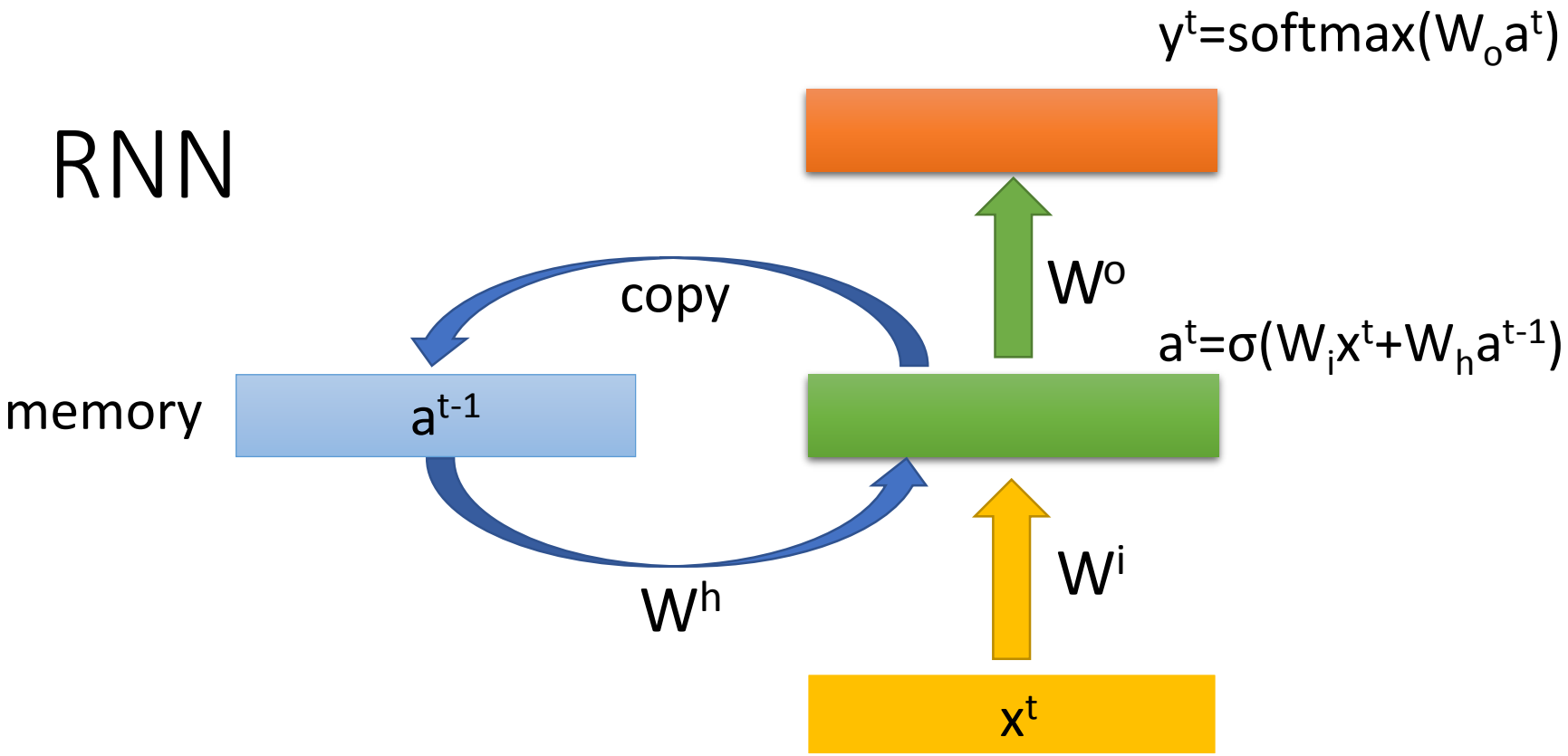
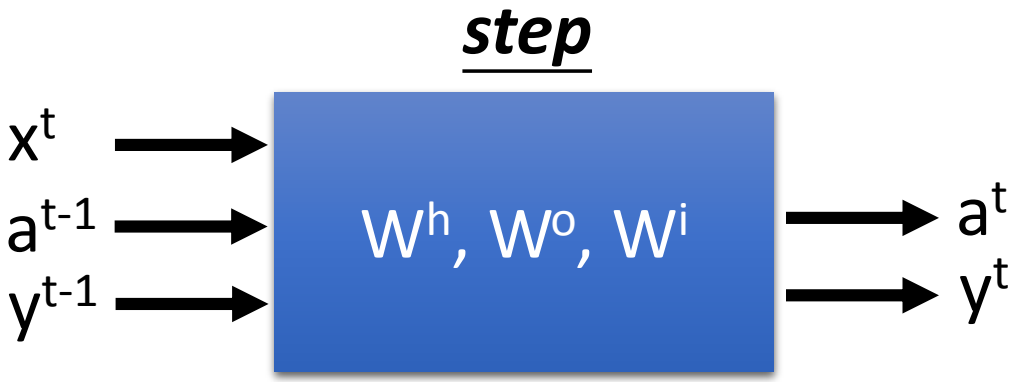
Introduction of Theano: scan

