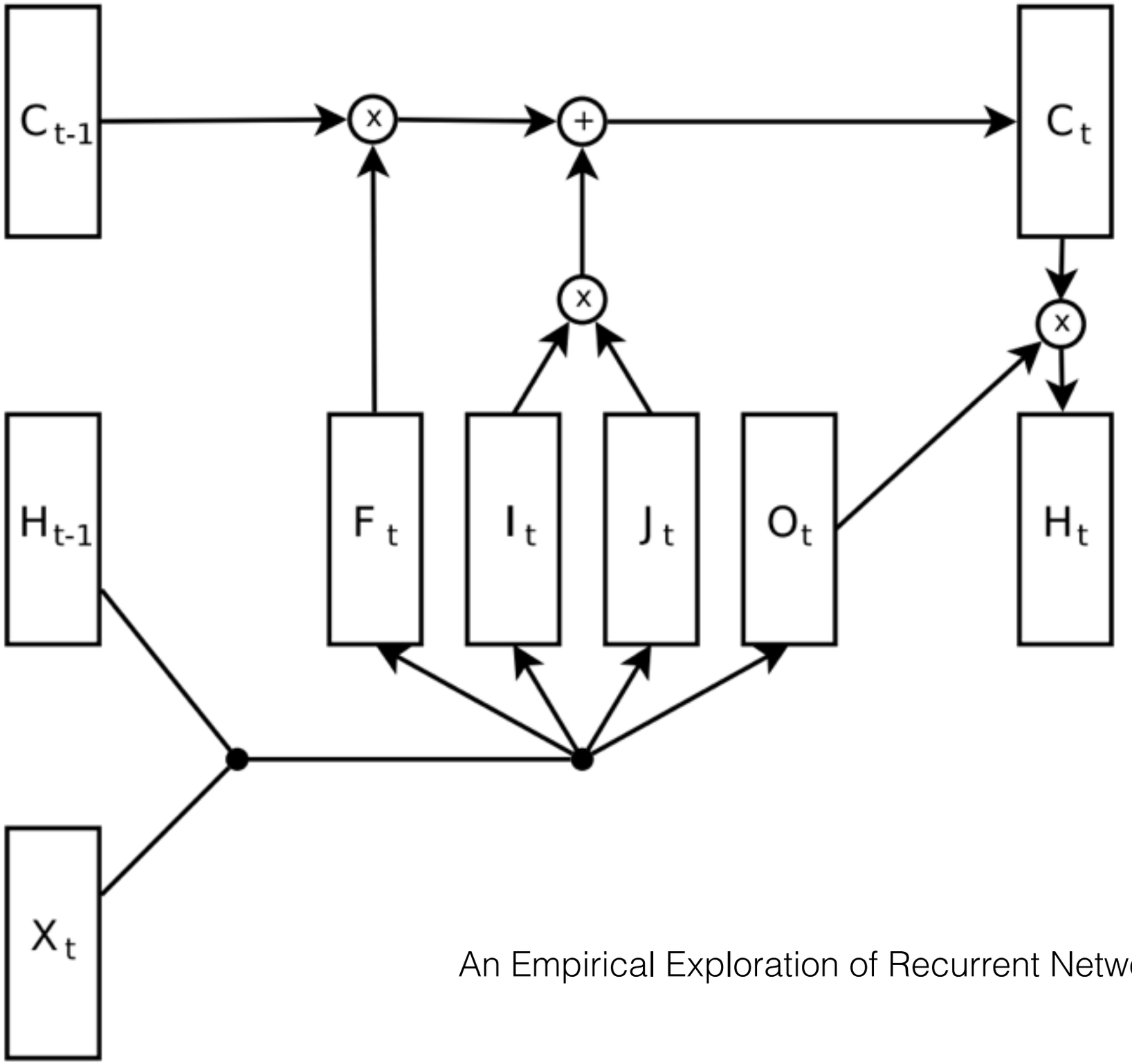
RNN Encoder-Decoder

Model and Implementation

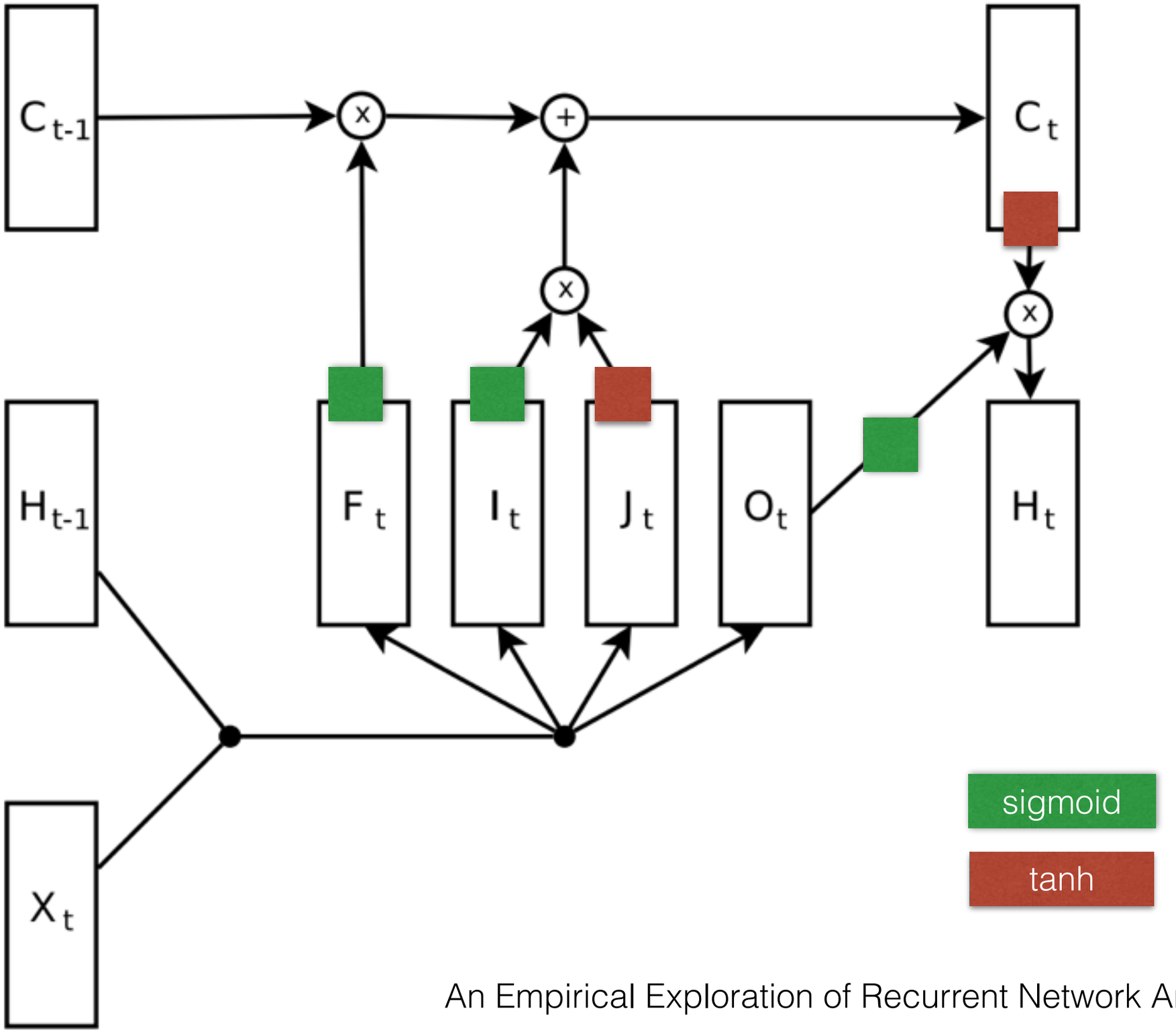
Shi Feng
