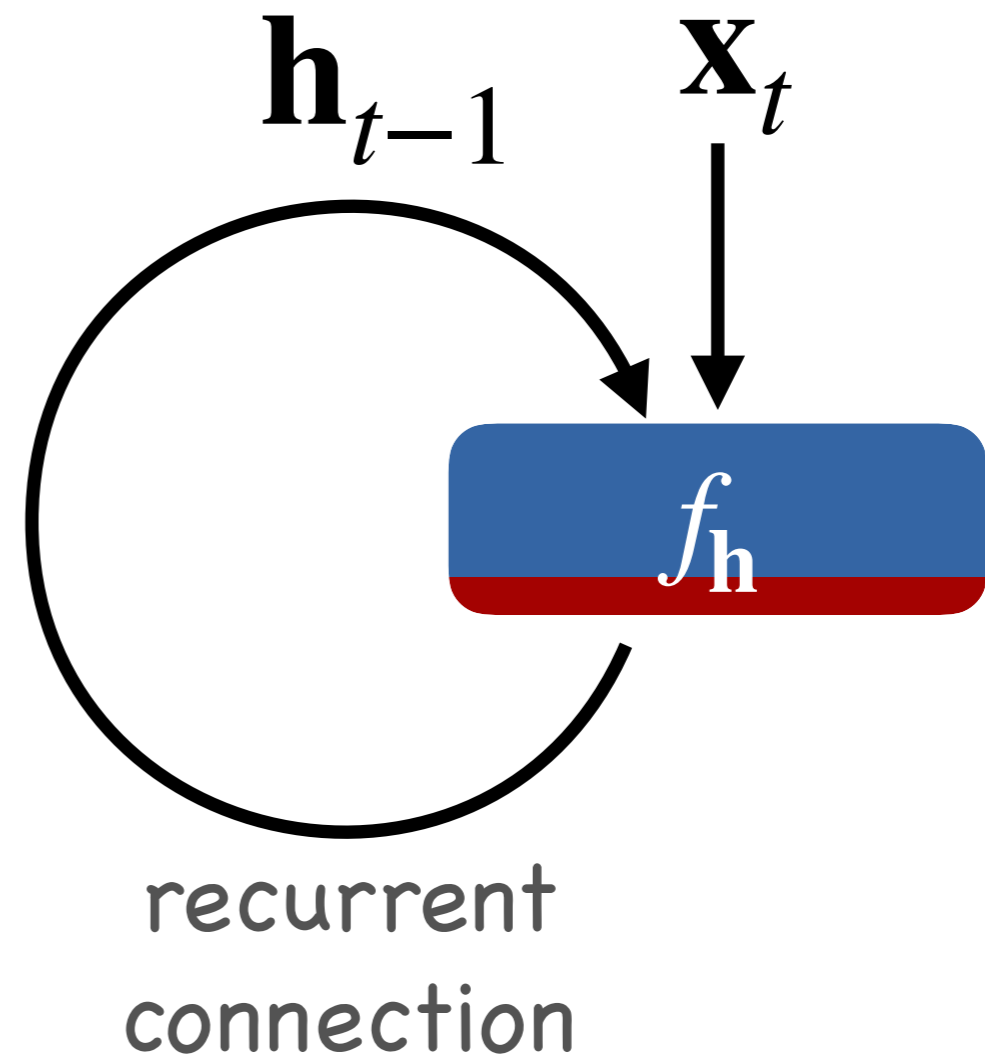


Summary

© 2019 Philipp Krähenbühl and Chao-Yuan Wu

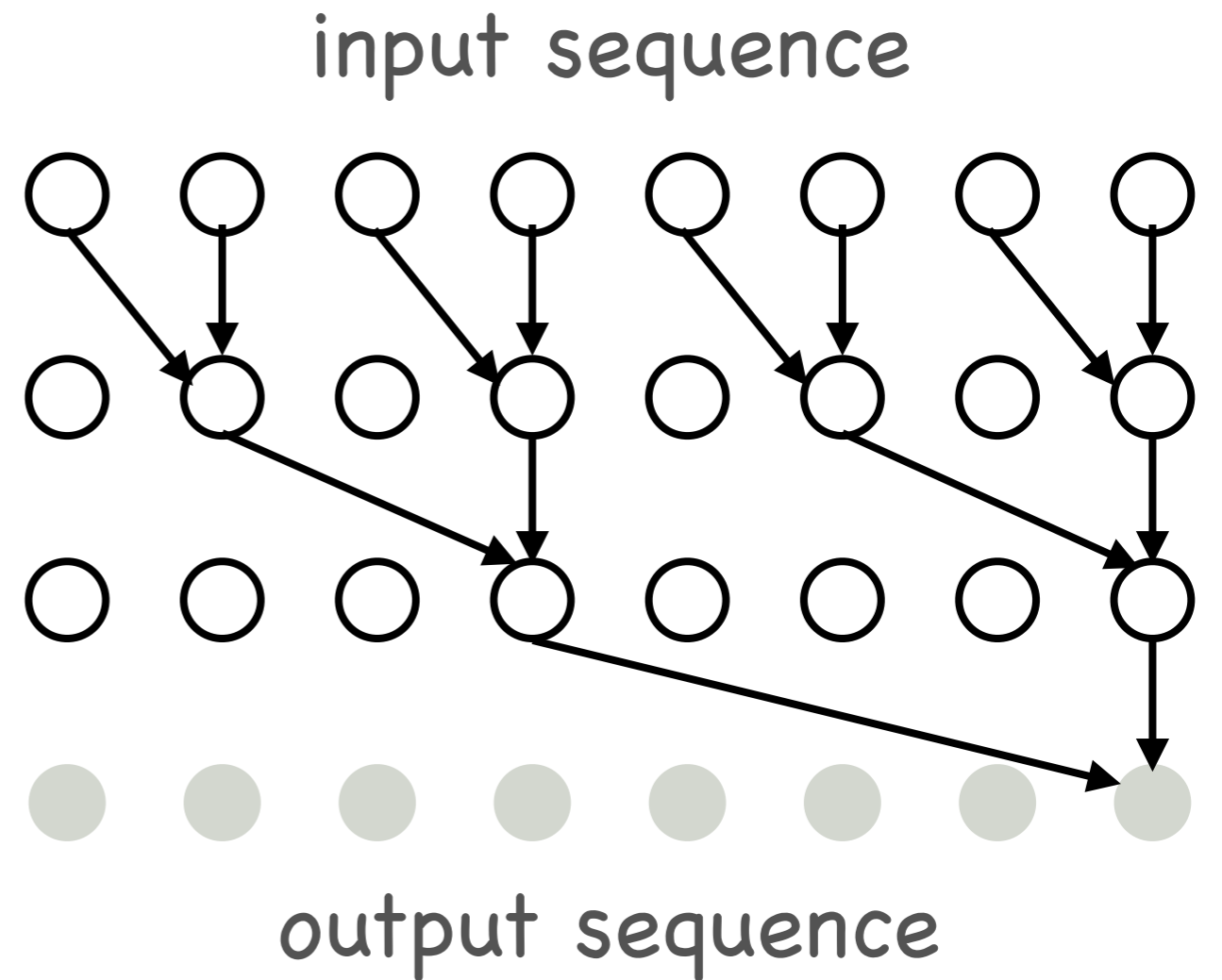
RNNs / LSTMs / GRUs

- Seem like a good idea
- Too hard to train
- No longer widely used



Temporal Convolutional Networks

- Fast and efficient training
- Work well for structured data



Attention / Transformers

- Fast and efficient training
- Better deals with irregular spacing
- State-of-the-art in NLP

