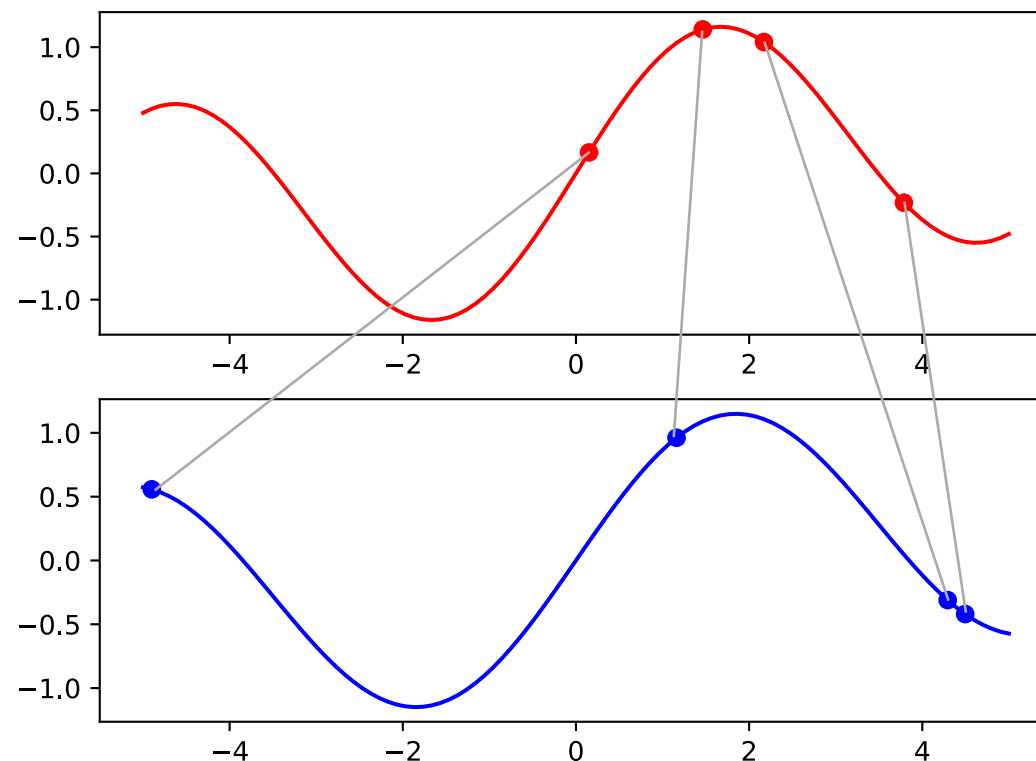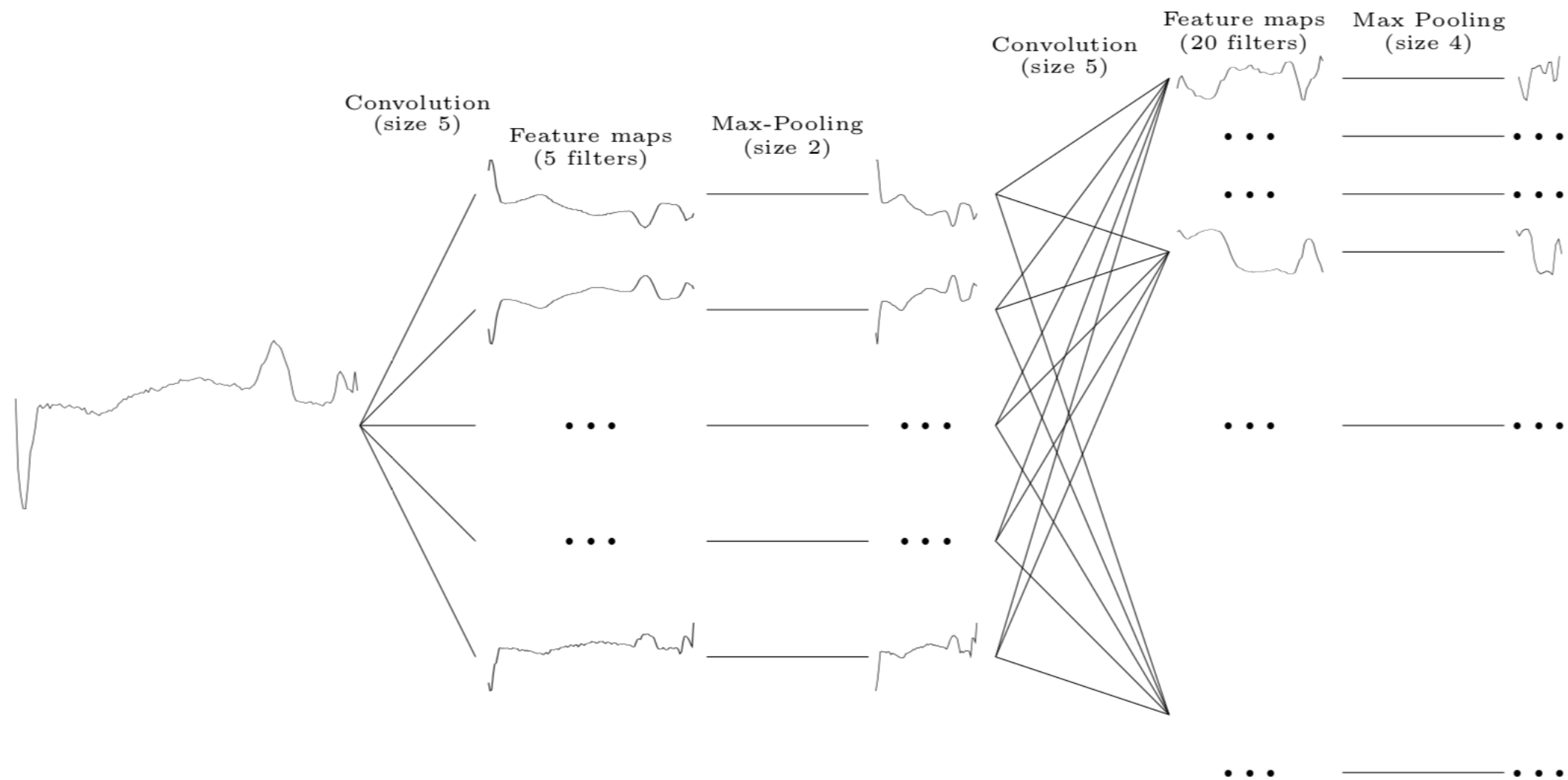# Neural nets for sequences
Text and time series