Zhirui Zhang

LSTM and GRU



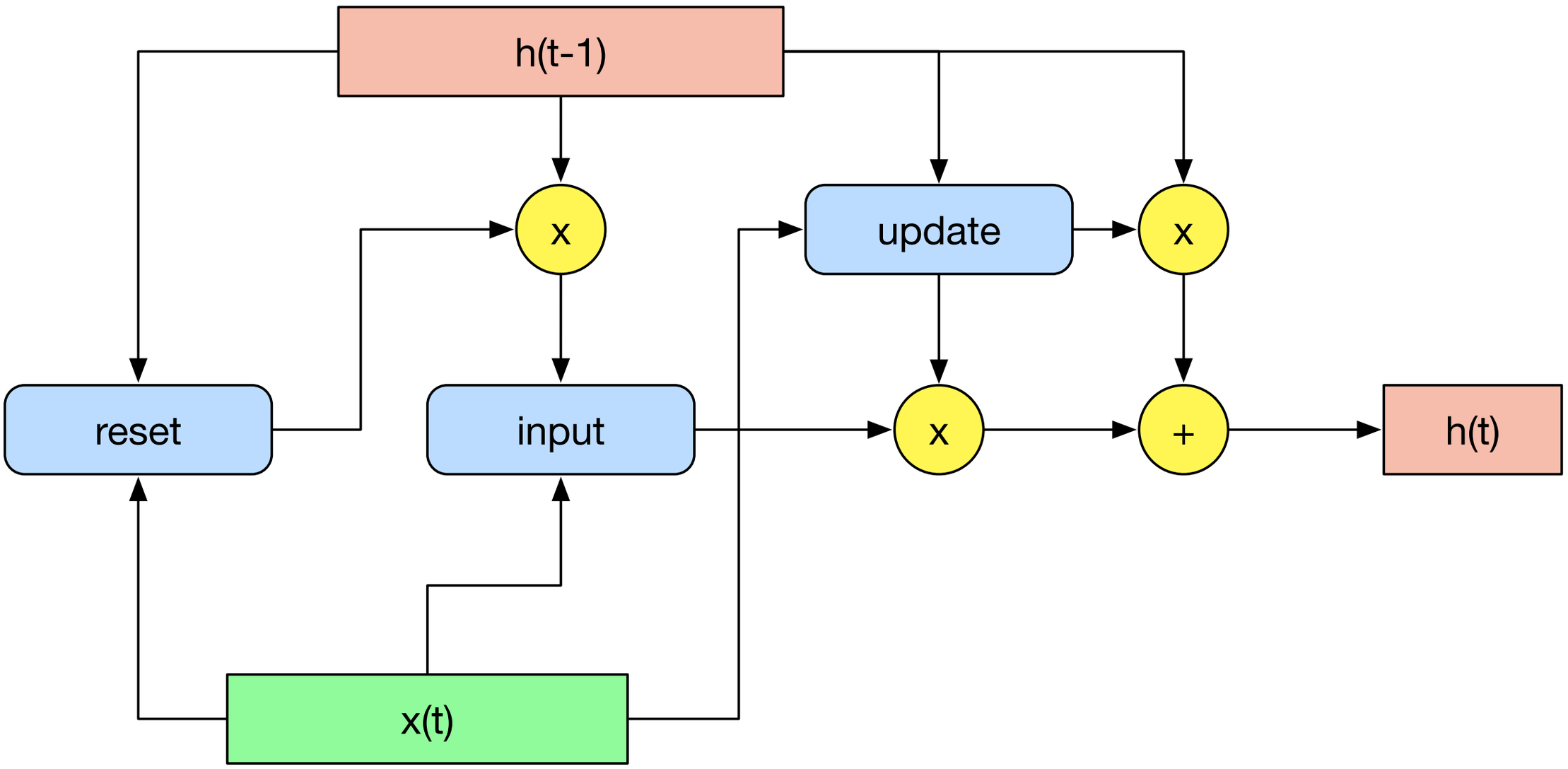
An Empirical Exploration of Recurrent Network Architectures

