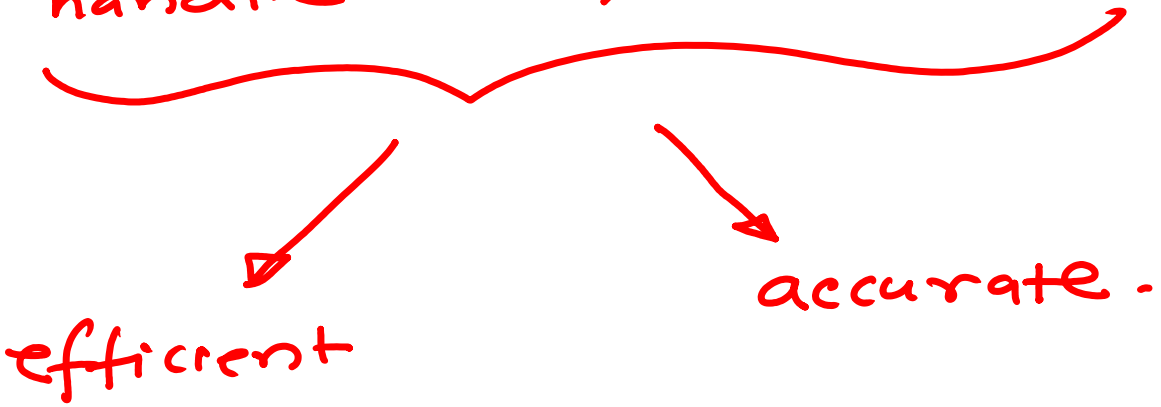
RNN → Recurrent neural network

RNN cell :-

Sequence learning :- deals with sequential data.

Example of sequential data :- Writing / Language, time-series (stock market)

* The idea in sequence learning is to develop models (NN models) that can handle sequential data.



* RNN is the most basic form of sequence learning model.

* Other sequence learning models include

(1) LSTM ← Long Short Term Memory.

