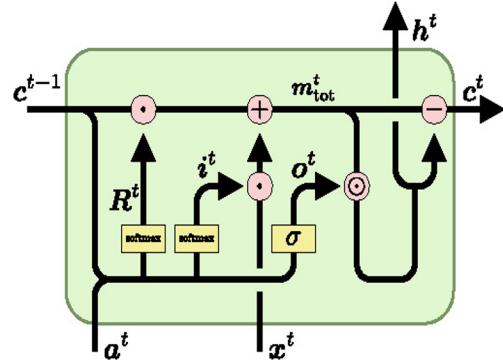


Seminar 8: RNN

Actually, AI does not exist. It is just ML.



99% toxic comment

Byte Pair Encoding

Tokenization

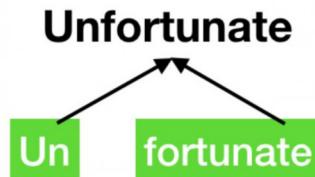
Word-level

- fixed vocabulary
- can process only a fixed number of words

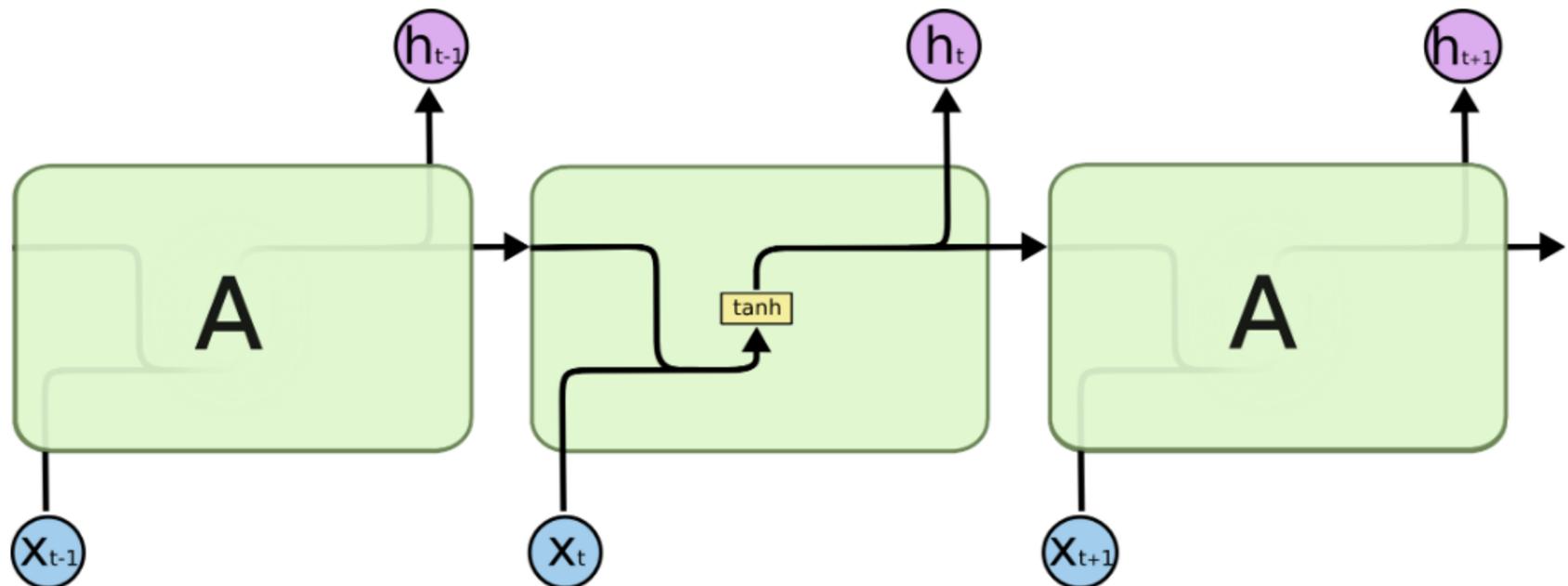
Subword-level

- open vocabulary
- rare and unknown tokens are encoded as sequences of subword units