LSTM and GRU

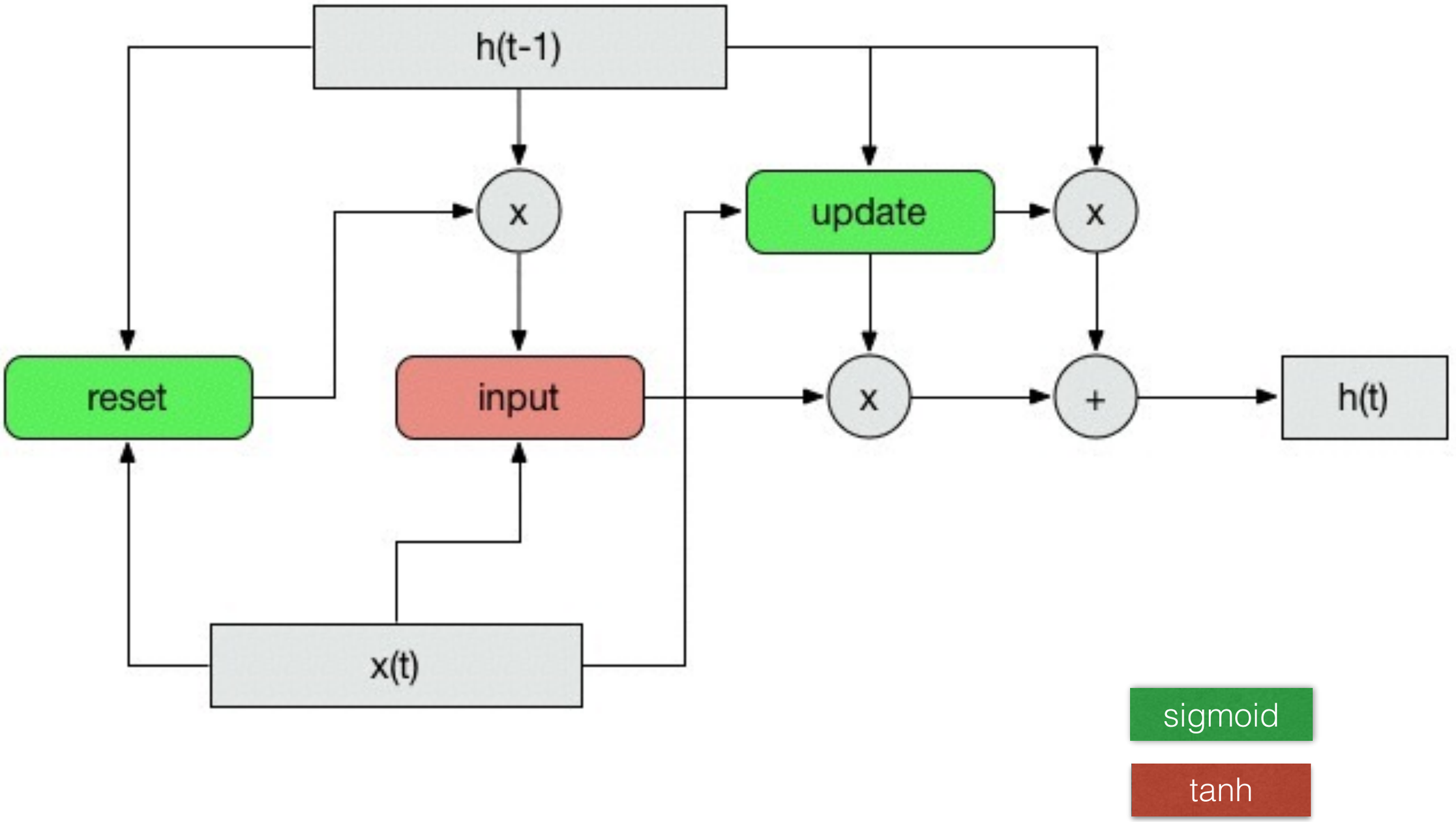


An Empirical Exploration of Recurrent Network Architectures

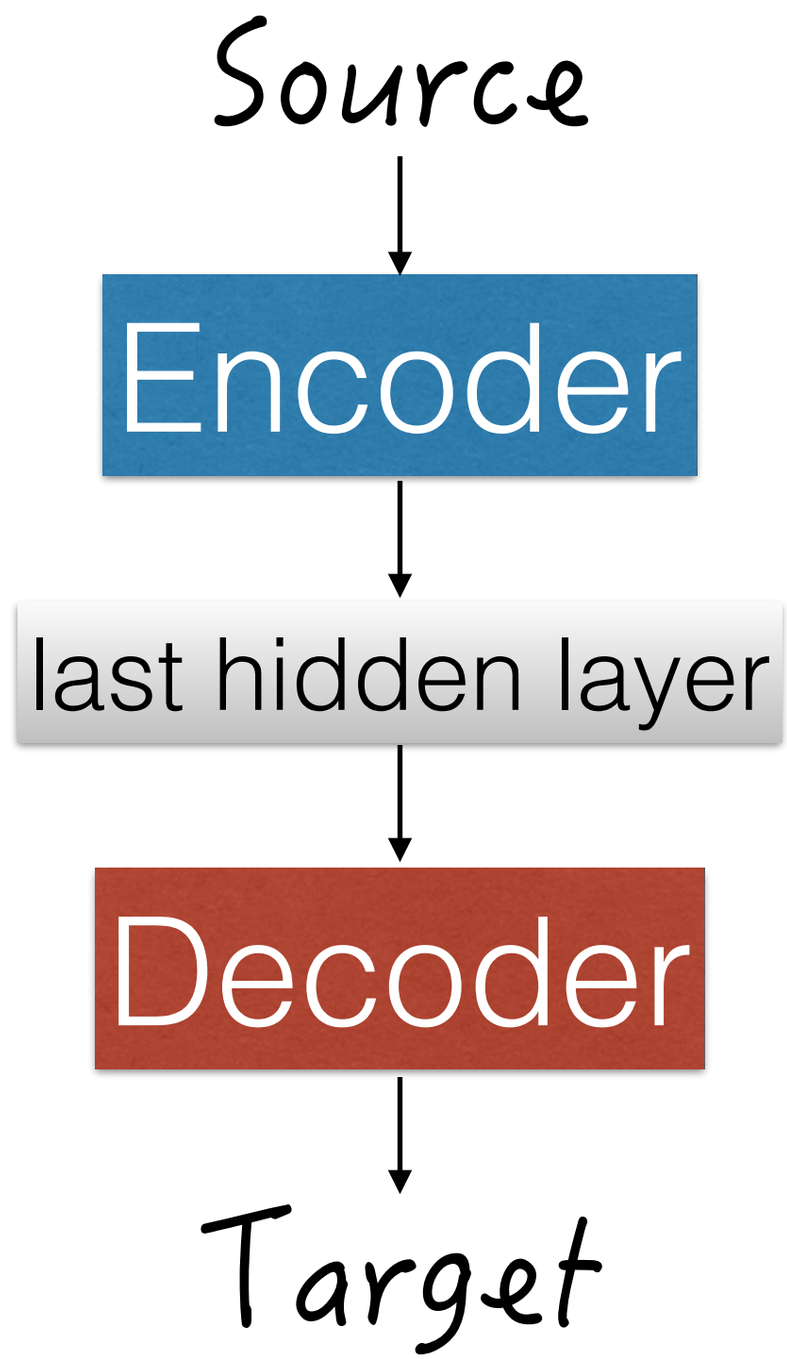
LSTM and GRU



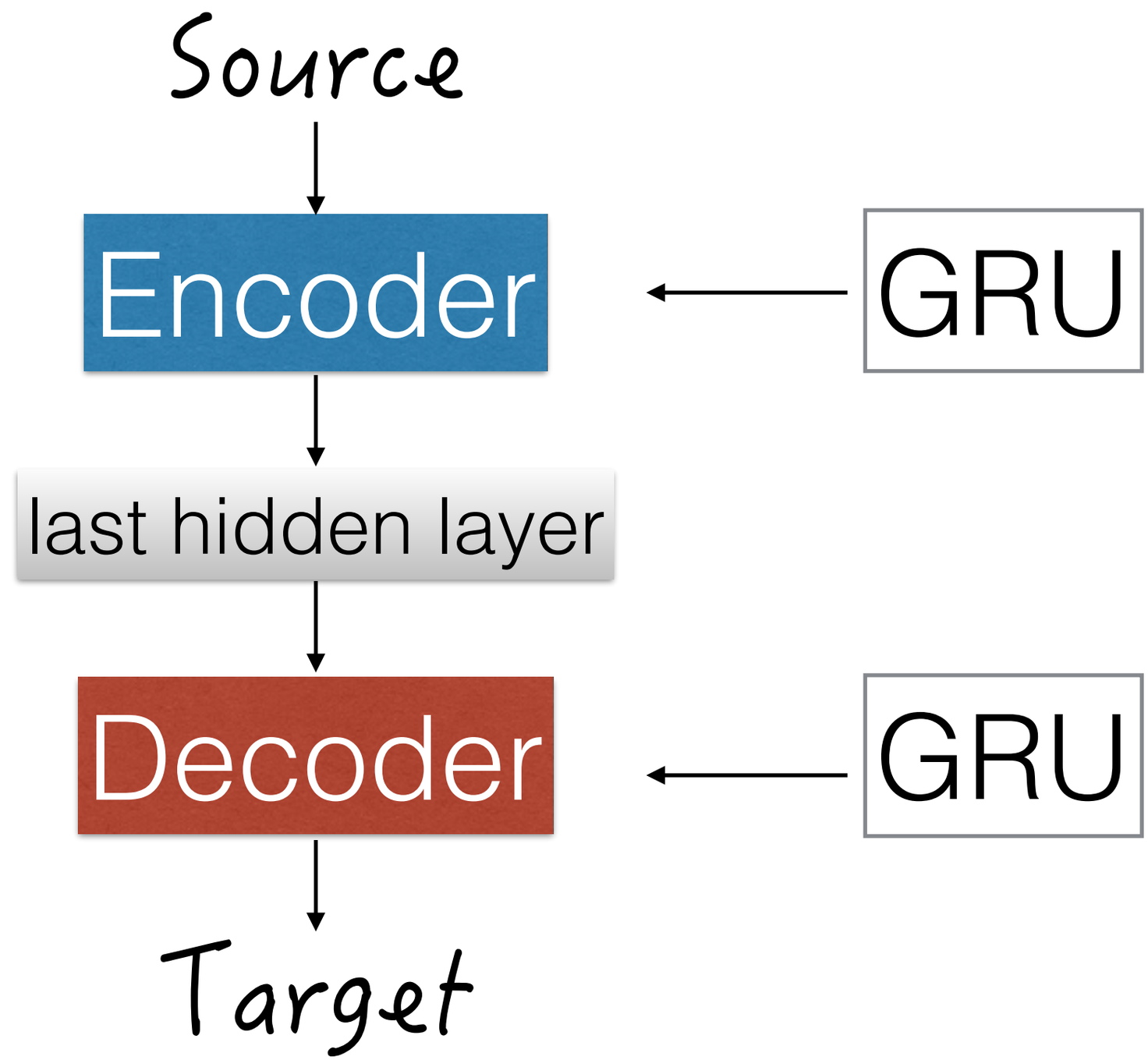
LSTM and GRU

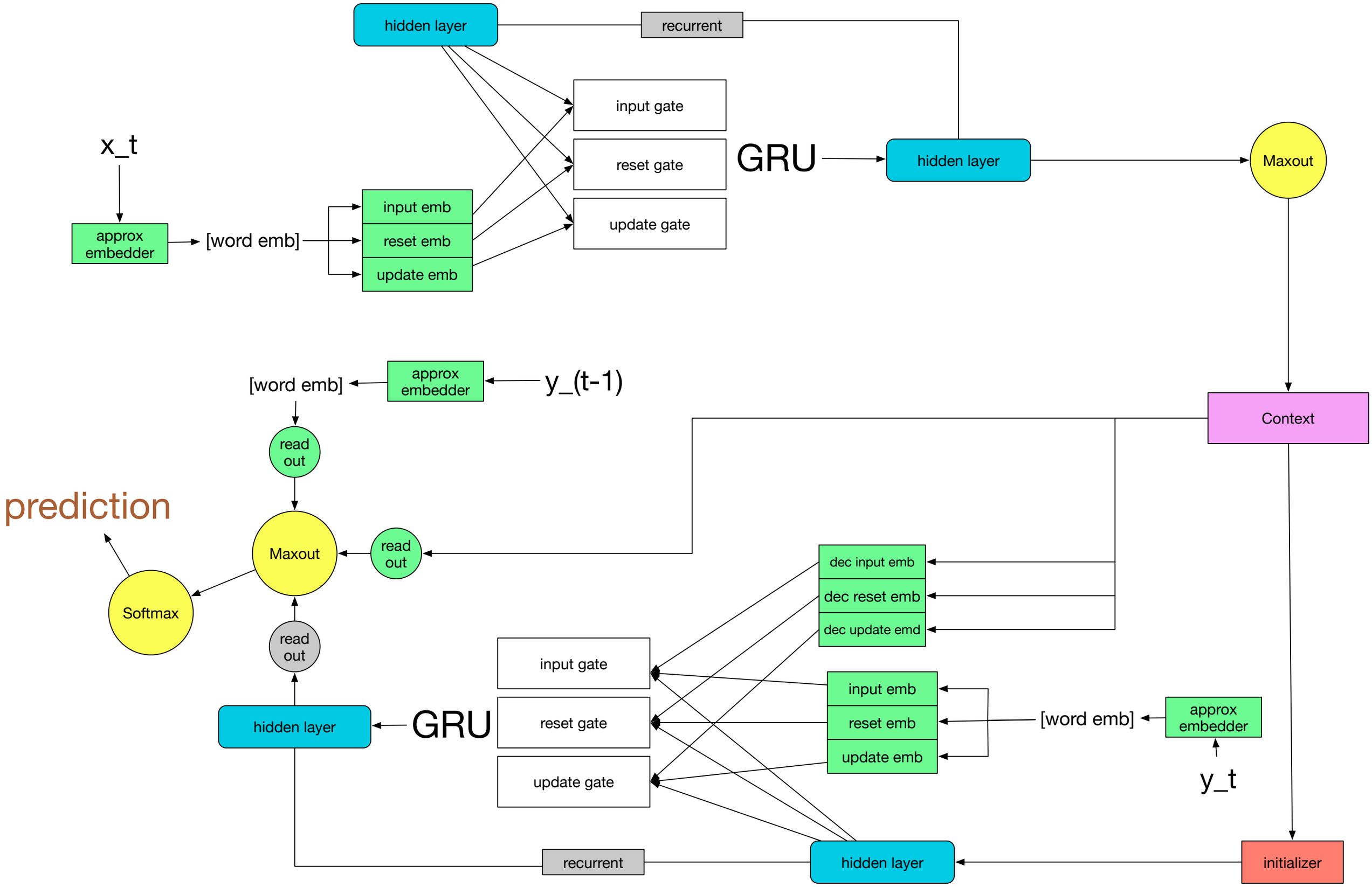


