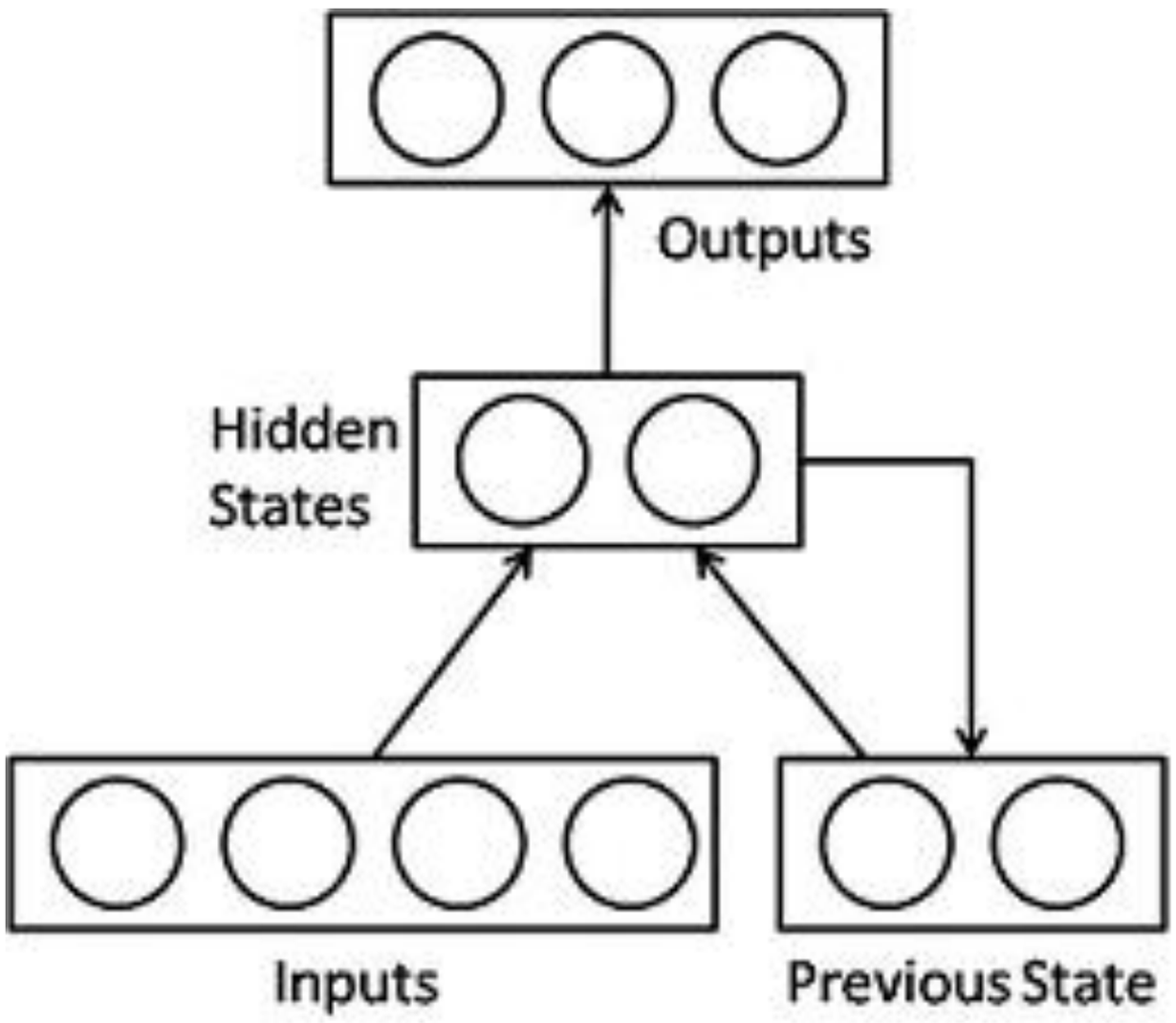
Deep Learning: Theory and Practice

Recurrent Neural Networks

04-04-2018

deeplearning.cce2018@gmail.com

Recurrent Operations



Feedback

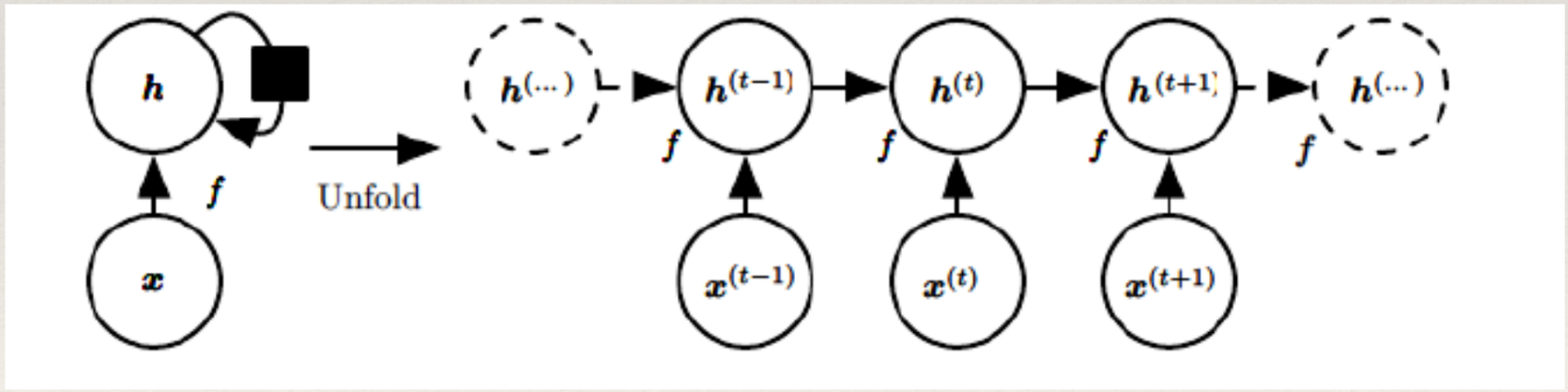
Recurrent Networks

$$\boldsymbol{s}^{(t)} = f(\boldsymbol{s}^{(t-1)}; \boldsymbol{\theta}),$$