Unfortunately

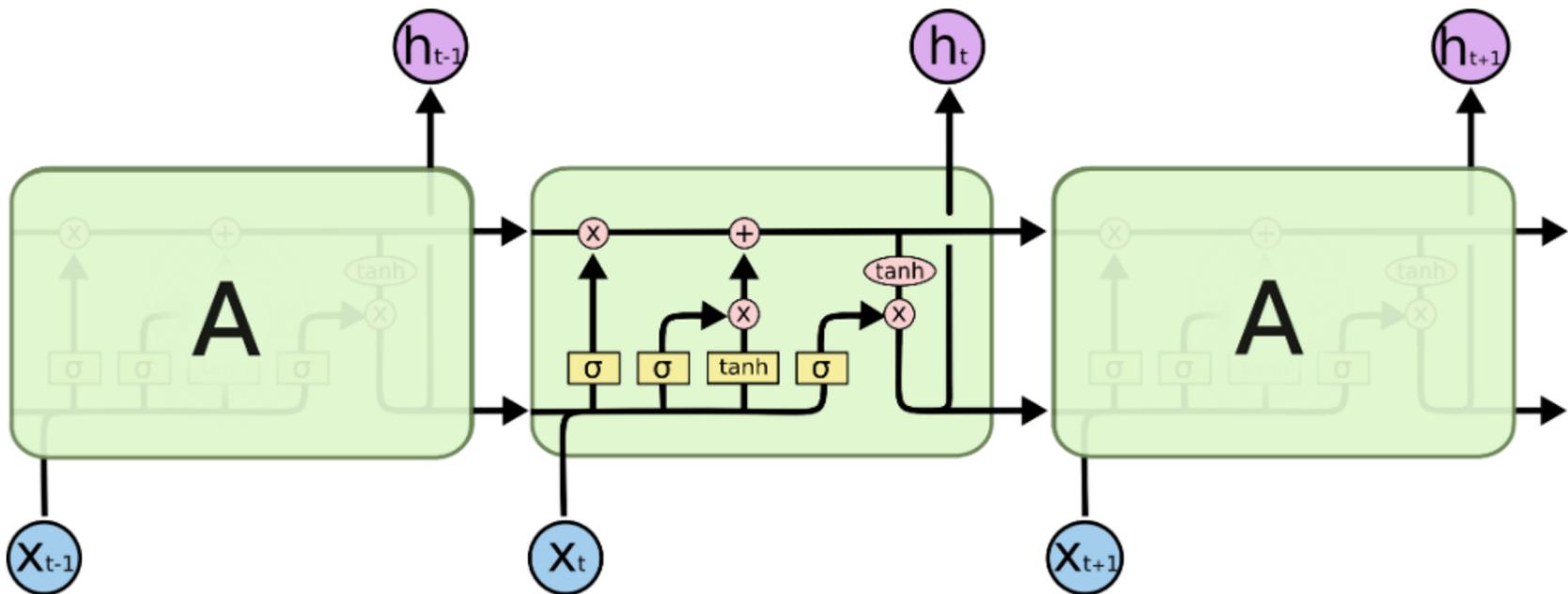


RNN



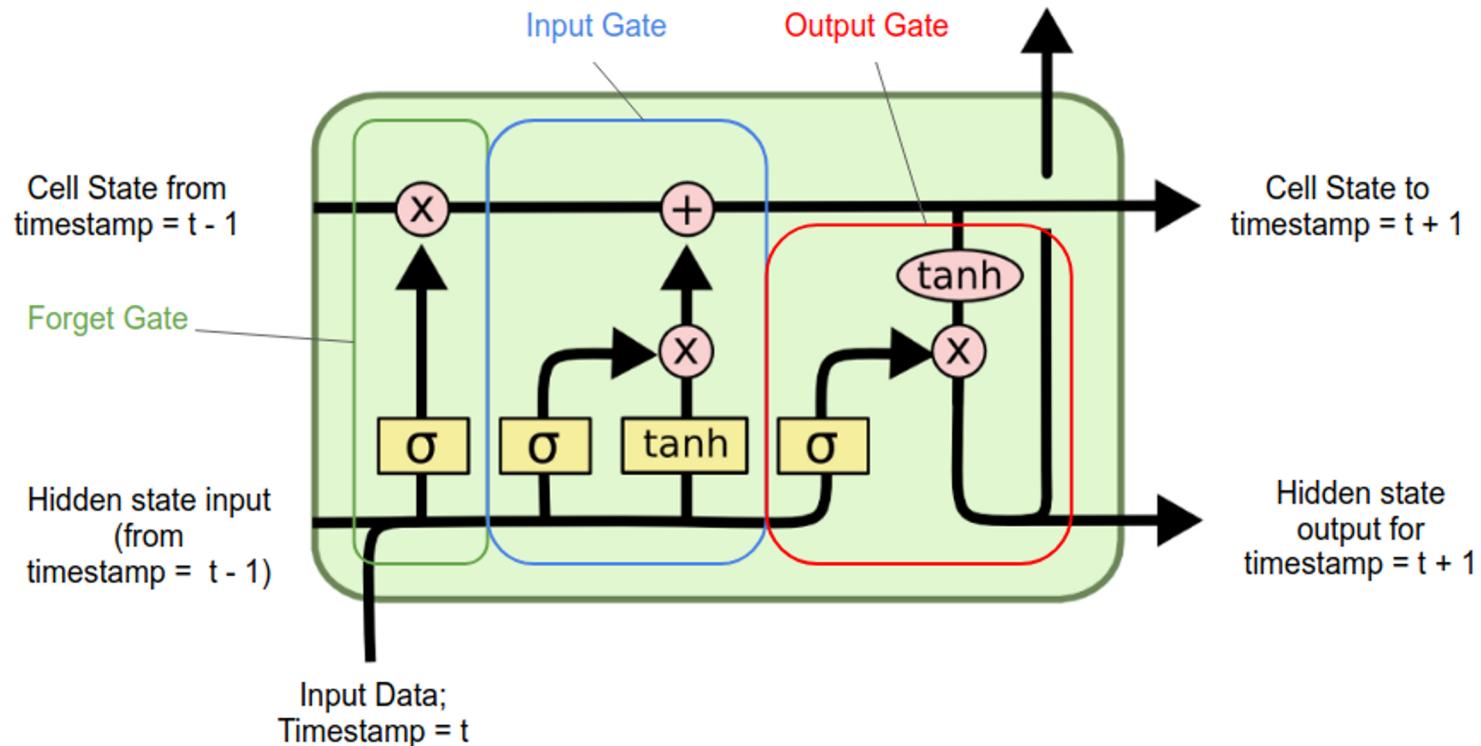
The repeating module in a standard RNN contains a single layer.