RNN Encoder-Decoder

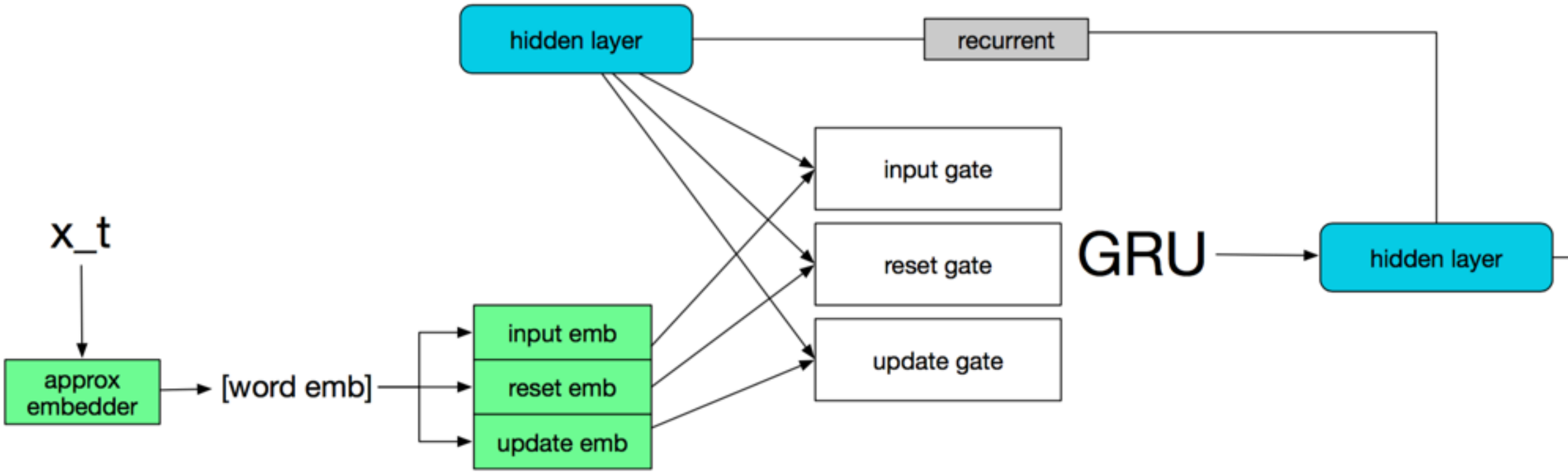


RNN Encoder-Decoder

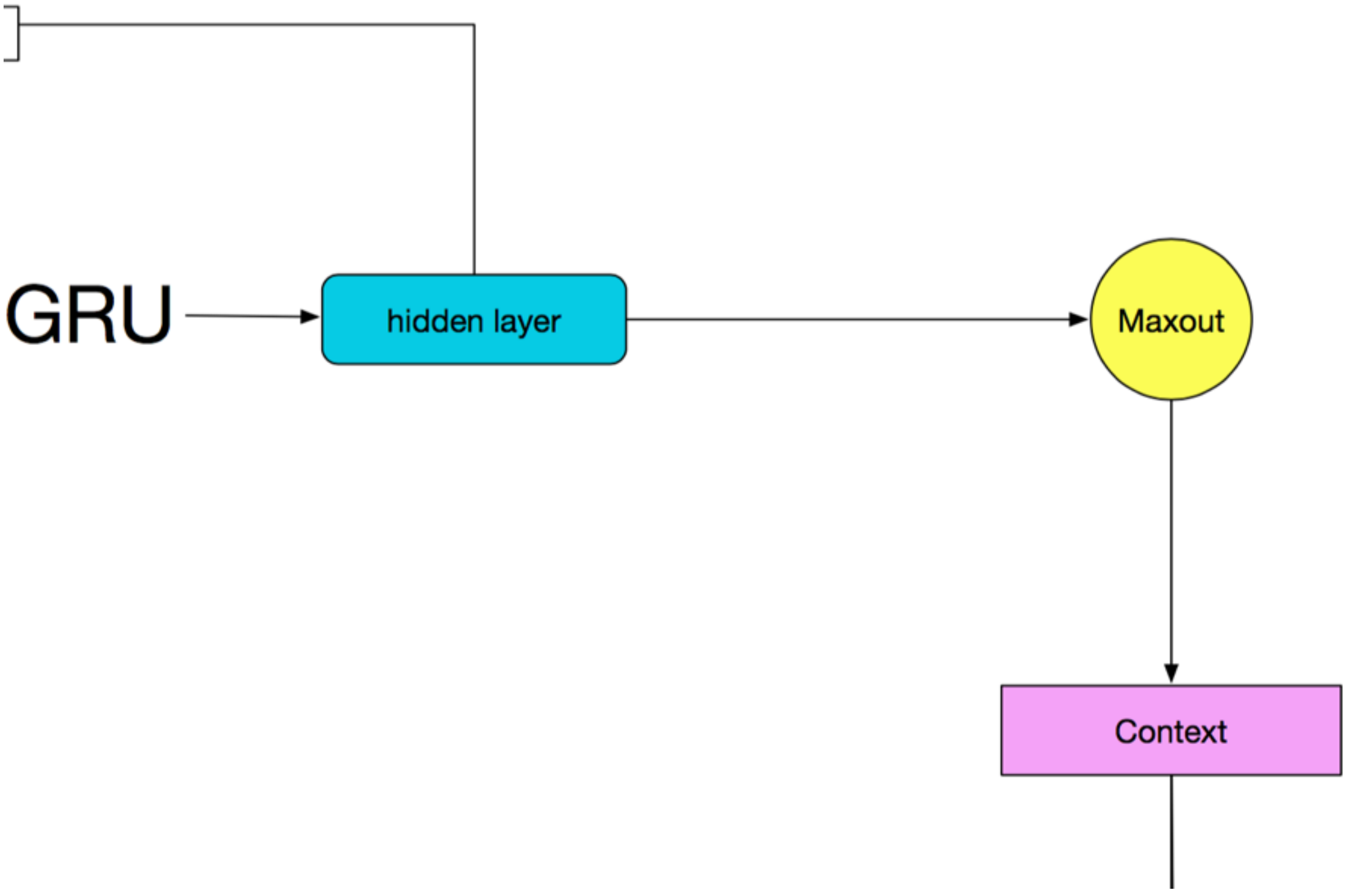



hidden layer
recurrent
input gate
reset gate
update gate
GRU
hidden layer
Maxout
x_t
approx embedder
[word emb]
input emb
reset emb
update emb
[word emb]
approx embedder
y_(t-1)
Context
read out
prediction
Maxout
read out
Softmax
read out
dec input emb
dec reset emb
dec update emd
input gate
reset gate
update gate
input emb
reset emb
update emb
[word emb]
approx embedder
y_t
hidden layer
GRU
recurrent
hidden layer
initializer