$$\begin{aligned} \boldsymbol{s}^{(3)} &= f(\boldsymbol{s}^{(2)}; \boldsymbol{\theta}) \\ &= f(f(\boldsymbol{s}^{(1)}; \boldsymbol{\theta}); \boldsymbol{\theta}) \end{aligned}$$

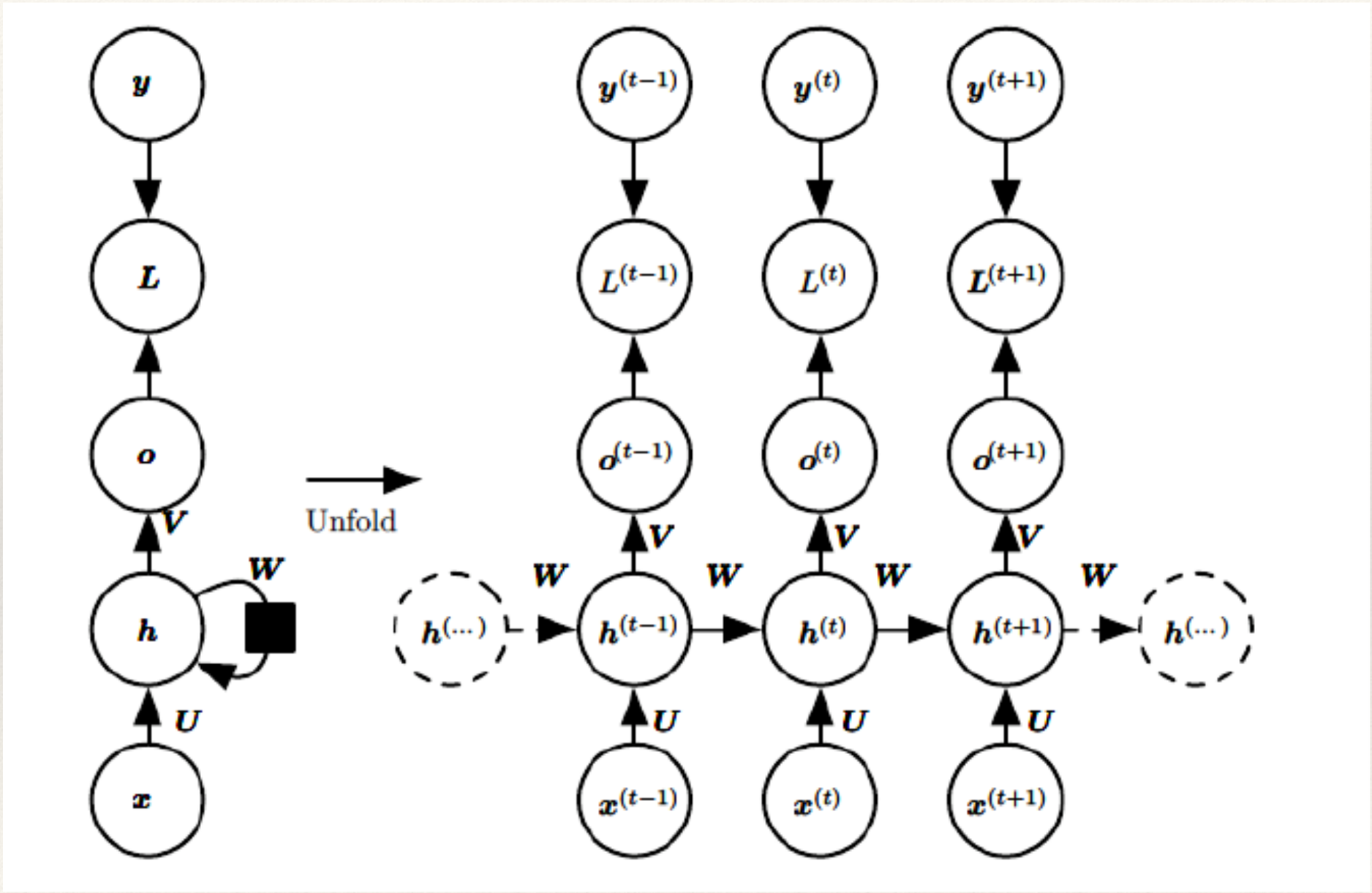
$$\boldsymbol{s}^{(t)} = f(\boldsymbol{s}^{(t-1)}, \boldsymbol{x}^{(t)}; \boldsymbol{\theta}),$$