LSTM

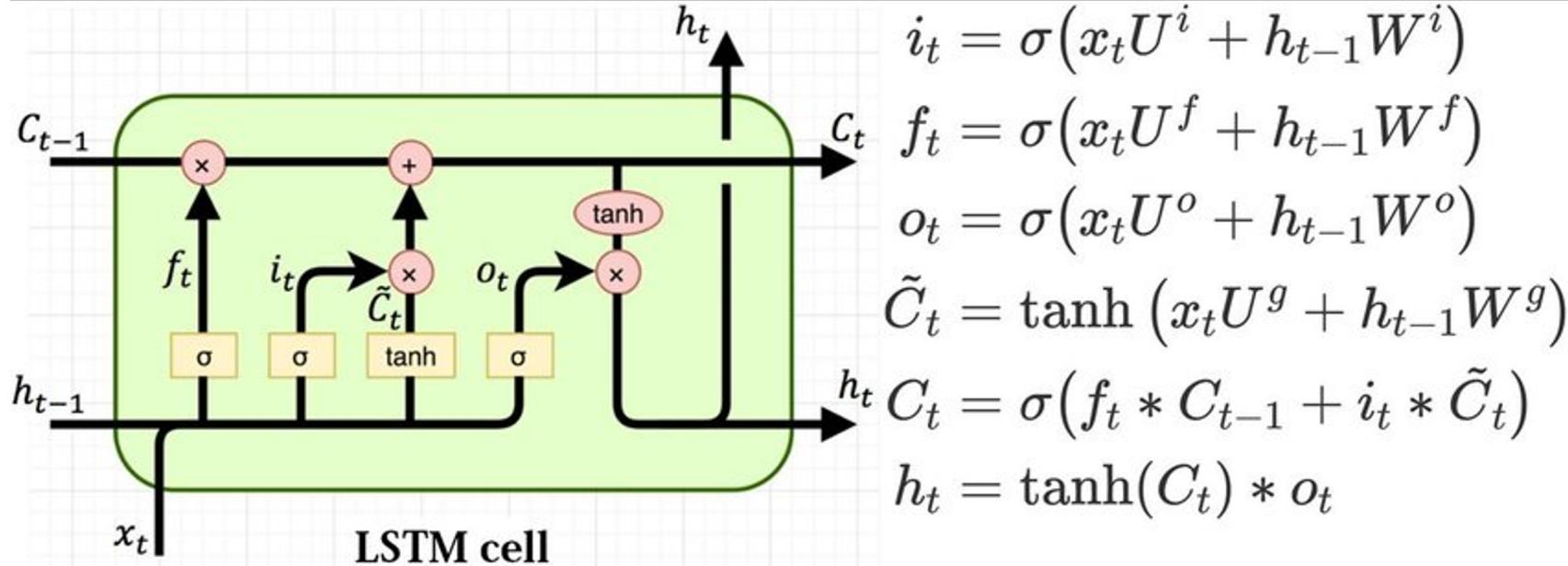


The repeating module in an LSTM contains four interacting layers.