(2) GRU ← Gated Recurrent unit

(3) Transformer → based on Attention

(4) MAMBA → based on state space model

* Different problem statement in Sequence learning :-

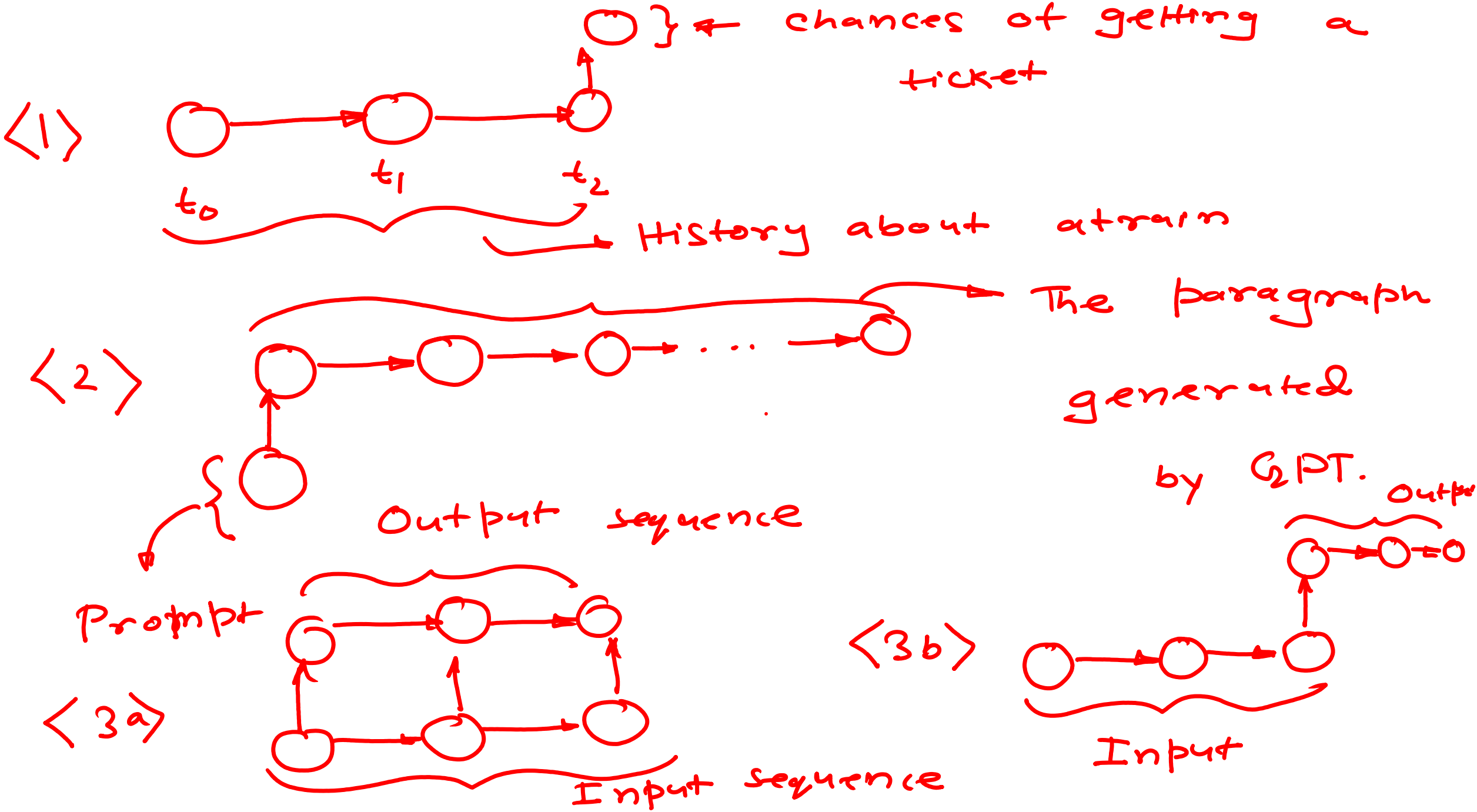
(i) Sequential → Scalar / vector

Input → output.

(ii) Scalar vector → sequential

(iii) Sequence to sequence learning.

- Input sequence is processed one at a time.
- The total input sequence is processed before it yield the output.


chances of getting a ticket
<1>
t_0
t_1
t_2
History about atrain
<2>
The paragraph generated by GPT.
Output sequence
Prompt
<3a>
Input sequence
<3b>
Input

