UNIVERSITÄT BONN

Sequence Processing

Moritz Wolter

September 25, 2022

Uni Bonn

Overview

Recurrent neural networks

Applications

Motivation

- Thus far we have never integrated information over time.
- We want the ability to create internal memory.
- Consider the sentence: I live in Paris. I speak ...
- ... French.
- Clearly it is likely for someone in Paris to speak French.
- Memory should help networks taking Paris into account when deciding what language is spoken.

Recurrent neural networks

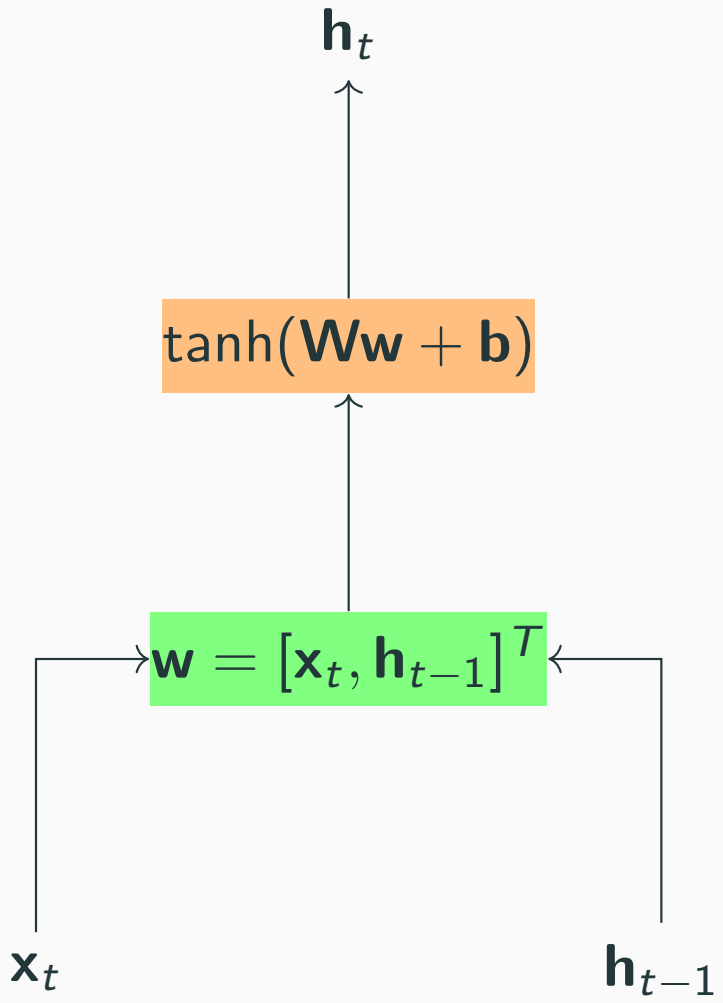
Elman-recurrent neural networks

A simple solution is to add a state to the network and feed this state recurrently back into the network [Elm90],

$$\overline{\mathbf{h}_t} = \mathbf{W}_h \mathbf{h}_t + \mathbf{W}_x \mathbf{x}_t + \mathbf{b}, \tag{1}$$