LSTM

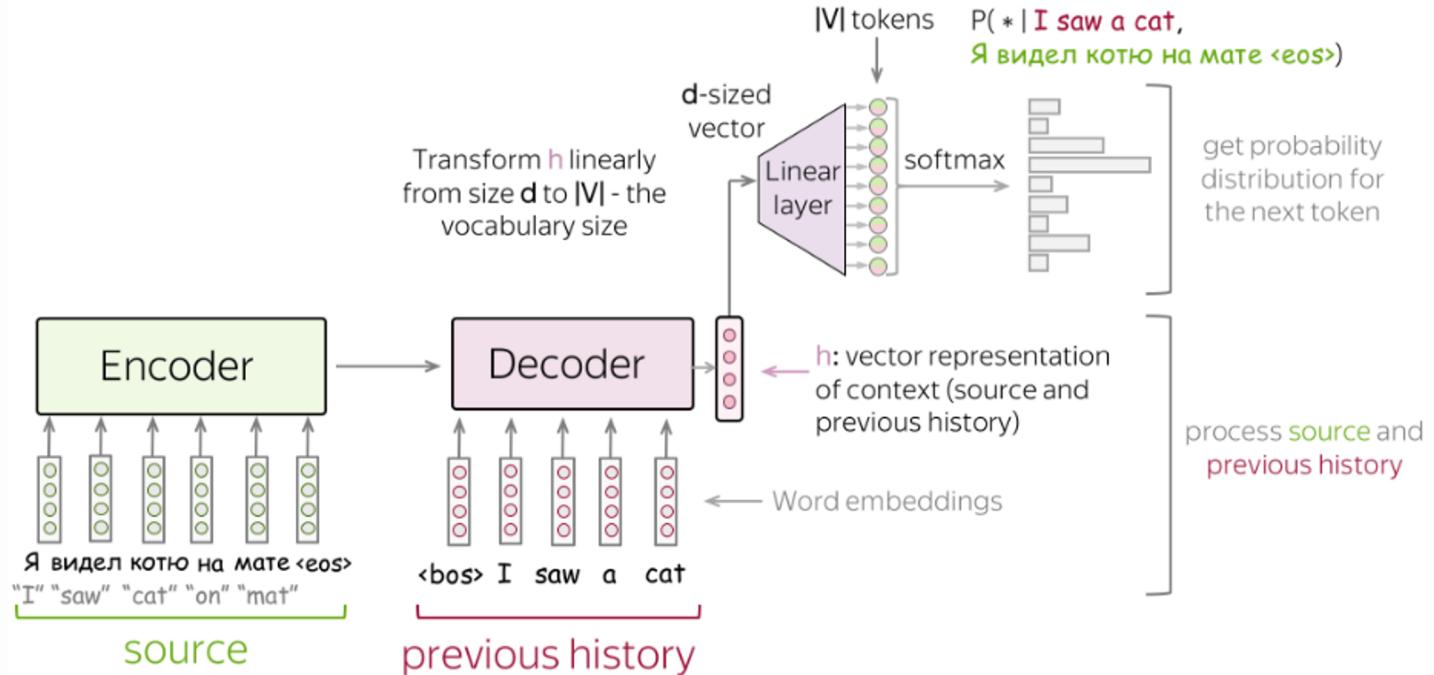


LSTM

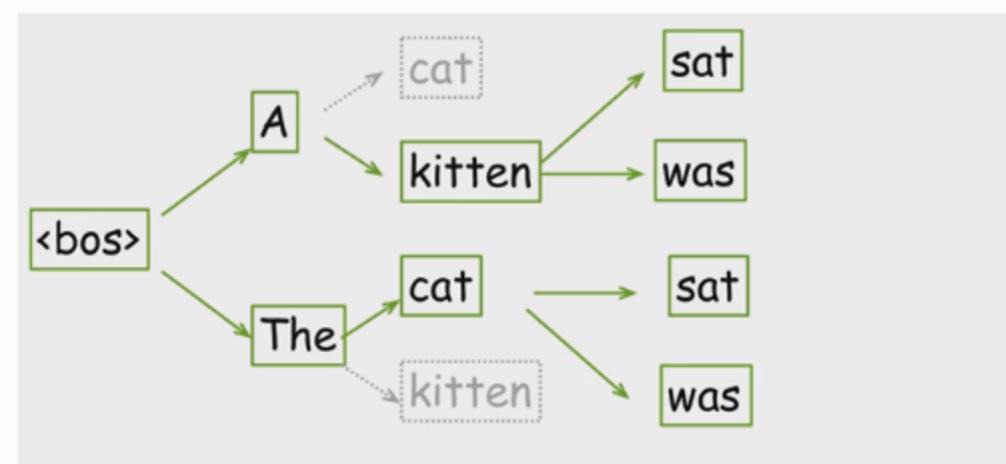


LSTM

seq2seq



Beam Search



https://lena-voita.github.io/nlp_course.html#preview_seq2seq_attn