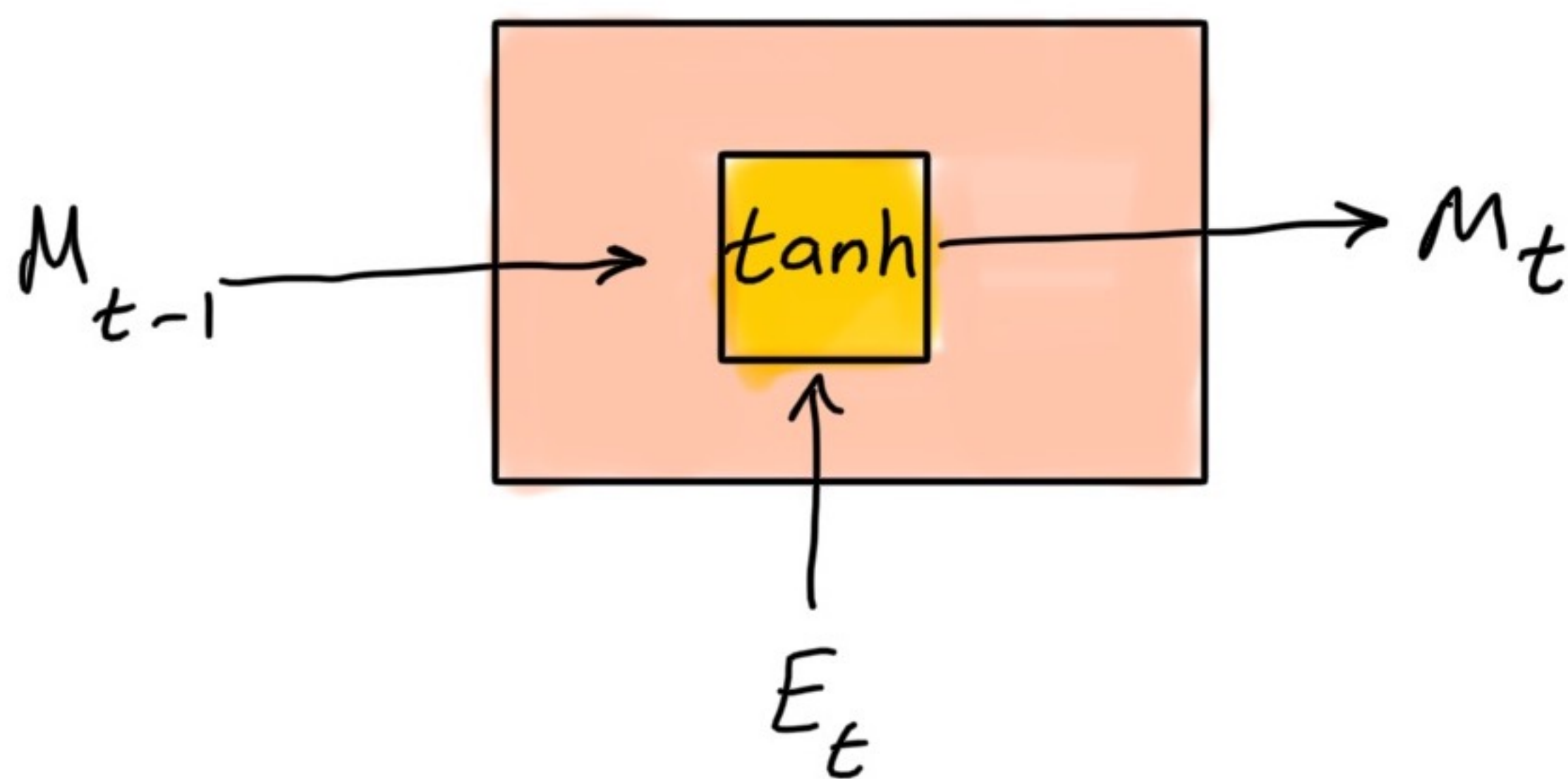
$\sigma$ 
 {
   
 $\sigma(x) \approx 1$  All data passes through
   
 $\sigma(x) \approx 0$  No data passes through
   
 Gating Functions
   
 ↓
   
 They decide what data to retain.