$$\mathbf{h}_{t+1} = f(\overline{\mathbf{h}_t}). \tag{2}$$

Elman-recurrent neural networks



Unrolling in Time

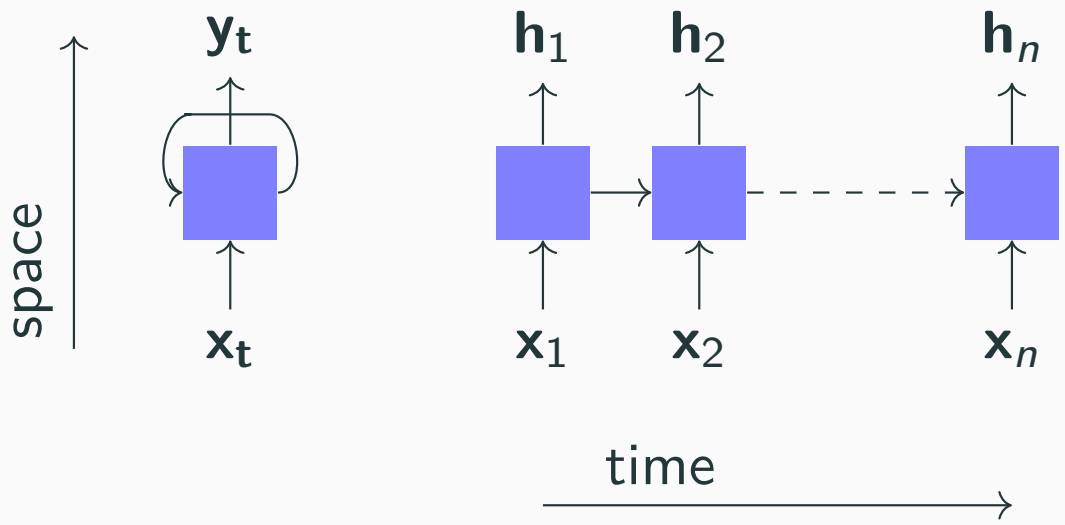


Figure: The rolled (left) cell can be unrolled (right) by considering all inputs it saw during the current gradient computation iteration.

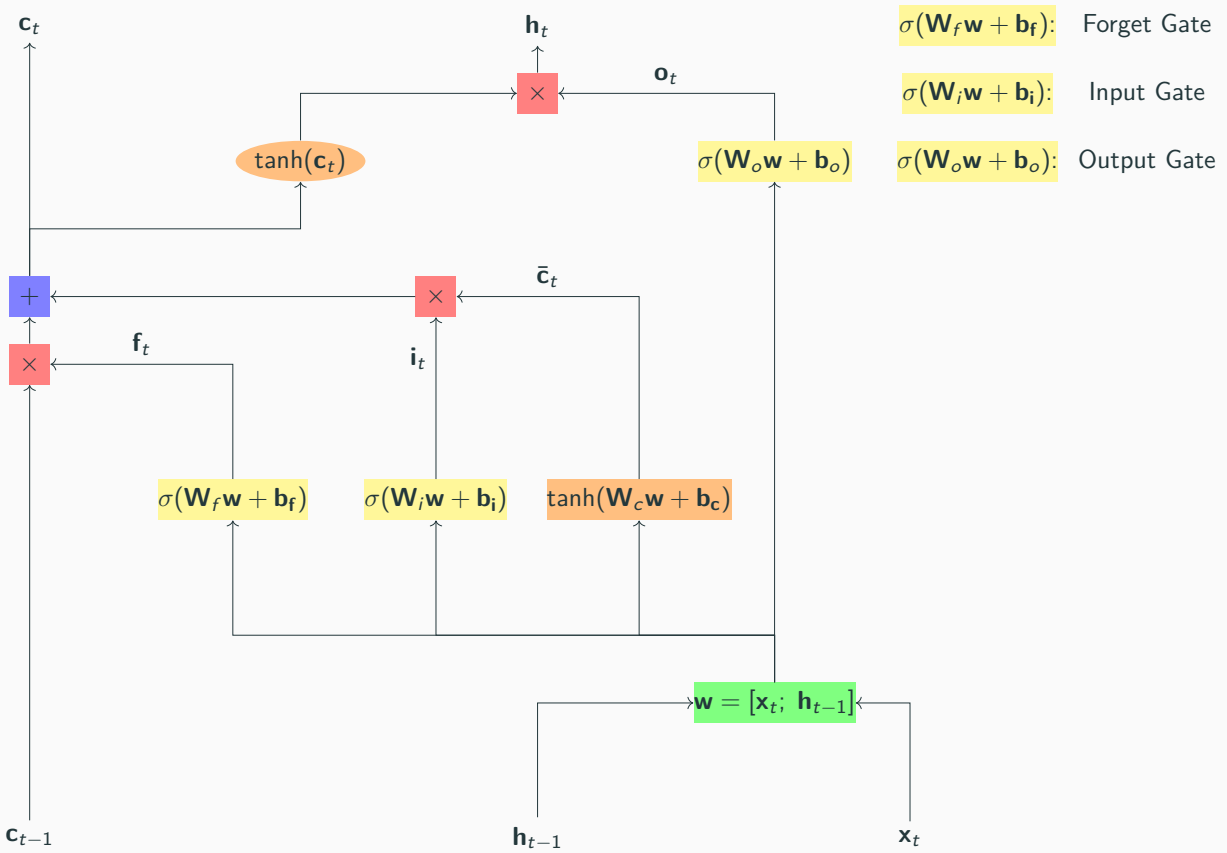
Stability of recurrent connections

For an intuition. Consider a linear network without activations or inputs.