$$\boldsymbol{h}^{(t)} = f(\boldsymbol{h}^{(t-1)}, \boldsymbol{x}^{(t)}; \boldsymbol{\theta}),$$



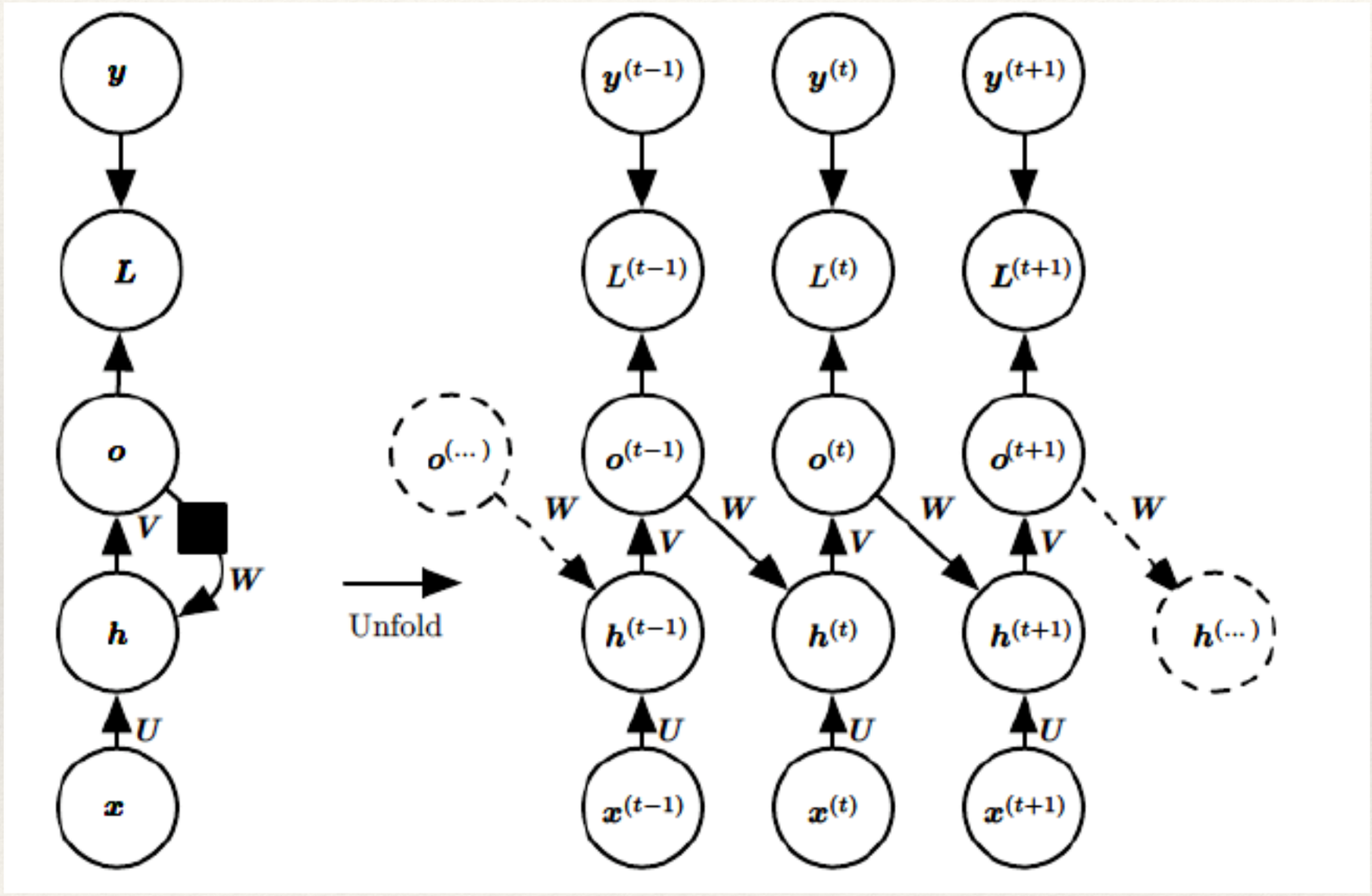
"Deep Learning", Ian Goodfellow, Yoshua Bengio, Aaron Courville

