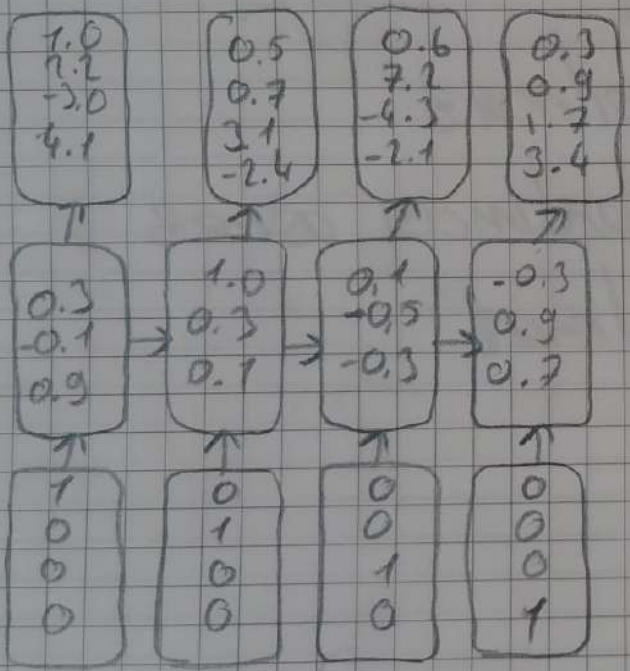
Karphaty RNNs

What are RNNs

Traditional NNs are static by input, output or layer size.
While RNNs works with sequences.

If training vanilla NNs are optimization over function, training RNNs are optimization over programs

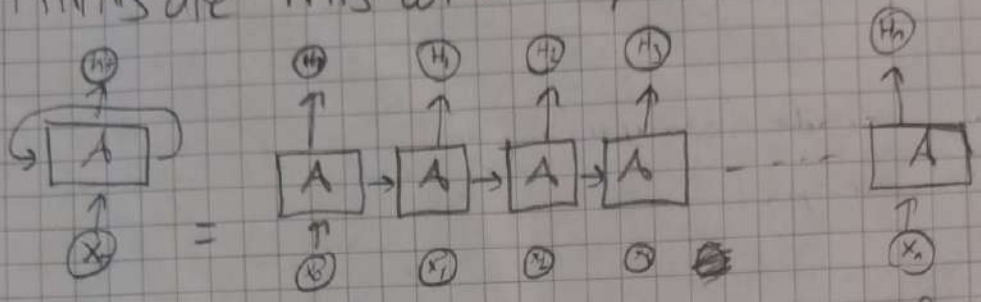


RNNs works in a recursive way. It inputs the current element with the previous elements at each step.

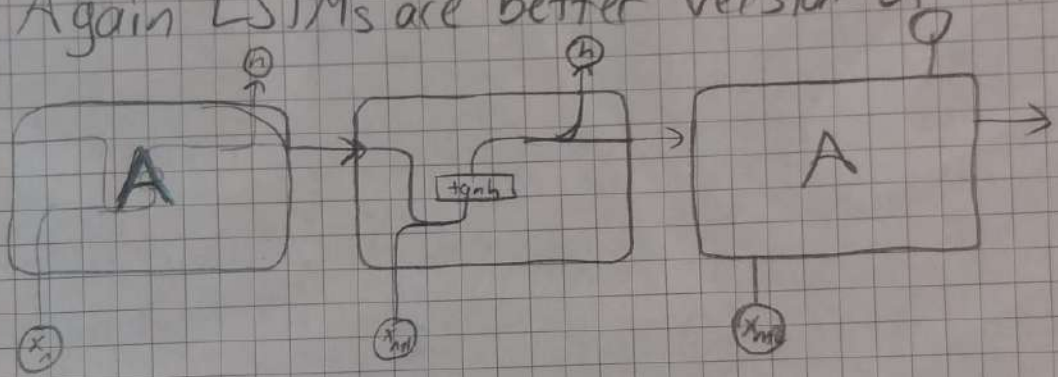
LSTM is ~~ed~~ enhanced model of RNN. Handles long-term memory.

Understanding LSTM Networks

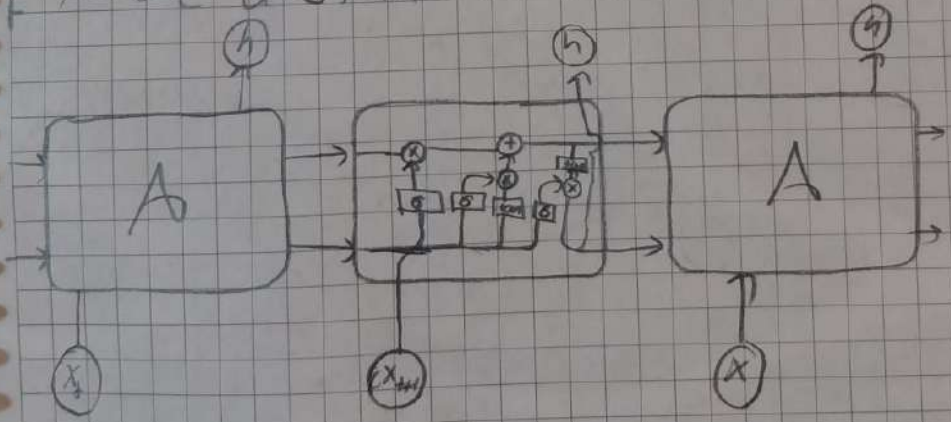
RNNs are NNs with loop.



Again LSTMs are better version of RNNs.



Above a Standard RNN



First sigmoid is "should we forget?" function deciding to forget or not the previous cell state.

Second sigmoid is what we keep in mind from current input function.

Last sigmoid is concluding the y from tanh of cell state and importance of input x.

Neural Image Caption Generation with Visual Attention

2 attention based image captioning introduced:
1) "soft" attention mechanism with back propagation
2) "hard" attention mechanism with REINFORCE

Q1) What is REINFORCE rule?
Q2) How encode-decode works.

Decoder selects the 'focus' areas by looking to lower CNN layers

Q3) What is Monte Carlo based sampling.
Q4) What is multinoulli distribution