* Recurrent Neural Network :-

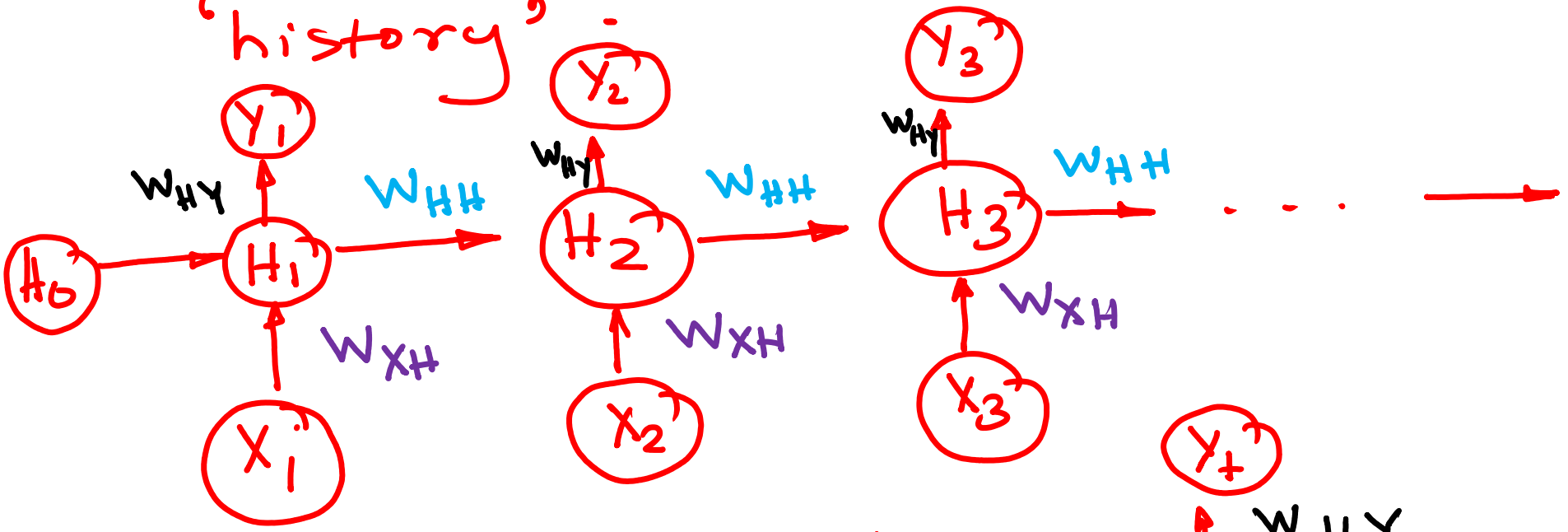
RNNs have three components :-

<1> Input node
<2> Hidden node
<3> Output node

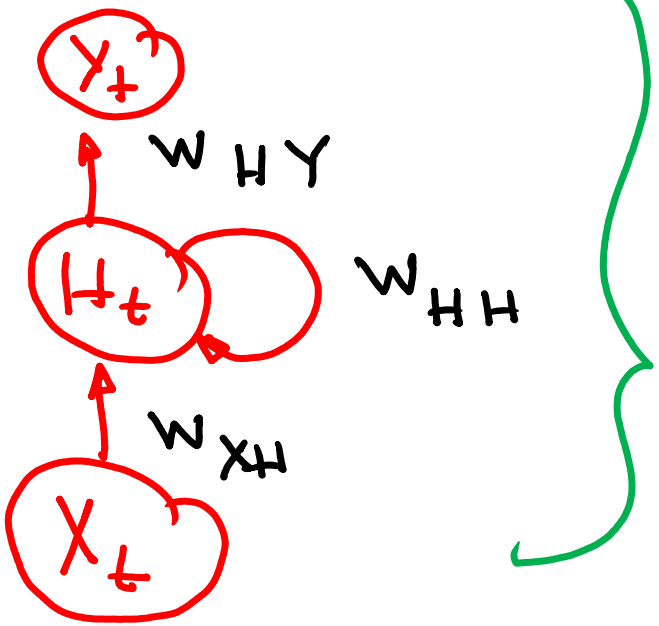
} ← These components are same as ANN

* RNN suggests a specific way of connecting input to output via hidden nodes.

* The hidden nodes are responsible for 'history'.