# LSTM Explained (visually!)



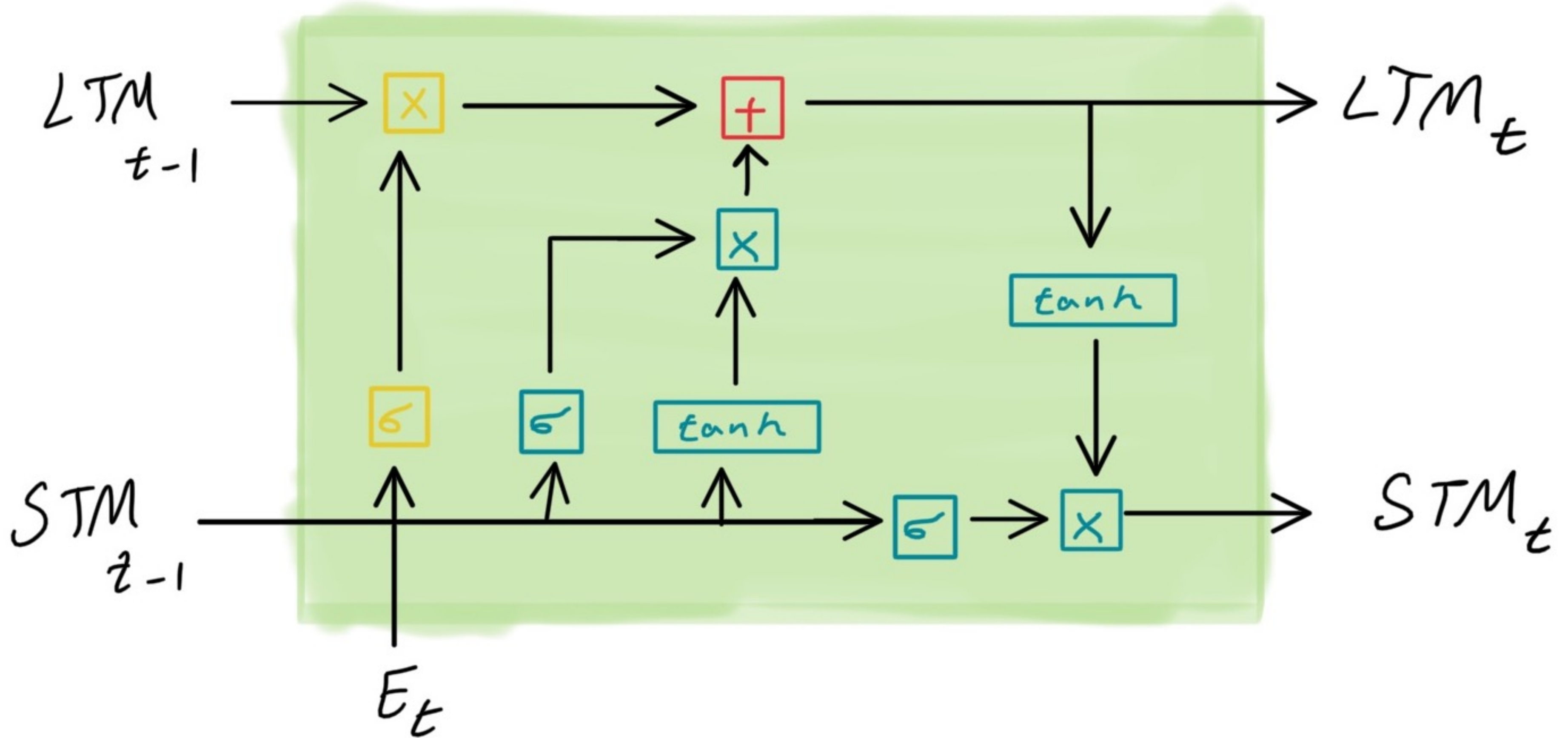
## Reminder for RNN



$$M_t = \tanh(W[STM_{t-1}, E_t] + b)$$



So now Let's see how LSTM looks like:



Let's look at each component individually