Recurrent Networks



"Deep Learning", Ian Goodfellow, Yoshua Bengio, Aaron Courville

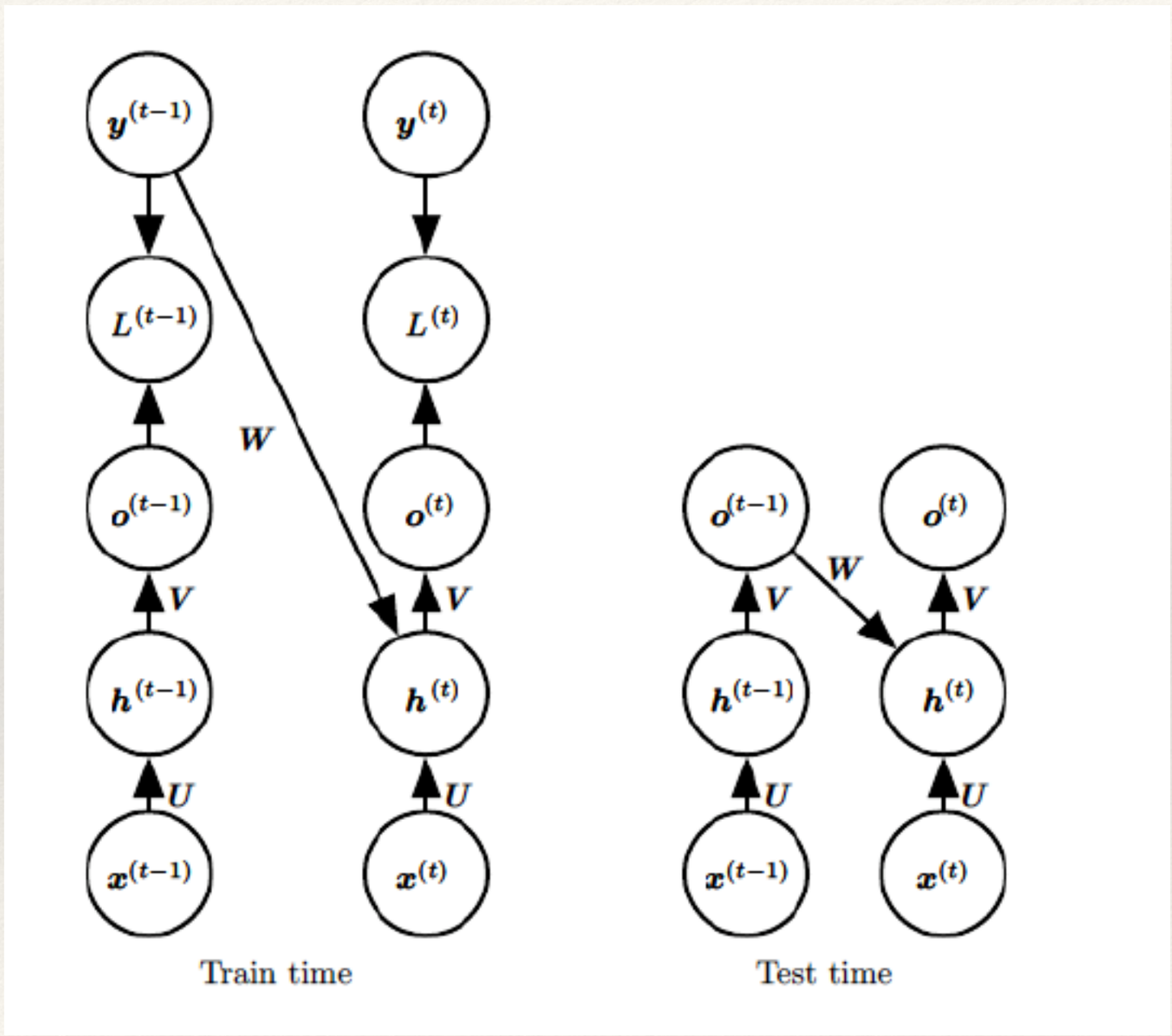
Recurrent Networks



Teacher Forcing Networks

"Deep Learning", Ian Goodfellow, Yoshua Bengio, Aaron Courville