Romain Tavenard (Université de Rennes 2)
Deep Learning course @EDHEC

Standard issues with sequences

- Variable number of observations per sequence

  - `the cat eats the mouse`

  - `at the moment, the cat is eating the mouse`

- Segmentation (starting/end points)

- Irregular sampling (time series)

# Solution #1: NN with 1d-convolutions

Source: [Le Guennec *et al.*, 2014]

Solution #1: NN with 1d-convolutions

- Variable number of observations per sequence

  - 0 padding

- Segmentation (starting/end points)

  - Data augmentation

  - Pooling

- Irregular sampling (time series)

  - Not robust to that!
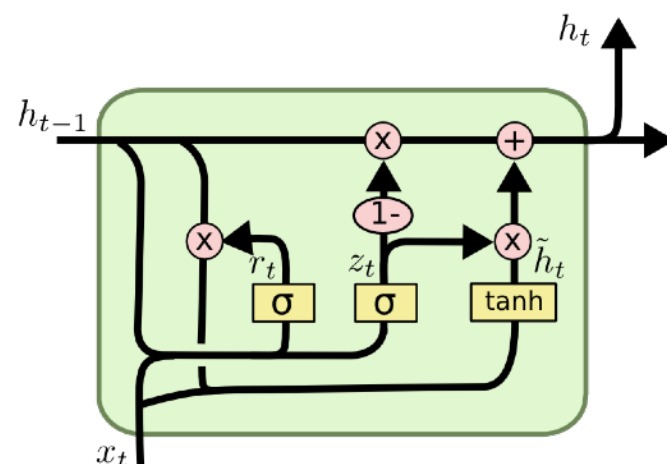
Solution #2: NN with recurrent units

- Very flexible model (any length, let the model learn its memory needs, …)

- Difficult to learn in practice

  - Slow (lack of parallelism)

