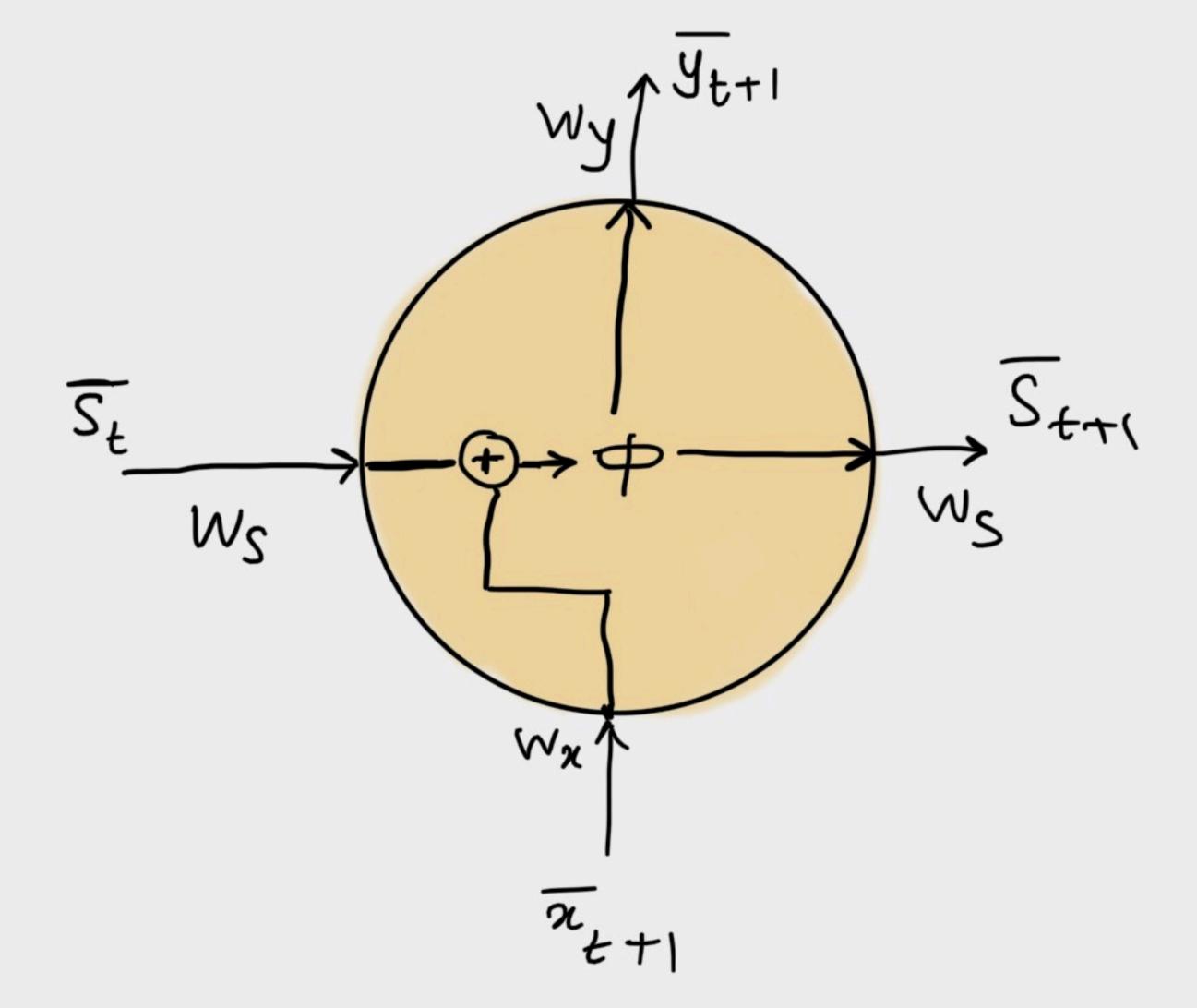
LSTM

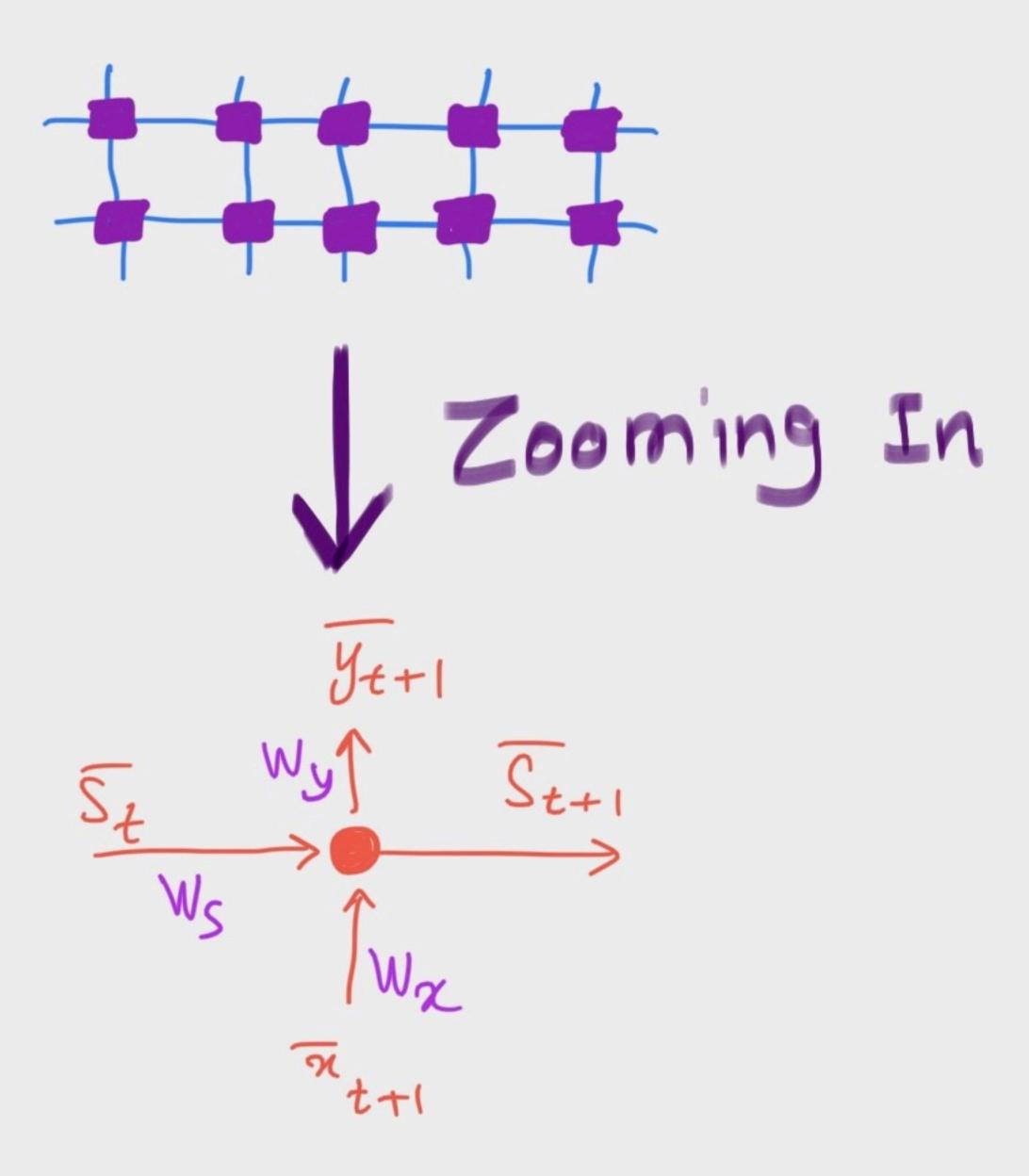