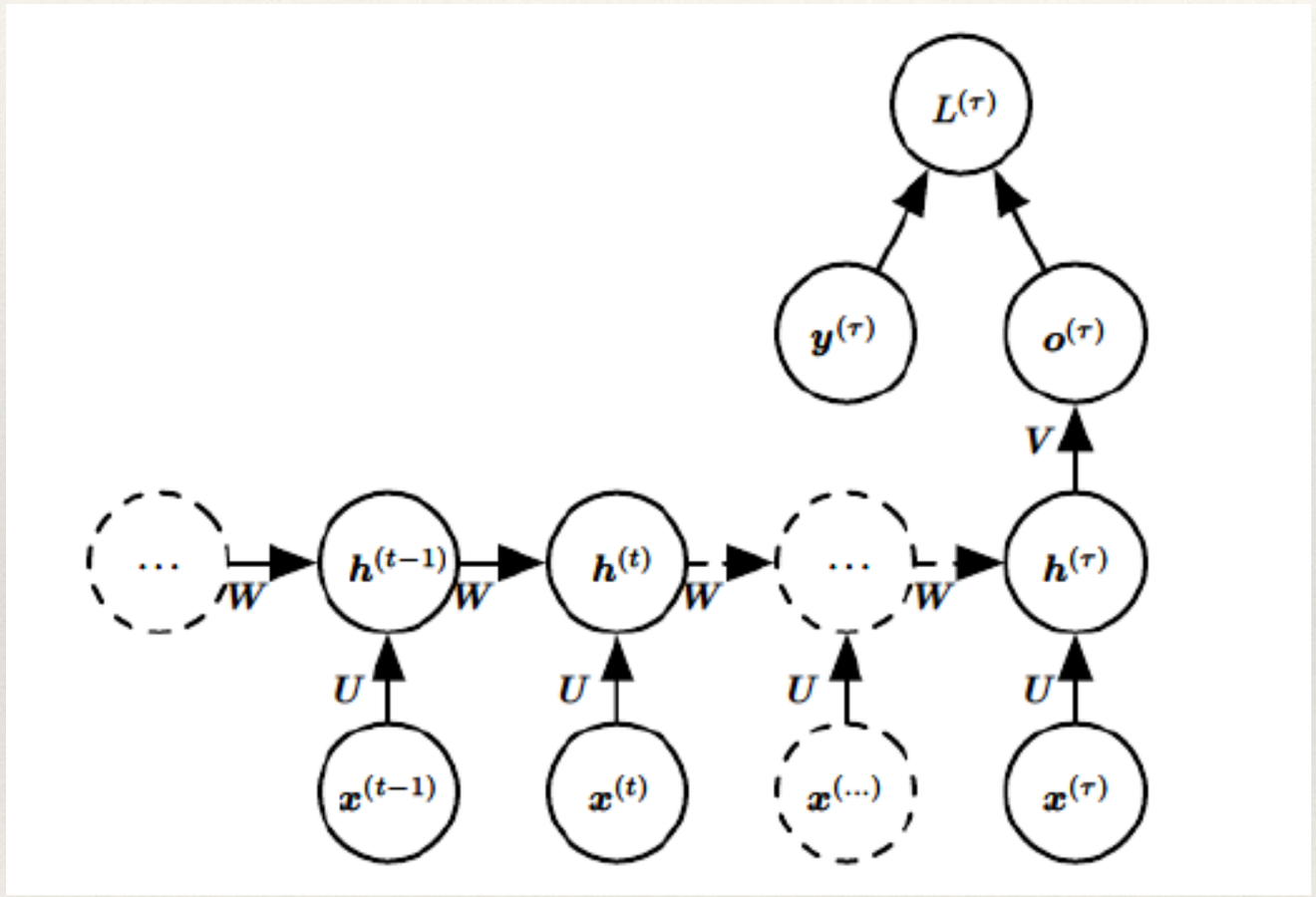
Recurrent Networks



Teacher Forcing Networks

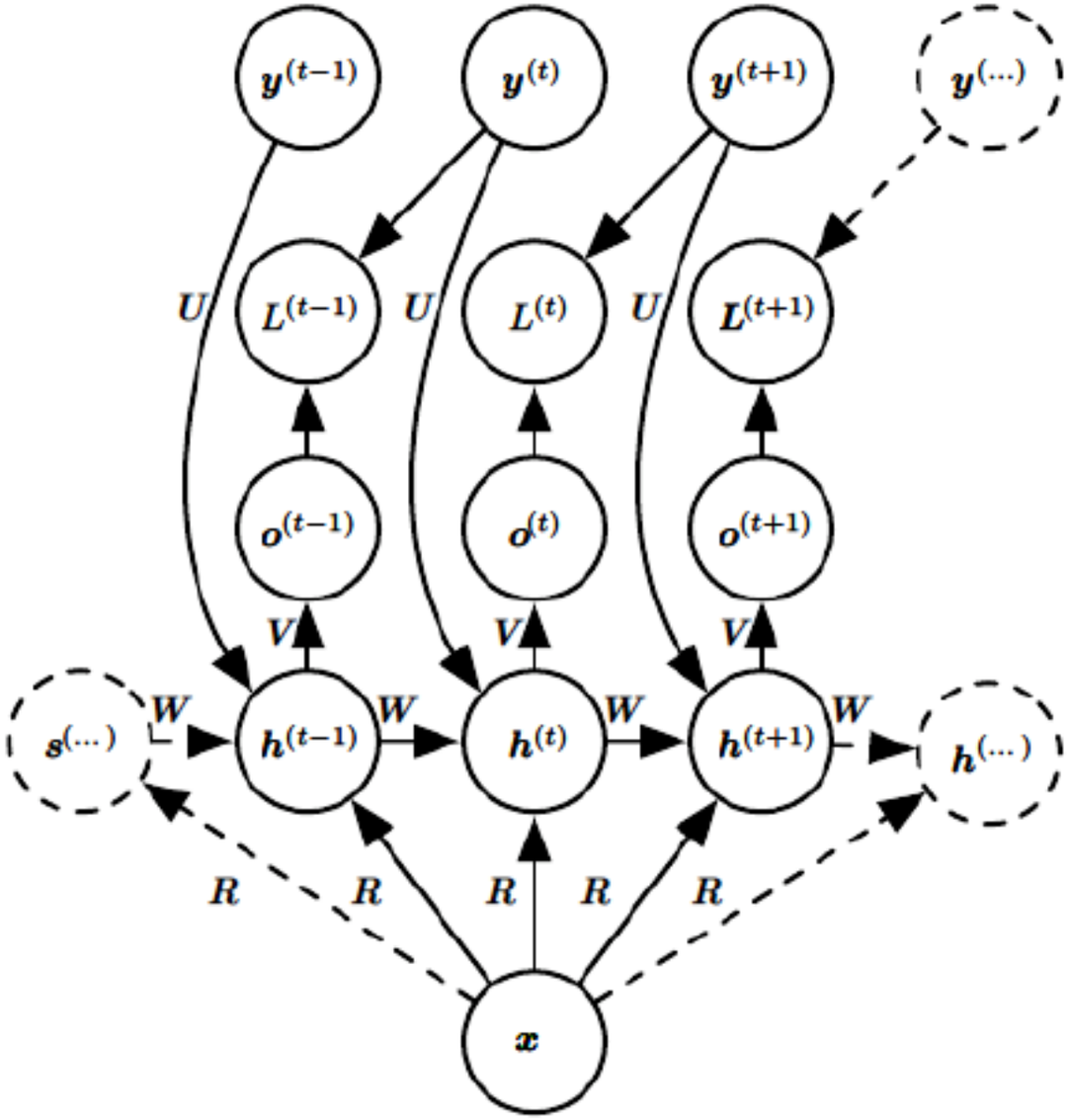
"Deep Learning", Ian Goodfellow, Yoshua Bengio, Aaron Courville