$$\mathbf{h}_{t+1} = \mathbf{W}_h \mathbf{h}_t \tag{3}$$

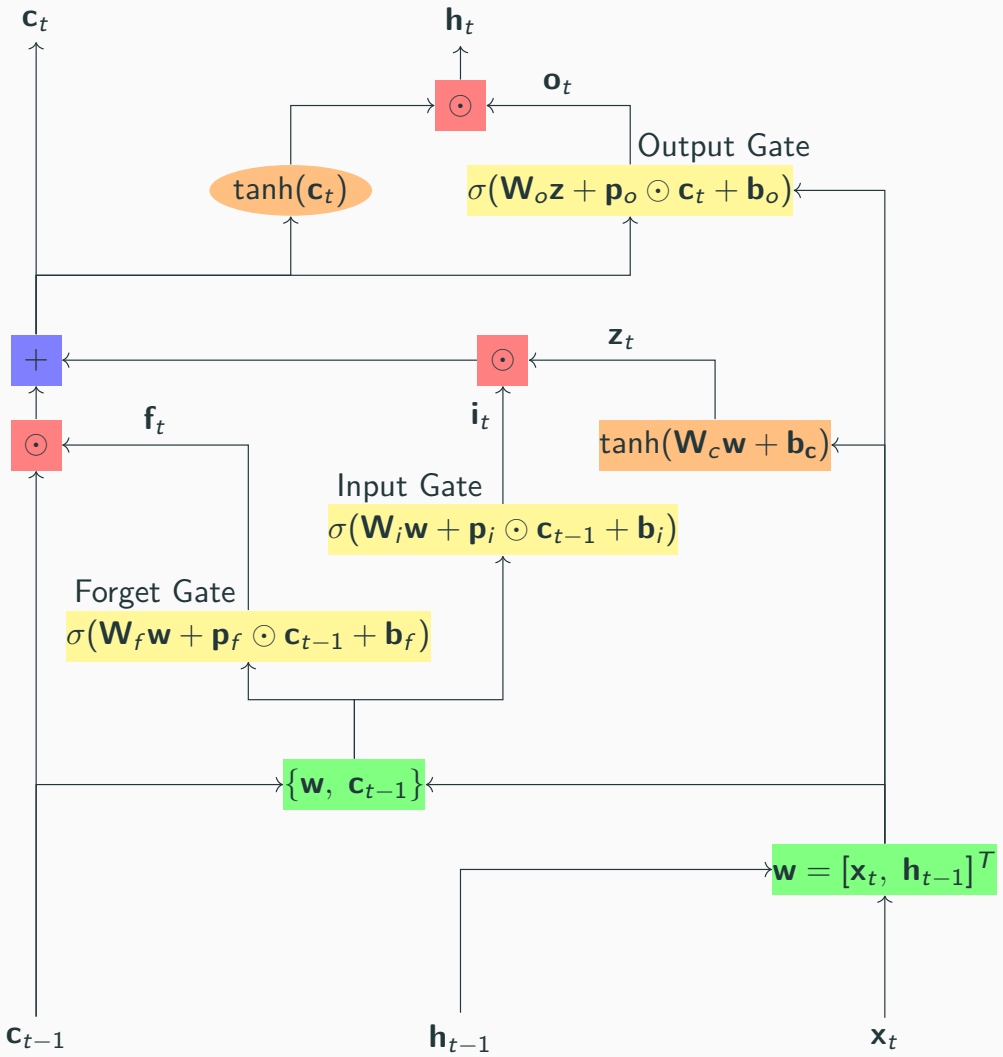
The evolution of the $\mathbf{h}$-sequence is guided by it's largest eigenvalue. If an eigenvalue larger than one exists. The state explodes. If all eigenvalues are smaller than one the state vanishes [GBC16].