or long short-term memory cell

let's Compane LSTM and RNN RNN Zooming In

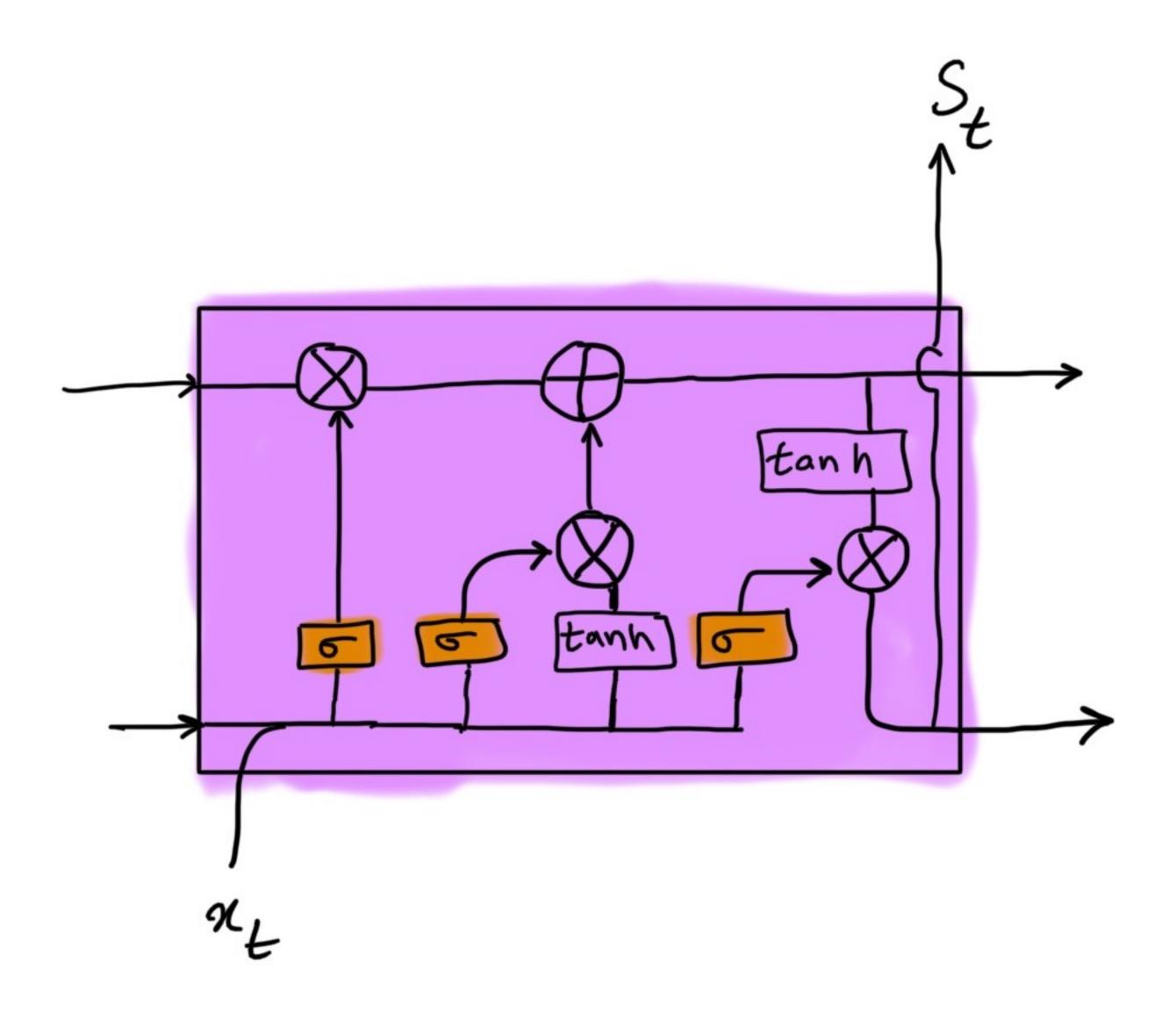