Recurrent Networks



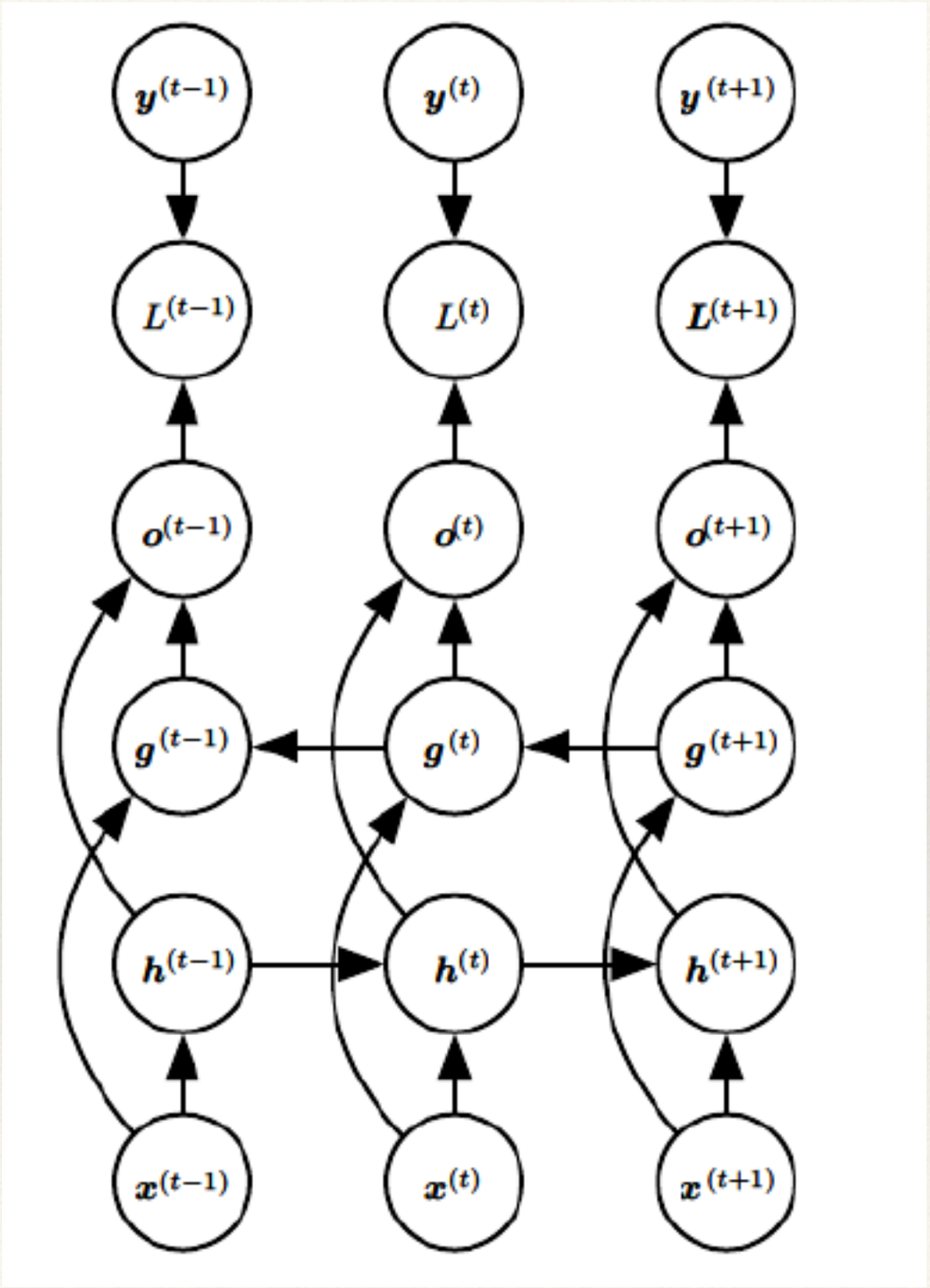
Multiple Input
Single Output

Recurrent Networks



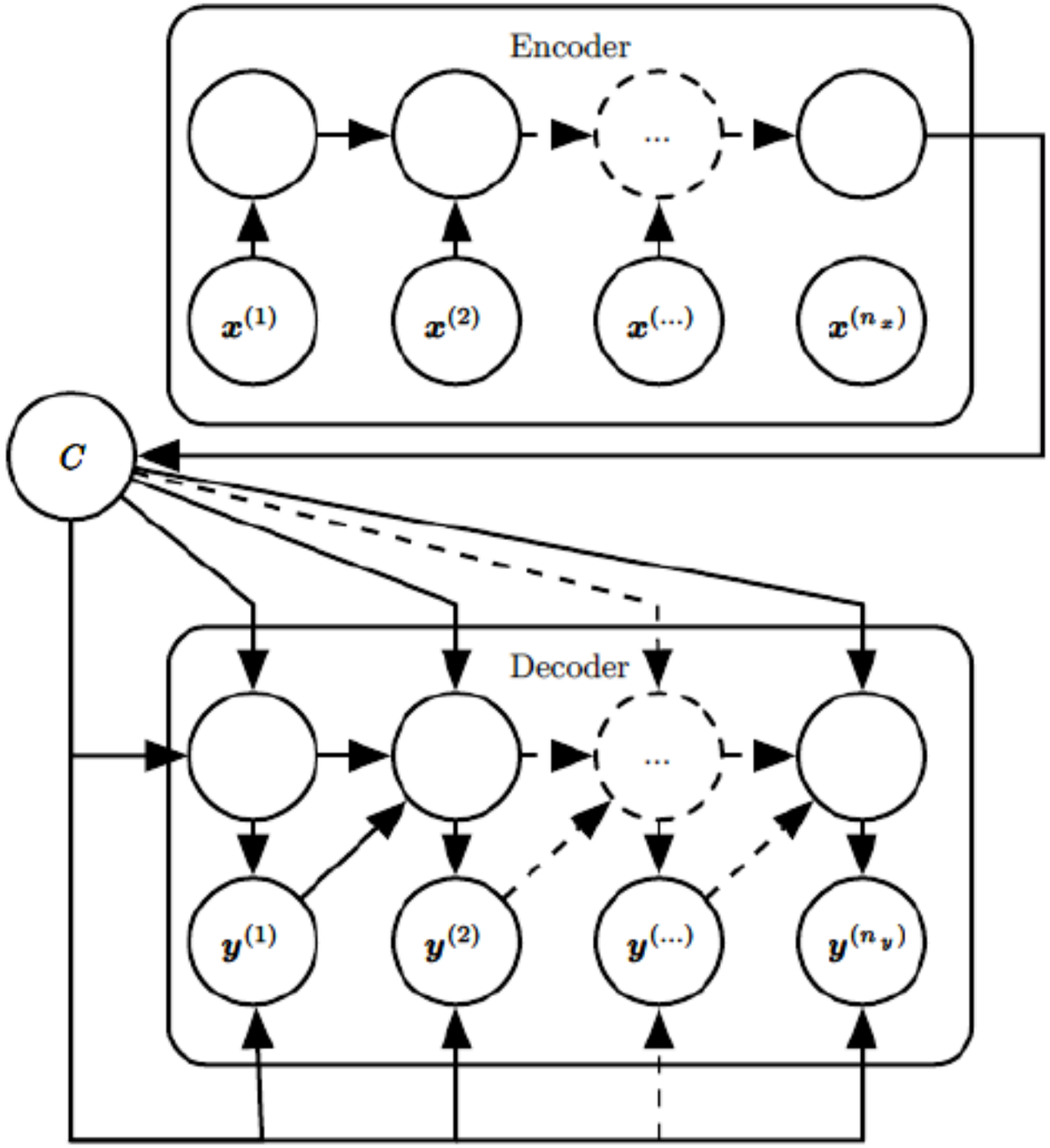
Single Input
Multiple Output

Recurrent Networks



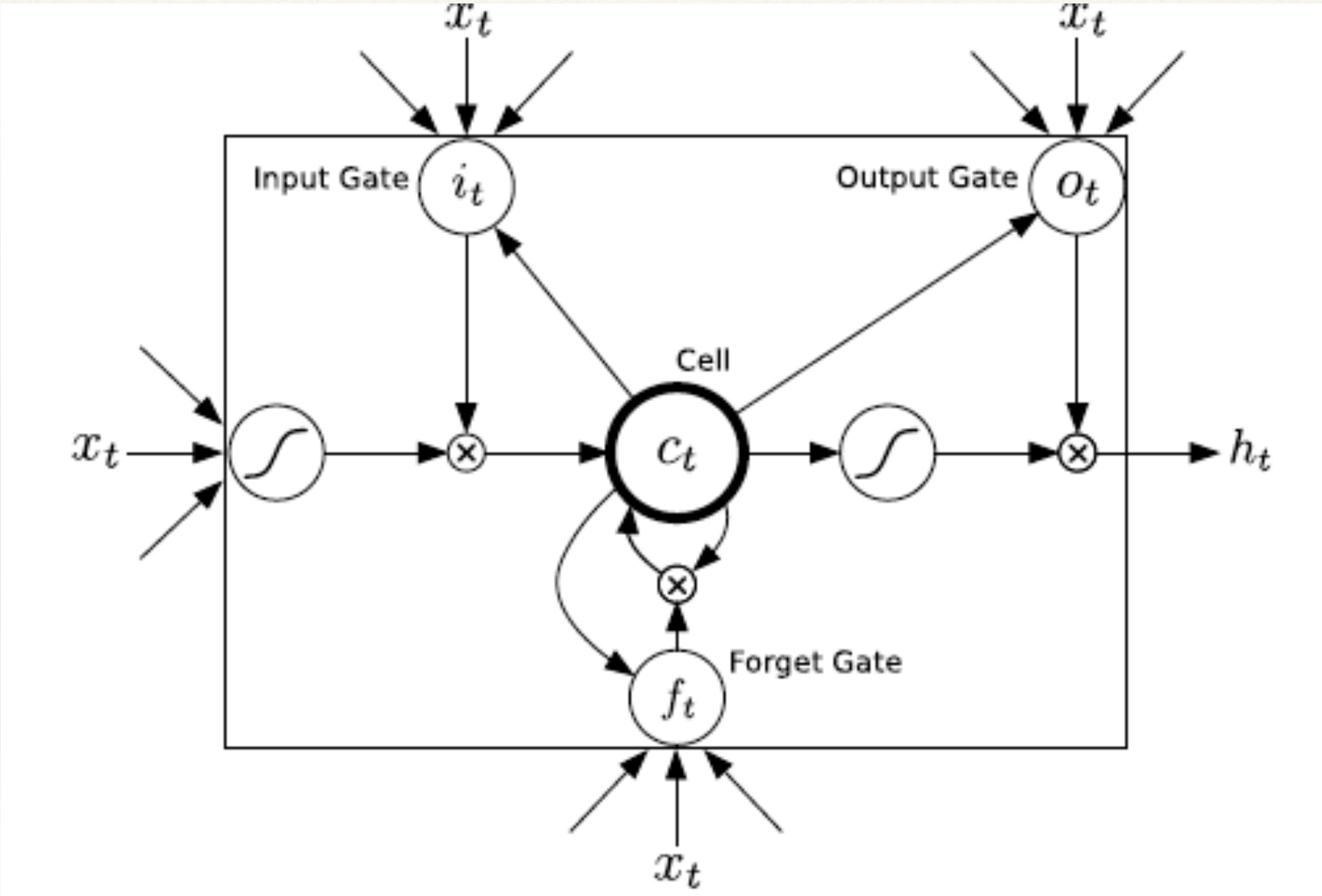
Bi-directional Networks

Recurrent Networks



Sequence to Sequence Mapping Networks

LSTM


x_t
x_t
Input Gate
i_t
Output Gate
o_t
Cell
x_t
c_t
h_t
Forget Gate
f_t
x_t

LSTM

$$
\begin{aligned}
i_t &= \sigma\left(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right) \\
f_t &= \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \\
c_t &= f_t c_{t-1} + i_t \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \\
o_t &= \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o\right) \\
h_t &= o_t \tanh(c_t)
\end{aligned}
$$

Bidirectional LSTMs