Hung-yi Lee

http://deeplearning.net/software/theano/library/scan.html

RNN

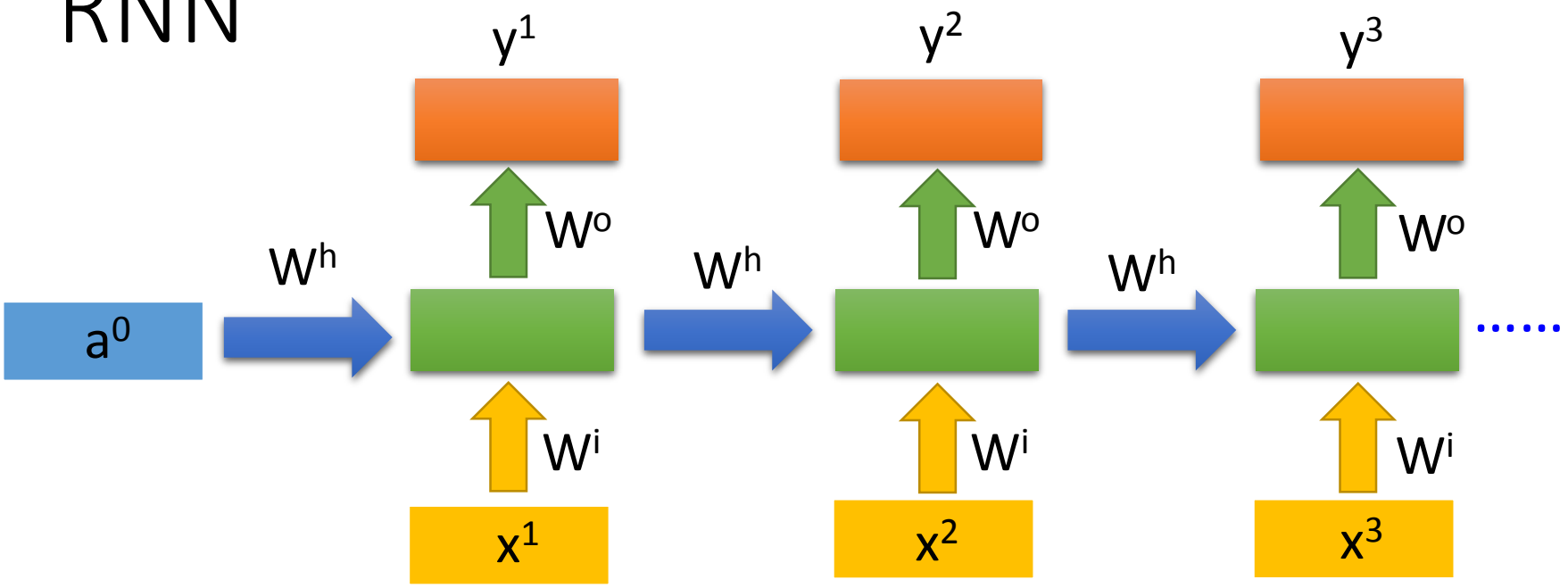

$y^t=softmax(W_o a^t)$
W^o
$a^t=\sigma(W_i x^t+W_h a^{t-1})$
copy
memory
a^{t-1}
W^h
W^i
x^t