Long Short Term Memory (LSTM)



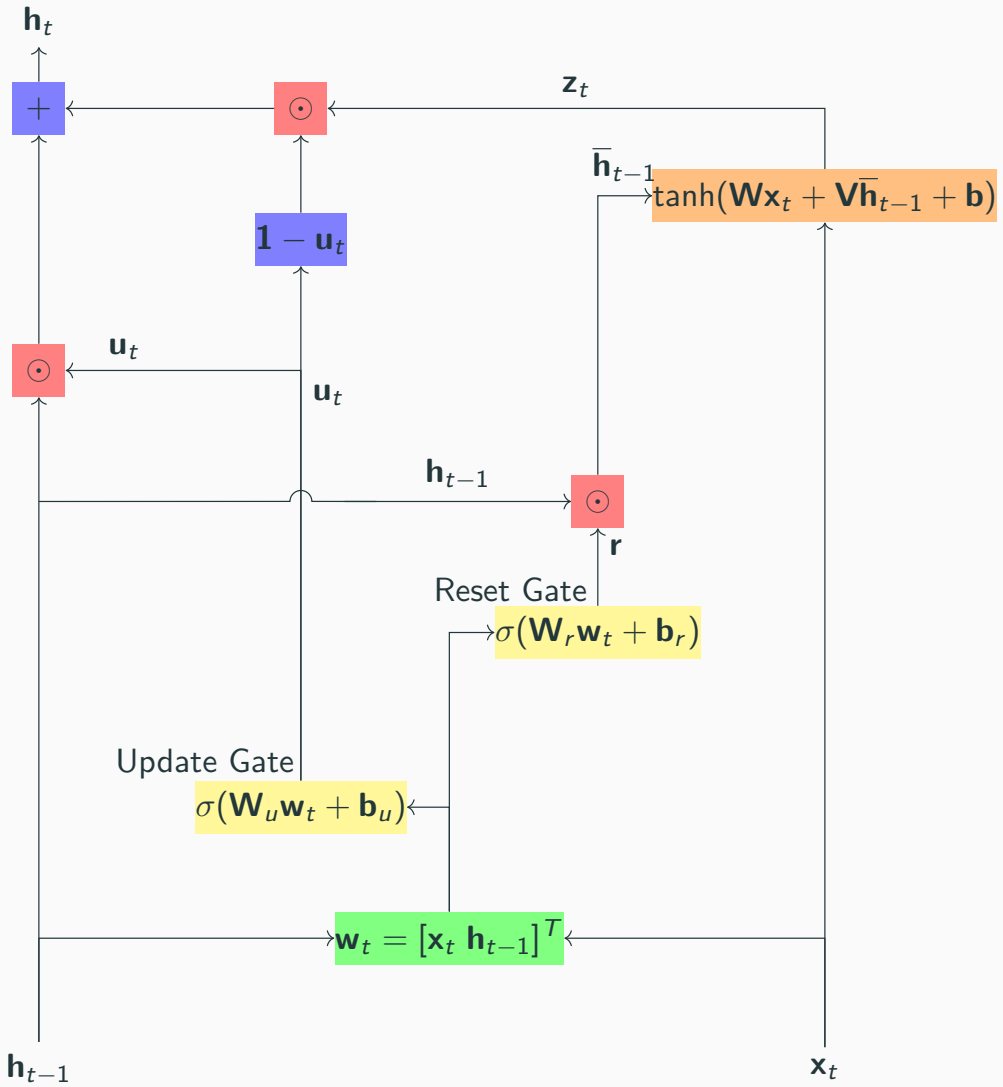
[Gre+16]

Long Short Term Memory (LSTM)



[Gre+16]

Gated recurrent units



Orthogonal networks

TODO

Summary

TODO

Applications

Language Processing

TODO

Speech Processing

TODO

Time-series forecasting

TODO

Conclusion

TODO

References

[Elm90] Jeffrey L Elman. "Finding structure in time." In: *Cognitive science* 14.2 (1990), pp. 179–211.

[GBC16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[Gre+16] Klaus Greff, Rupesh K Srivastava, Jan Koutnik, Bas R Steunebrink, and Jürgen Schmidhuber. "LSTM: A search space odyssey." In: *IEEE transactions on neural networks and learning systems* 28.10 (2016), pp. 2222–2232.