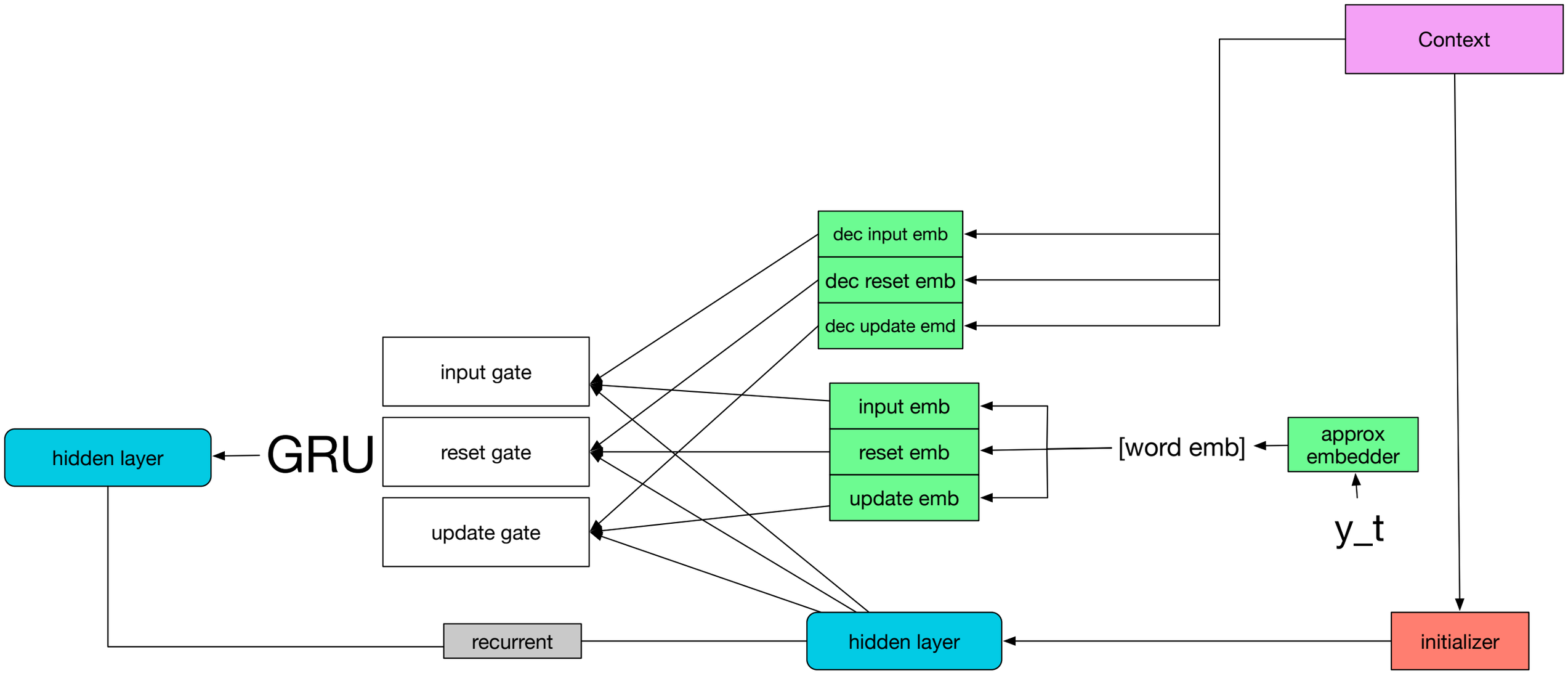
Encoder


hidden layer
recurrent
input gate
reset gate
update gate
GRU
hidden layer
x_t
approx embedder
[word emb]
input emb
reset emb
update emb