Zooming In



LSTM

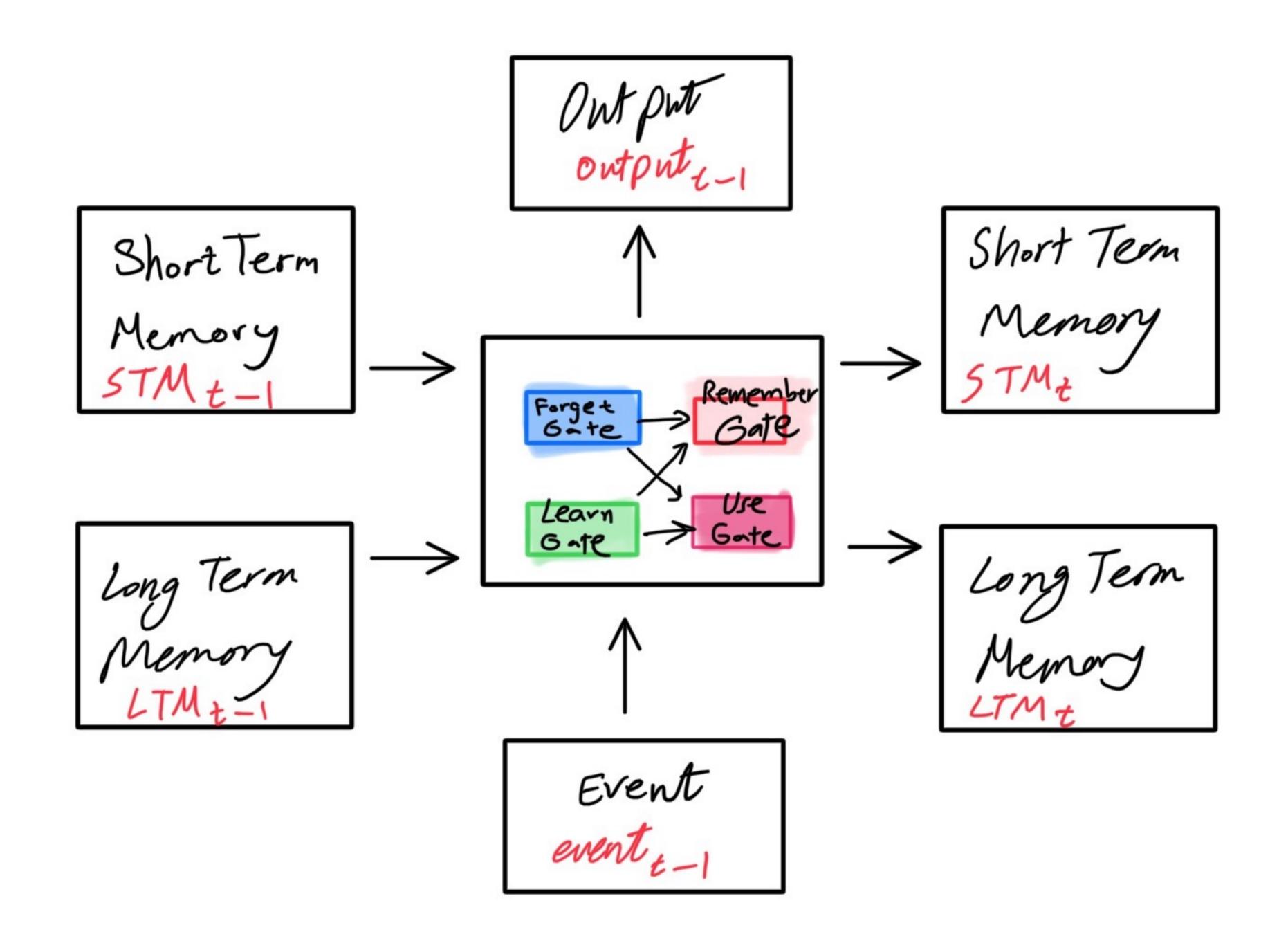






Gating
Functions (x) & O No data passes through
They decide what data to retain.

LSTM Explained (visually!)