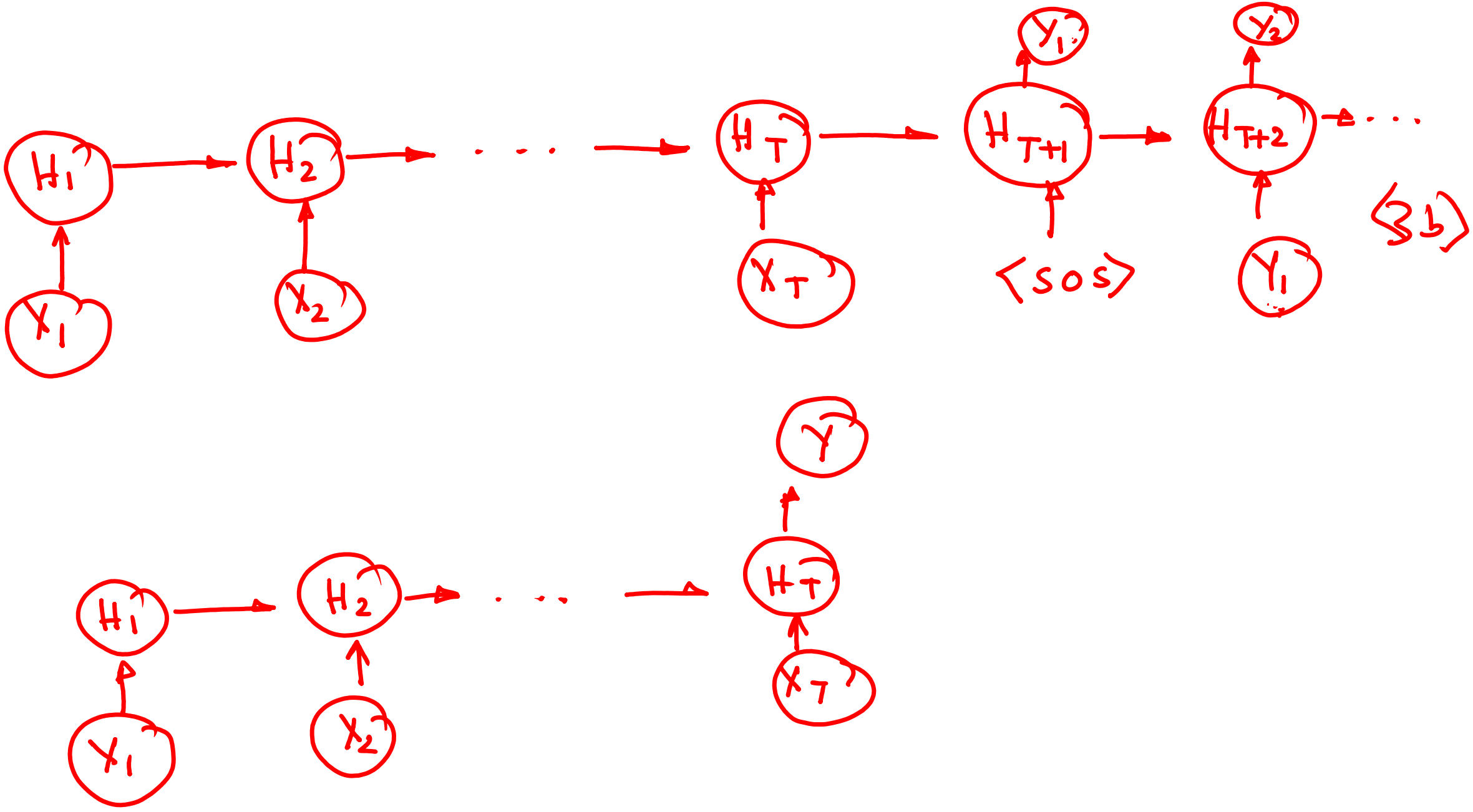


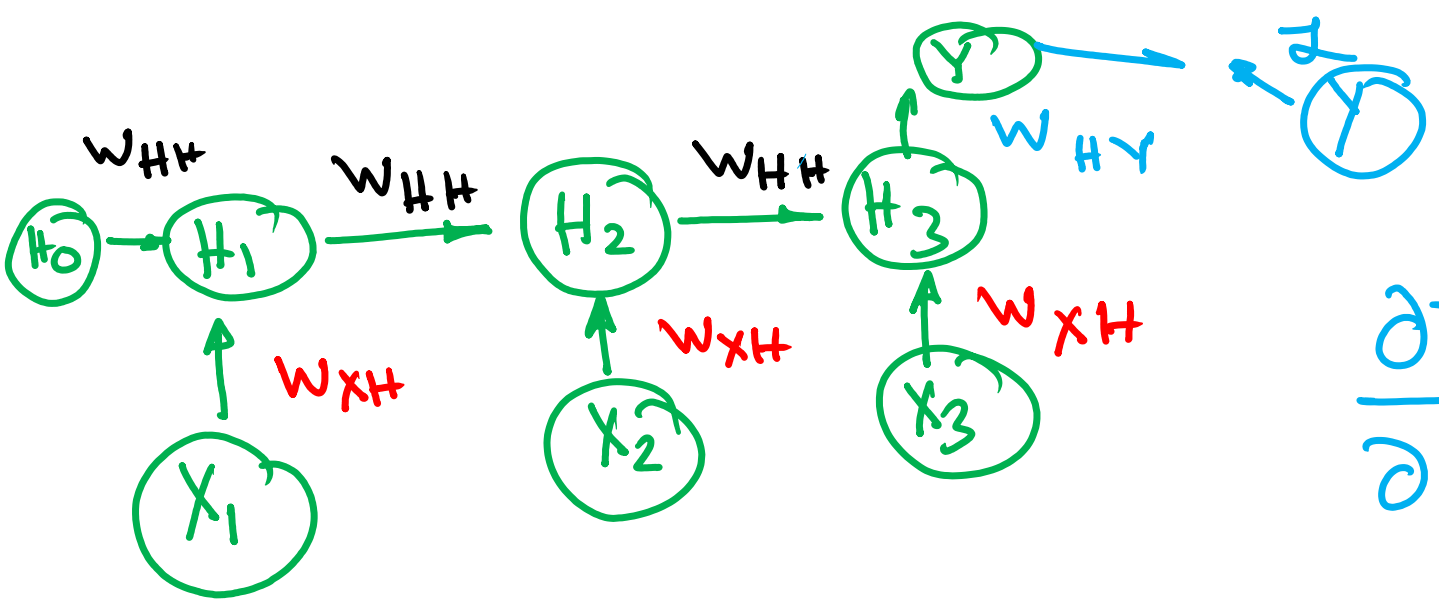
(3 of)



$$H_t = \sigma\left(W_{XH} X_t + W_{HH} H_{t-1}\right)$$

$$Y_t = \sigma\left(W_{HY} H_t\right)$$


H1
H2
HT
HT+1
HT+2
X1
X2
XT
Y1
Y2
<SOS>
Y1
<SOS>
Y
H1
H2
HT
X1
X2
XT



$$\frac{\partial L}{\partial W_{HY}}, \quad \frac{\partial L}{\partial W_{HH}}, \quad \frac{\partial L}{\partial W_{XH}}$$

$$\frac{\partial L}{\partial W_{HY}} = \frac{\partial}{\partial W_{HY}} \left(Y - \tilde{Y} \right)^2 \qquad \tilde{Y} = \sigma\left(W_{HY} H_3 \right)$$

$$= -2 \left(Y - \tilde{Y} \right) \frac{\partial \tilde{Y}}{\partial W_{HY}} = -2 \left(Y - \tilde{Y} \right) \sigma' \left(W_{HY} H_3 \right)$$

$$\frac{\partial L}{\partial W_{xH}} = \frac{\partial L}{\partial Y} \frac{\partial Y}{\partial H_3} \frac{\partial H_3}{\partial W_{xH}} +$$

$$\frac{\partial L}{\partial Y} \cdot \frac{\partial Y}{\partial H_3} \cdot \frac{\partial H_3}{\partial H_2} \cdot \frac{\partial H_2}{\partial W_{xH}} +$$