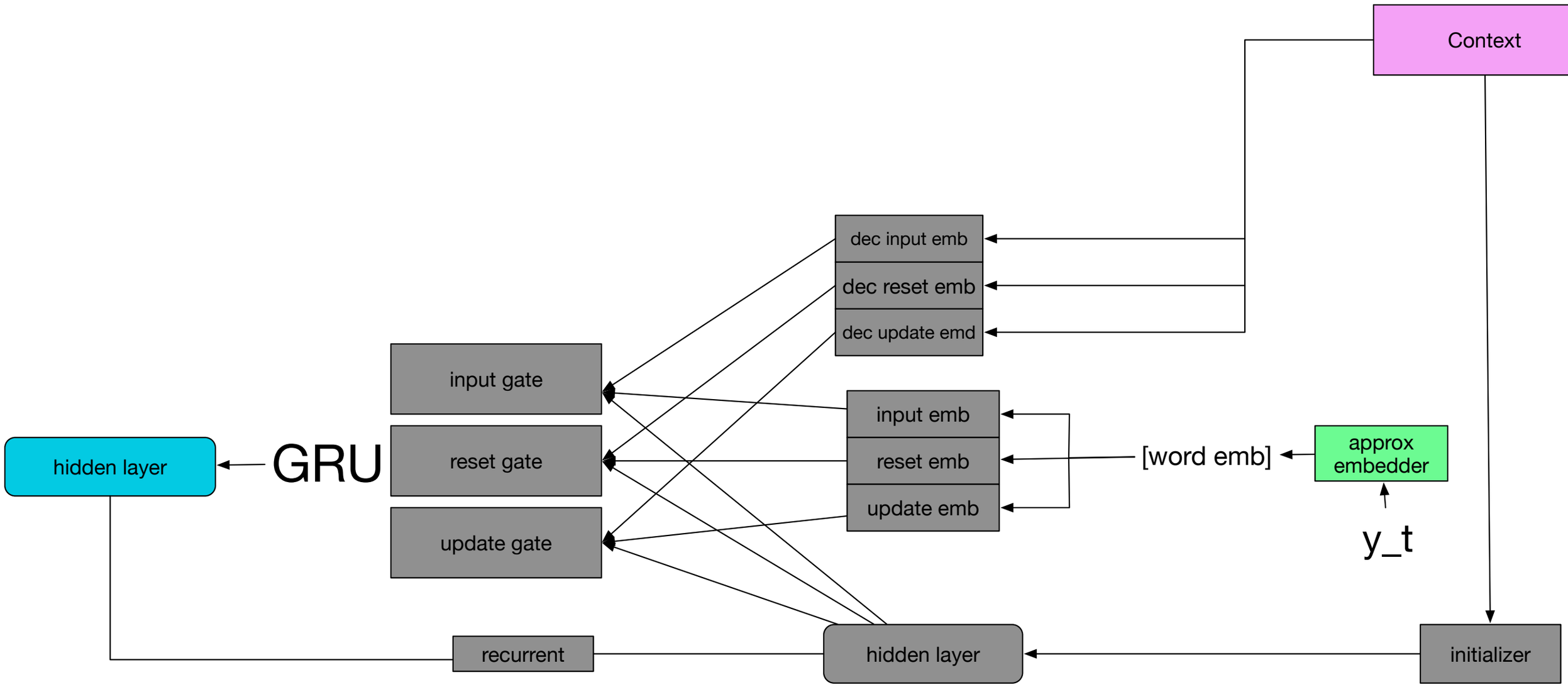
From Encoder to Context



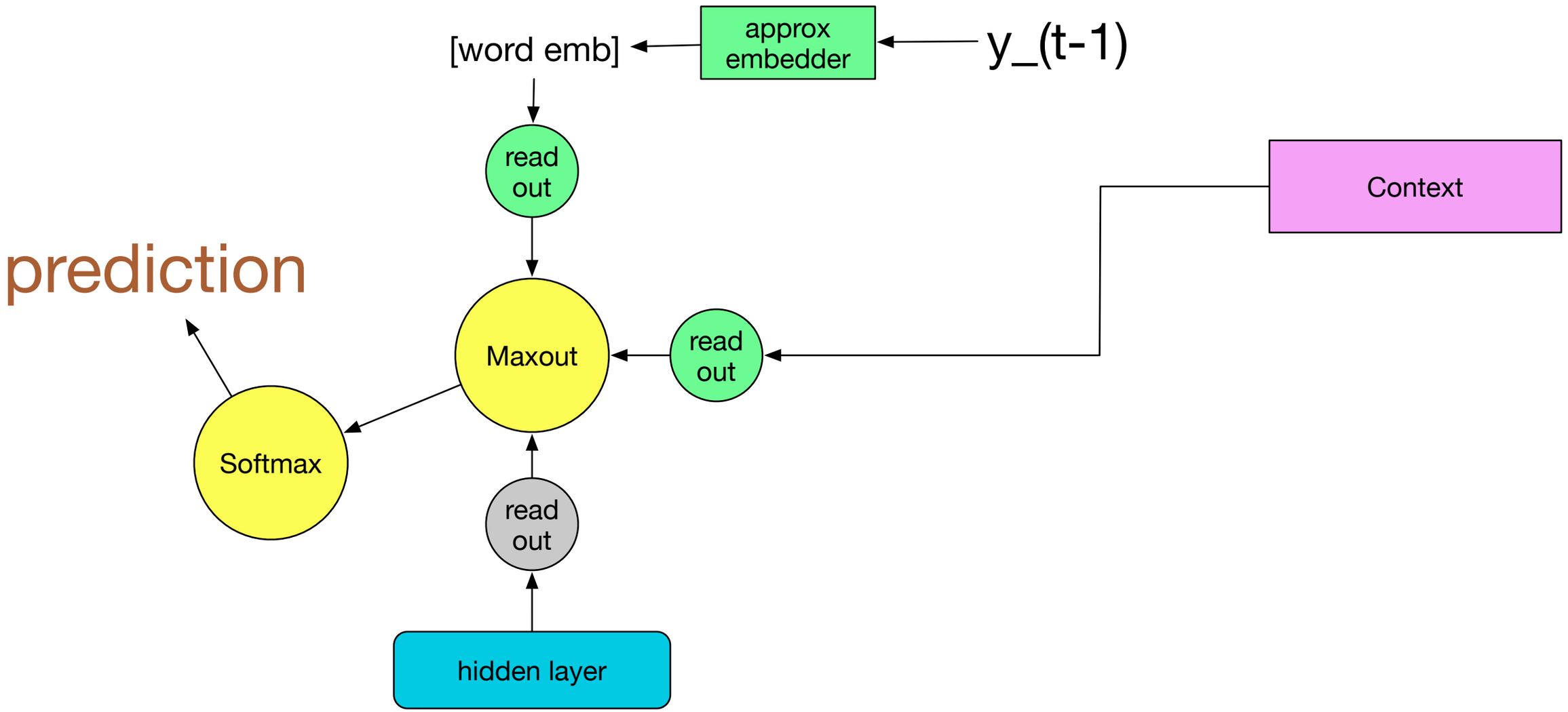
Decoder

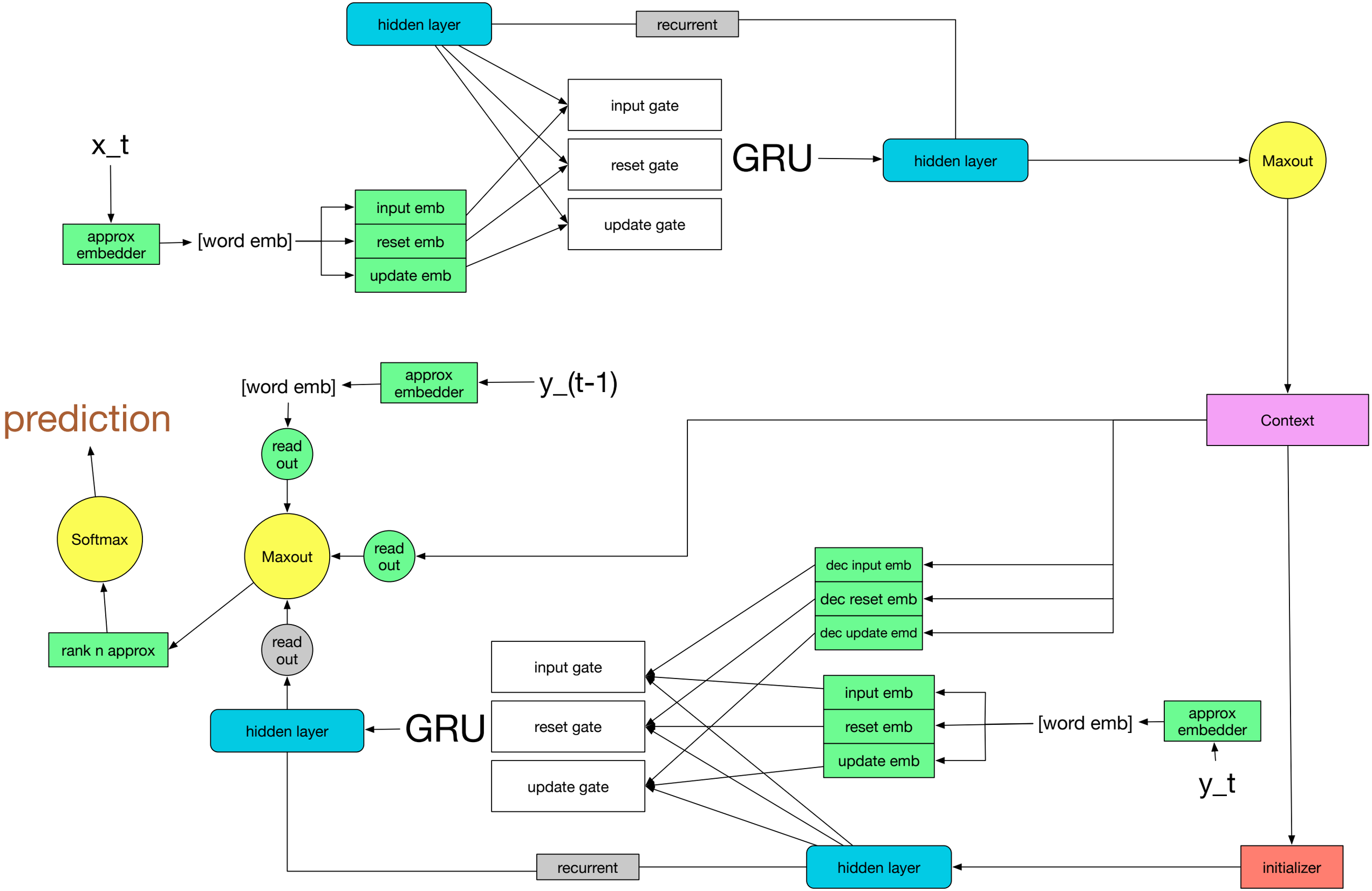

Context
dec input emb
dec reset emb
dec update emd
input gate
reset gate
update gate
input emb
reset emb
update emb
[word emb]
approx embedder
y_t
hidden layer
GRU
recurrent
hidden layer
initializer

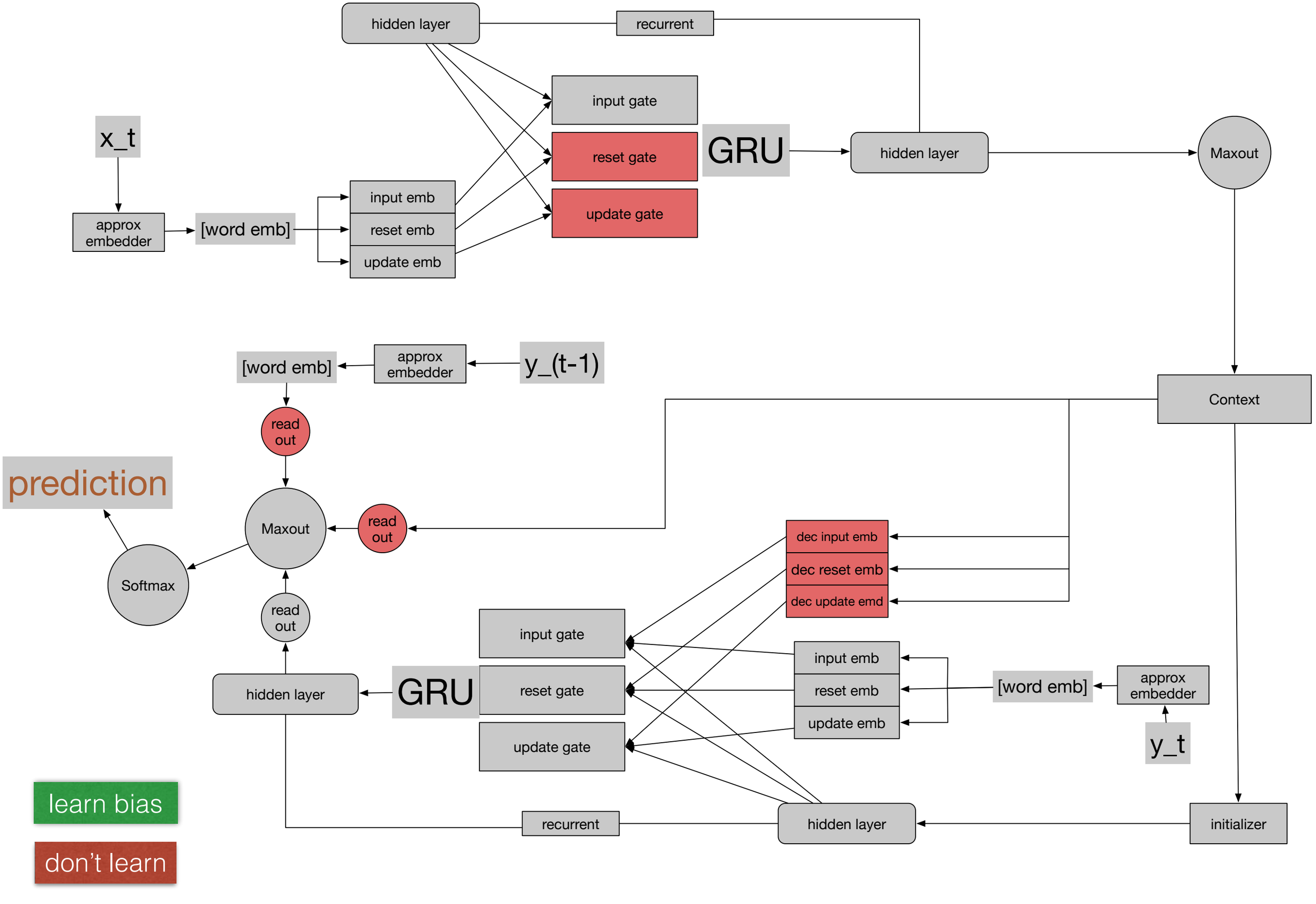
From Decoder to Output



From Decoder to Output




hidden layer
recurrent
input gate
reset gate
update gate
x_t
GRU
hidden layer
Maxout
approx embedder
[word emb]
input emb
reset emb
update emb
[word emb]
approx embedder
y_(t-1)
Context
prediction
read out
Softmax
Maxout
read out
dec input emb
dec reset emb
dec update emd
rank n approx
read out
input gate
input emb
hidden layer
GRU
reset gate
reset emb
[word emb]
approx embedder
update gate
update emb
y_t
recurrent
hidden layer
initializer


hidden layer
recurrent
input gate
reset gate
update gate
GRU
hidden layer
Maxout
x_t
approx embedder
[word emb]
input emb
reset emb
update emb
[word emb]
approx embedder
y_(t-1)
Context
read out
prediction
Maxout
read out
Softmax
read out
dec input emb
dec reset emb
dec update emd
input gate
reset gate
update gate
GRU
hidden layer
input emb
reset emb
update emb
[word emb]
approx embedder
y_t
learn bias
don't learn
recurrent
hidden layer
initializer

Encoder hidden states

h1 h2 … hT

Alignments (tanh)

Softmax

x

hidden t-1

Averaged Context t

read out

Output …

read out

read out

dec input emb

dec reset emb

dec update emd

input gate

reset gate

update gate

input emb

reset emb

update emb

[word emb]

approx embedder

y_t

hidden layer

GRU

recurrent

hidden layer

initializer

Variations & Extensions

- Replace GRU with LSTM
    - *Sequence to Sequence Learning with Neural Networks*
- Translate image to language
    - *Show and Tell: A Neural Image Caption Generator*
    - *Show, Attend and Tell: Neural Image Caption Generation with Visual Attention*