  - Vanishing gradients (hard to learn long-term dependencies)
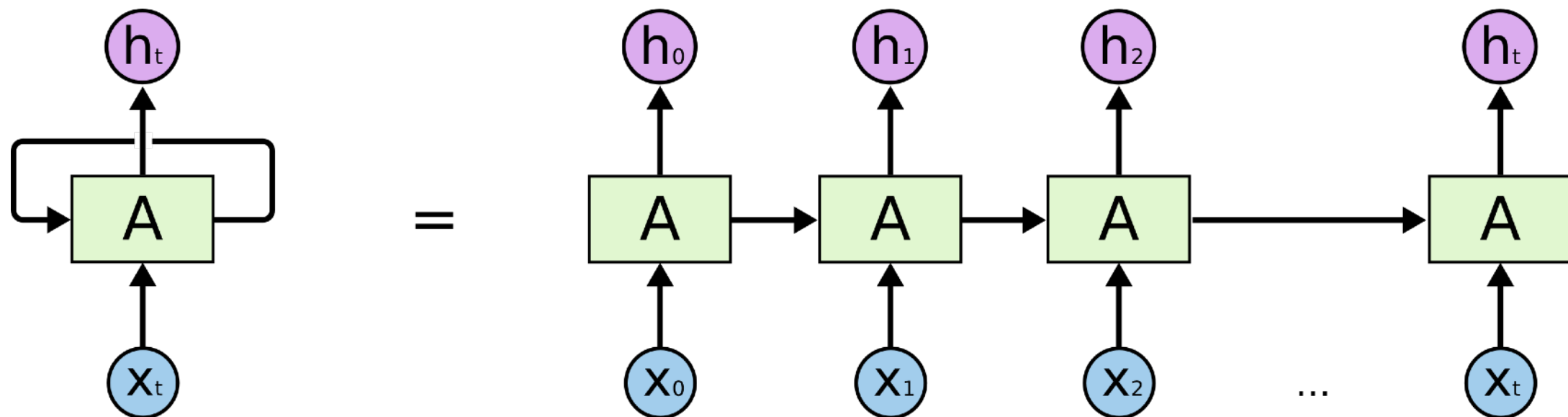
Source: Christopher Olah's blog

Solution #2: NN with recurrent units

- Variants that work better in practice

  - Long Short Term Memory (LSTM)

  - Gated Recurrent Unit (GRU)

- Principle
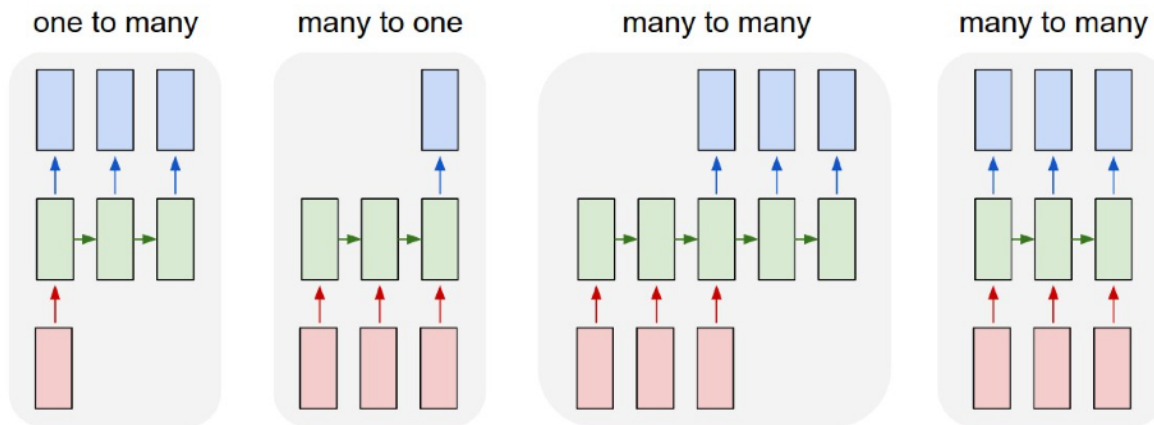
  - At each time step, keep only part of the information

$$z_t = \sigma \left(W_z \cdot [h_{t-1}, x_t] \right)$$

$$r_t = \sigma \left(W_r \cdot [h_{t-1}, x_t] \right)$$

$$\tilde{h}_t = \tanh \left(W \cdot [r_t * h_{t-1}, x_t] \right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Illustration: GRU cell, source: Christopher Olah's blog

Solution #2: NN with recurrent units

one to many　　many to one　　many to many　　many to many

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.
```

Sample text generated by a RNN
trained on Shakespeare words

Sample LaTeX generated by a RNN
trained on a book of algebraic geometry

Source: Andrej Karpathy's blog, http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Solution #2: NN with recurrent units

Memory encoded in variant RNNs,
Source (and interactive demo): https://distill.pub/2019/memorization-in-rnns/

Solution #2: NN with recurrent units

- Variable number of observations per sequence

  - OK

- Segmentation (starting/end points)

  - Not invariant to that!

- Irregular sampling (time series)

  - OK

Summary

- 1d-CNN and RNN can be used

  - depends on the context

  - proved equivalent for stable processes in many-to-one settings [Miller & Hardt, 2019]

  - 1d-CNN are faster to learn