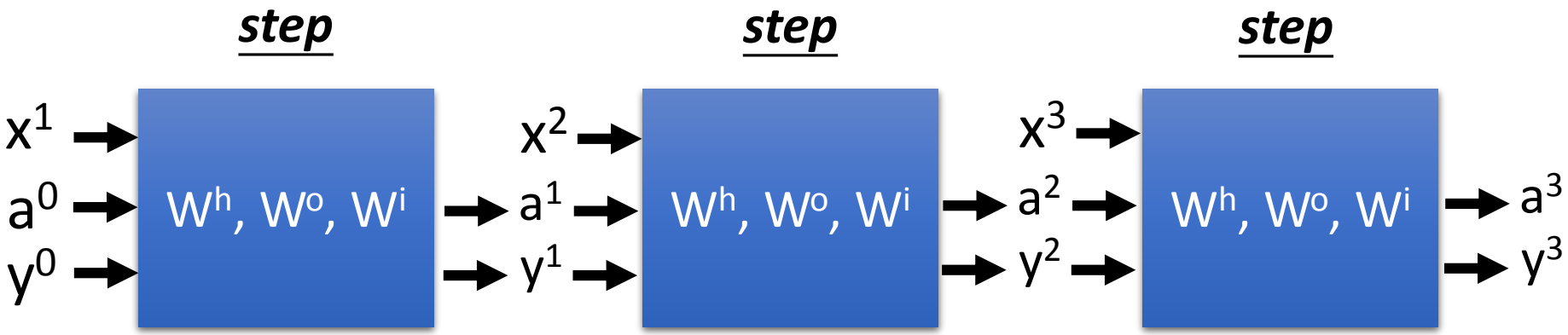

step
x^t
a^{t-1}
y^{t-1}
W^h, W^o, W^i
a^t
y^t

RNN


```
def step(x_t,a_tm1,y_tm1):
        a_t = sigmoid( T.dot(x_t,Wi) \
                    + T.dot(a_tm1,Wh) + bh )
        y_t = softmax( T.dot(a_t,Wo) + bo)
        return a_t, y_t
```

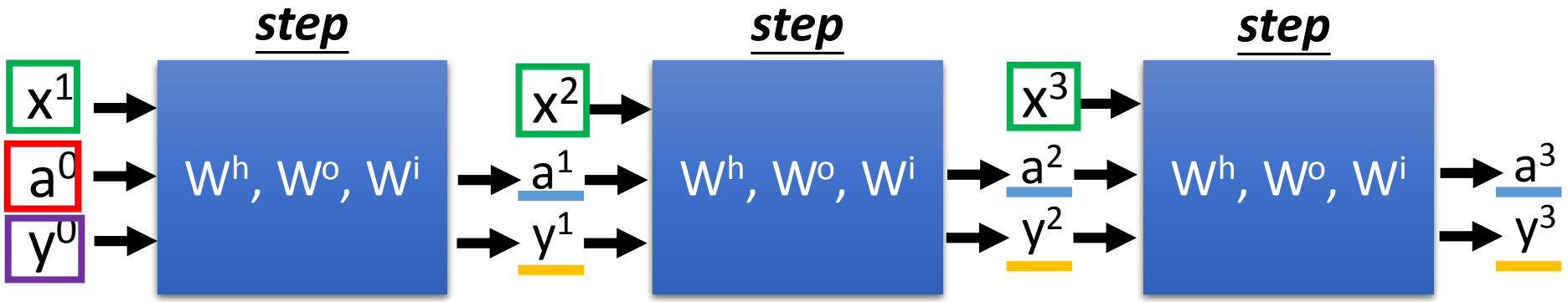
RNN


y1
y2
y3
Wo
Wo
Wo
Wh
Wh
Wh
a0
Wi
Wi
Wi
x1
x2
x3
.....

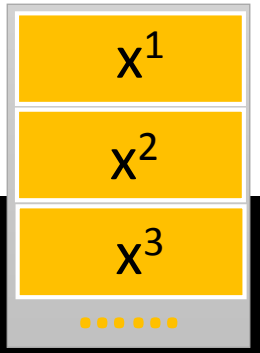

step
step
step
x1
a0
y0
Wh, Wo, Wi
a1
y1
x2
Wh, Wo, Wi
a2
y2
x3
Wh, Wo, Wi
a3
y3

theano.scan

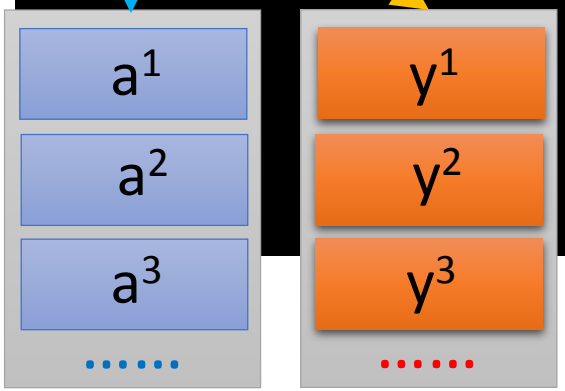
```
def step(x_t,a_tm1,y_tm1):
        a_t = sigmoid( T.dot(x_t,Wi) \
                + T.dot(a_tm1,Wh) + bh )
        y_t = softmax( T.dot(a_t,Wo) + bo)
        return a_t, y_t
```



Matrix:



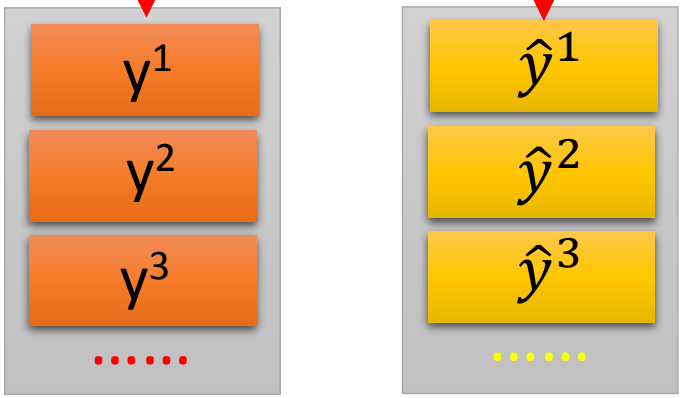

```
[a_seq,y_seq],_ = theano.scan(
                        step,
                        sequences = x_seq,
                        outputs_info = [ a_0, y_0 ],
                        truncate_gradient=-1
)
```



a_0, y_0 are shared variables

Computing Cost & Gradients

```
cost = T.sum( ( y_seq - y_hat_seq )**2 )
```



W^h, W^o, W^i .....

```
gradients = T.grad(cost,parameters)
```

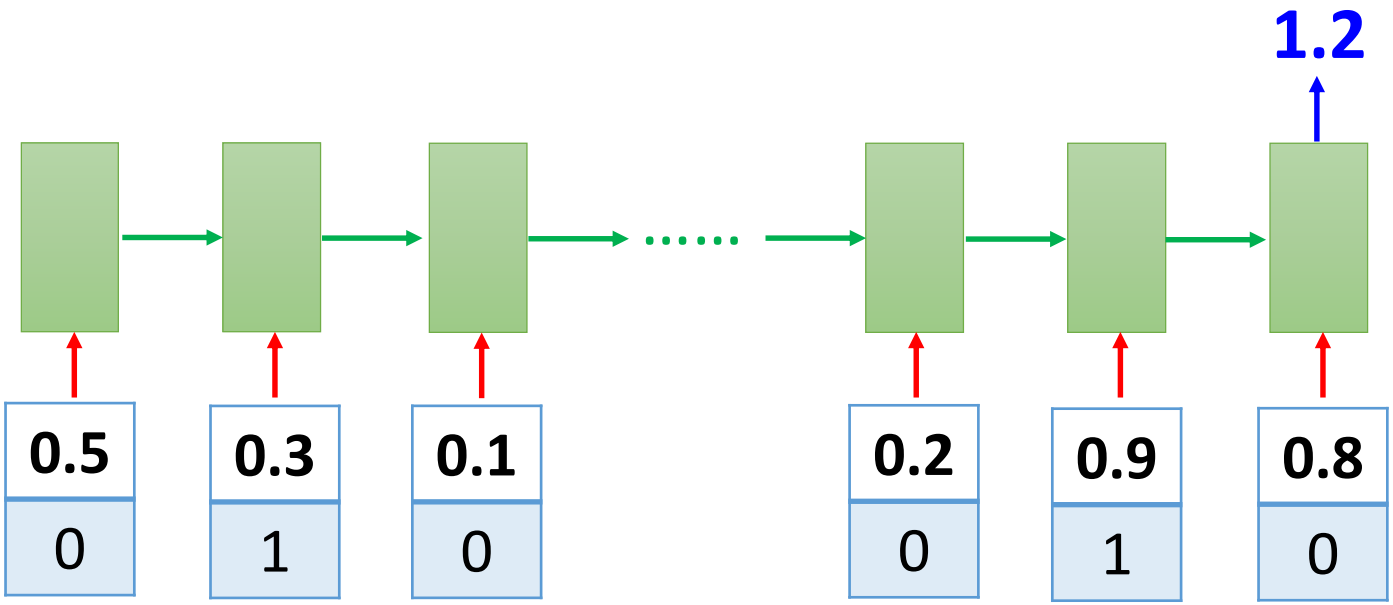
BPTT

That's it

```
rnn_train = theano.function(
        inputs=[x_seq,y_hat_seq],
        outputs=cost,
        updates=MyUpdate(parameters,gradients)
)
```

```
for i in range(10000000):
        x_seq, y_hat_seq = gen_data()
        print rnn_train(x_seq,y_hat_seq)
```

Example



Example:
http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/theano/rnn.example.py

Modified from
https://github.com/Lasagne/Lasagne/blob/master/examples/recurrent.py

Share your idea on FB Group

- [Theano-RNN Q1] Any idea to make RNN faster with GPU?
- [Theano-RNN Q2] How to implement bi-directional RNN?
- [Theano-RNN Q3] Can you implement “batch of sequences”?
  - [Theano-RNN Q3-1] All the sequences in a batch have the same length.
  - [Theano-RNN Q3-2] The sequences in a batch have different lengths.