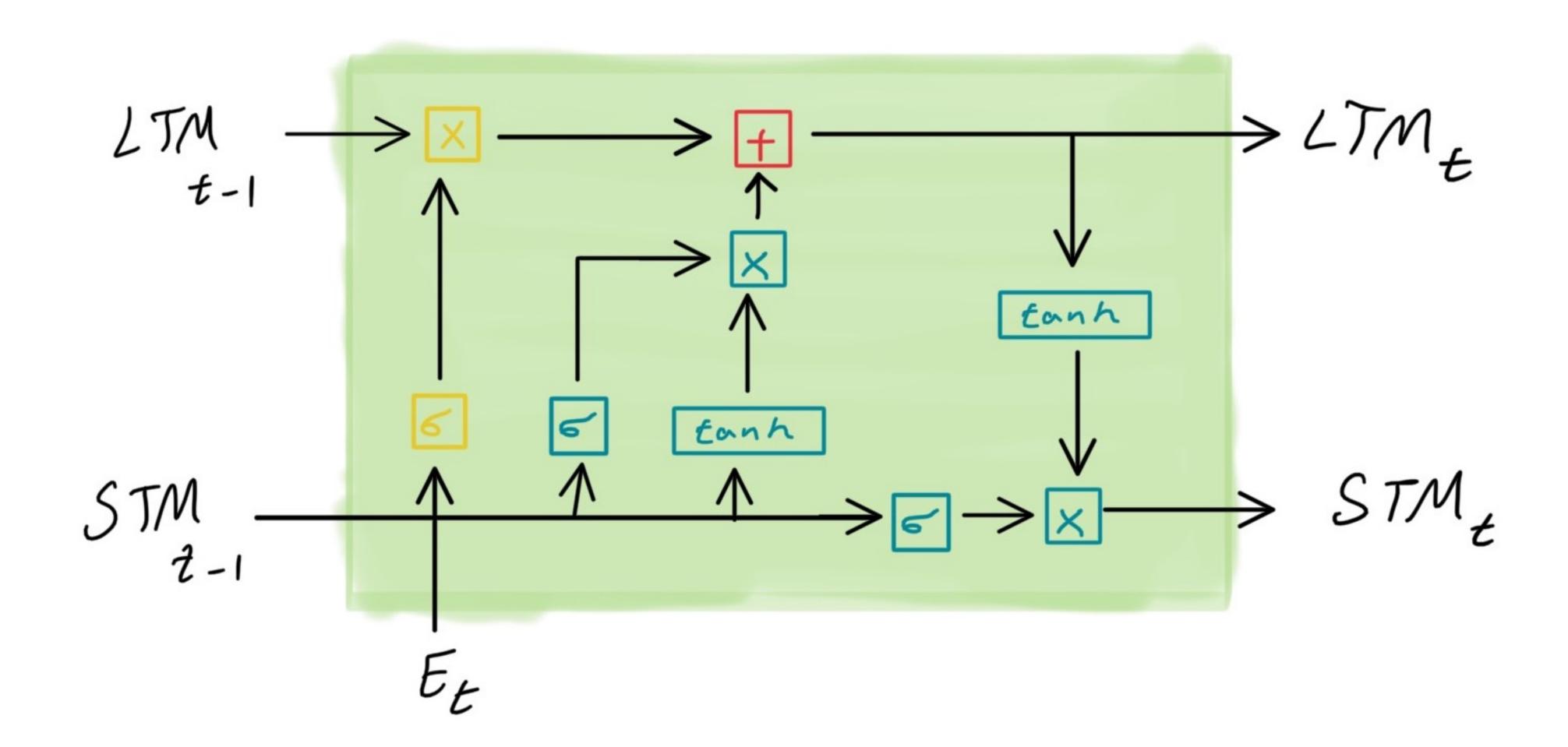


Reminder for RNN

$$M_{t-1}$$
 $tanh$
 E_t

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

So now let's see how LSTM Looks Like:



Let's book at each component individually

STM

Et

STM

Combine

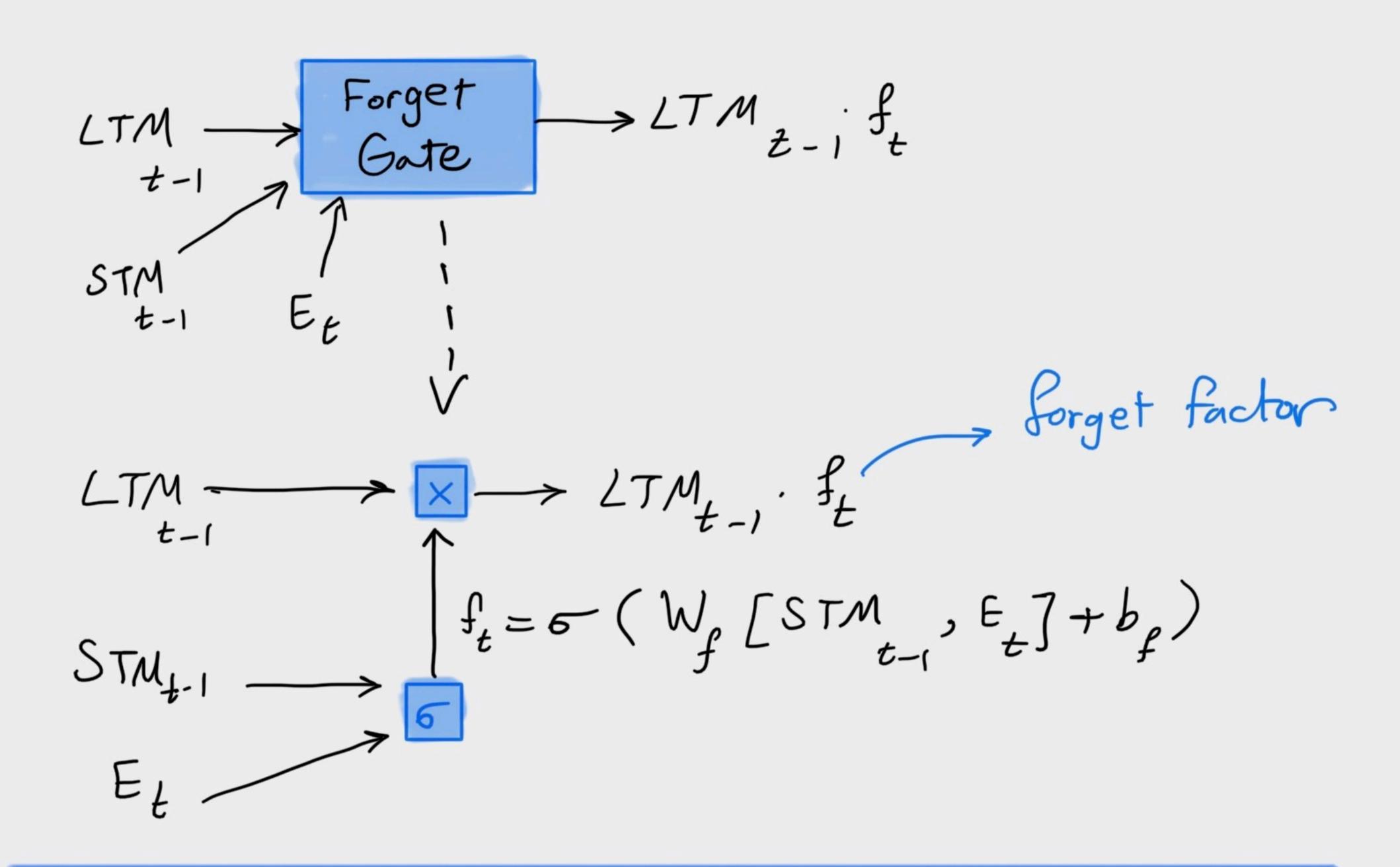
STM

$$t-1$$

Et

Anh

 $t = tanh(W_n[STM, \xi_1+b_n)]$
 $t = tanh(W_n[STM, \xi_1+b_n)]$
 $t = tanh(W_n[STM, \xi_1+b_n)]$
 $t = tanh(W_n[STM, \xi_1+b_n)]$



$$\begin{array}{c} LTM_{\xi-1} \longrightarrow Remember \\ STM \longrightarrow Gate \\ \downarrow \\ LTM \longrightarrow Forget \\ t-1 \end{array} \longrightarrow \begin{array}{c} Forget \\ gate \\ t-1 \end{array} \longrightarrow LTM = LTM \cdot f_t + N \cdot i_t \\ STM \longrightarrow gate \\ t-1 \end{array}$$

$$STM \longrightarrow Gate \longrightarrow LTM_{\xi}$$

$$LTM_{t-1} \rightarrow Use Gate \rightarrow LTM_{t}$$

$$STM_{t-1} \rightarrow Use Gate \rightarrow LTM_{t}$$

$$U_{t} = tanh (W_{t}LTM_{t-1} \cdot f_{t} + b_{t})$$

$$LTM_{t-1} \rightarrow STM_{t} = U_{t} \cdot V_{t}$$

$$STM_{t-1} \rightarrow V_{t} = G (W_{t}[STM_{t-1}, E_{t}] + b_{t})$$

$$E_{t} \qquad V_{t} = G (W_{t}[STM_{t-1}, E_{t}] + b_{t})$$