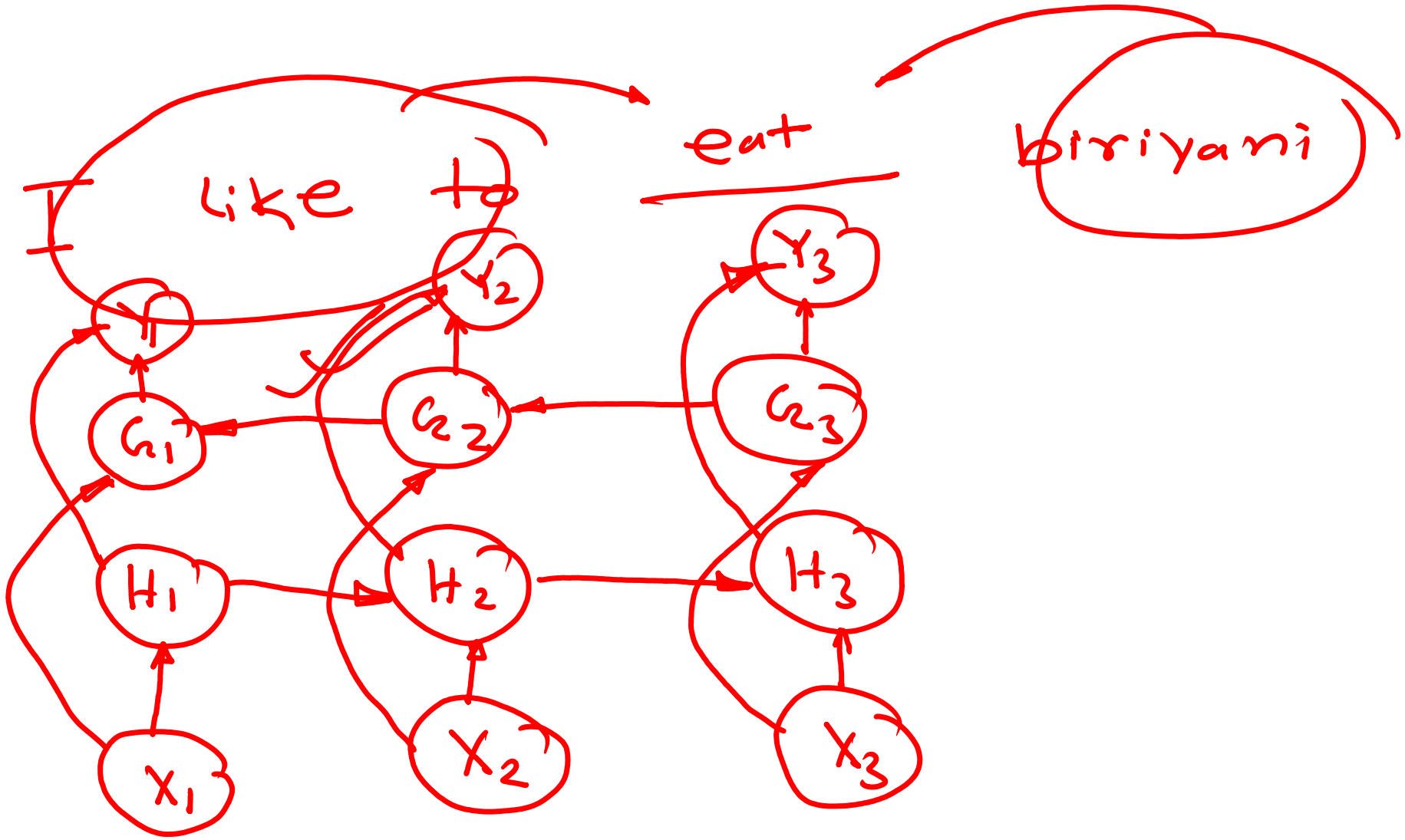
$$\frac{\partial L}{\partial Y} \cdot \frac{\partial Y}{\partial H_3} \cdot \frac{\partial H_3}{\partial H_2} \cdot \frac{\partial H_2}{\partial H_1} \cdot \frac{\partial H_1}{\partial W_{xH}}$$

* RNNs are trained using BPTT

↓

Backpropagation through time.

* Bidirectional RNN :-



- There was a king X.
- There was another king Y.
- X and Y were from enemy kingdoms
- X and Y were both powerful and they went into a war with each other.
- Y killed X

* X had a son Jr X.

① Jr X grew up to be a powerful king and attacked Y.

② Y killed Jr X.

③ Jr X had a son S Jr X.

④ He grew up and attacked Y and killed Y.

	Power	Conflict	Revenge
S1)	0.6	0	0
S2)	0.75	0.1	0
S3)	0.75	0 8	0
S4)	0.9	0.9	0
S5)	0. 6-	0.7	0.1
S6)	0.61	0.72	0.4