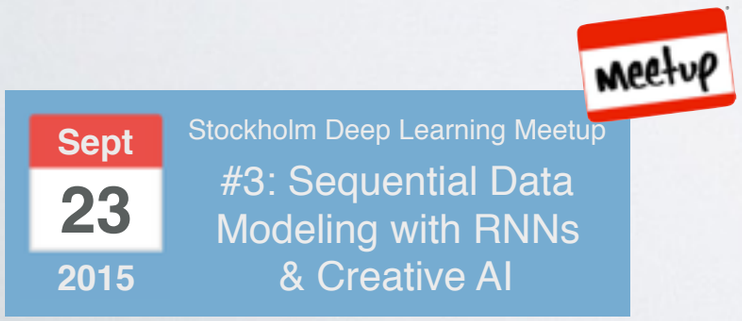
Intro to RNNs in Blocks

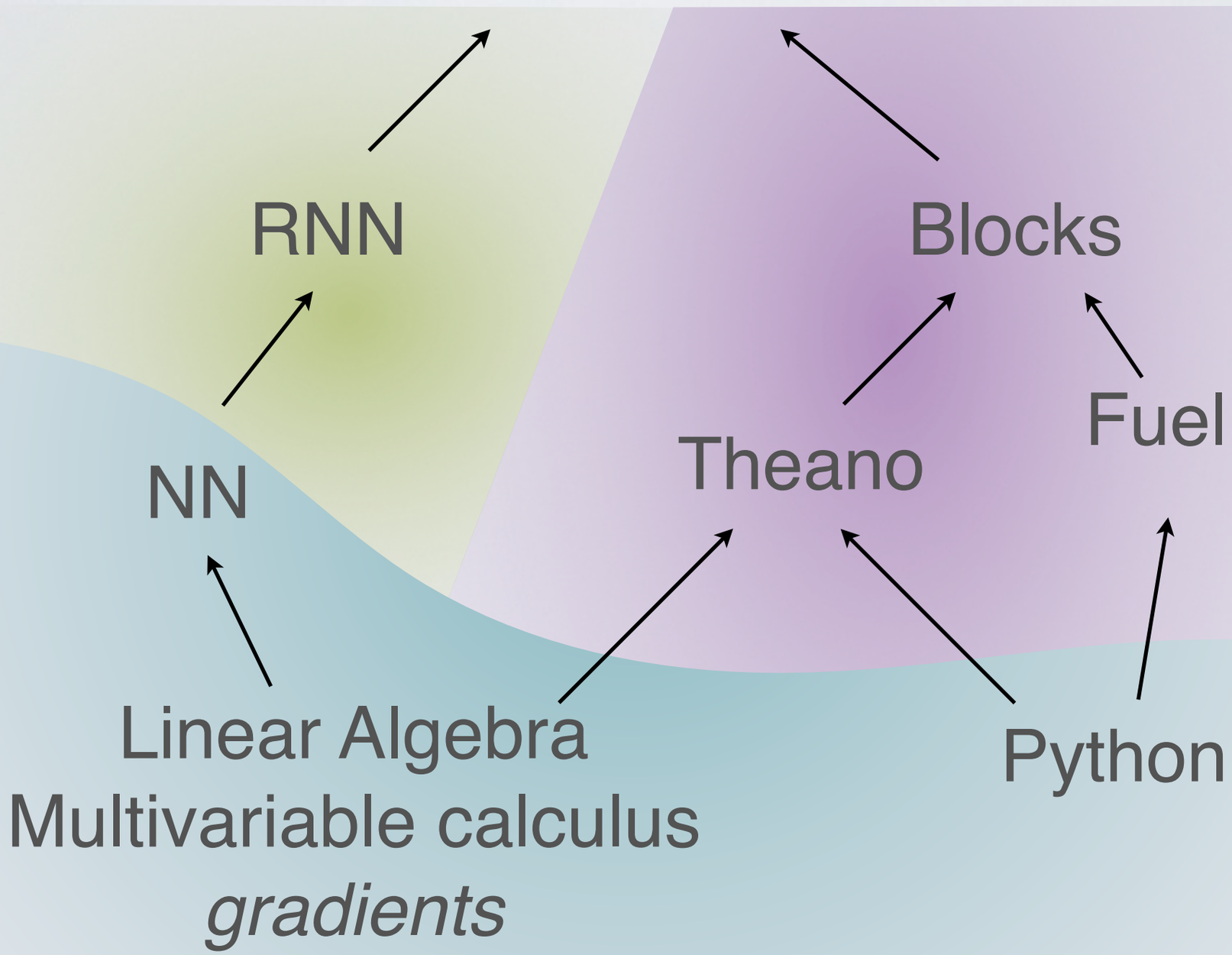
Anders Huss

anders@watty.se



© 2015 Watty AB. All Rights Reserved

Overview


RNNs in Blocks
RNN
Blocks
NN
Theano
Fuel
Linear Algebra
Multivariable calculus
gradients
Python

Overview

- Neural Networks (*the very basics*)
- Theano (*basics*)
- Blocks and Fuel
- Recurrent Neural Networks (RNN)
- RNNs in Blocks
  - ⋆ notebook tutorial
- Discussion/Questions

anders@watty.se

Overview

- **Neural Networks** (*the very basics*)
- Theano (*basics*)
- Blocks and Fuel
- Recurrent Neural Networks (RNN)
- RNNs in Blocks
  - ⋆ notebook tutorial
- Discussion/Questions

anders@watty.se

Neural Networks

…Or any *differentiable parameterised function*?

$$\hat{y} = f(x, \theta)$$

$$D = \{(x, y)\}$$

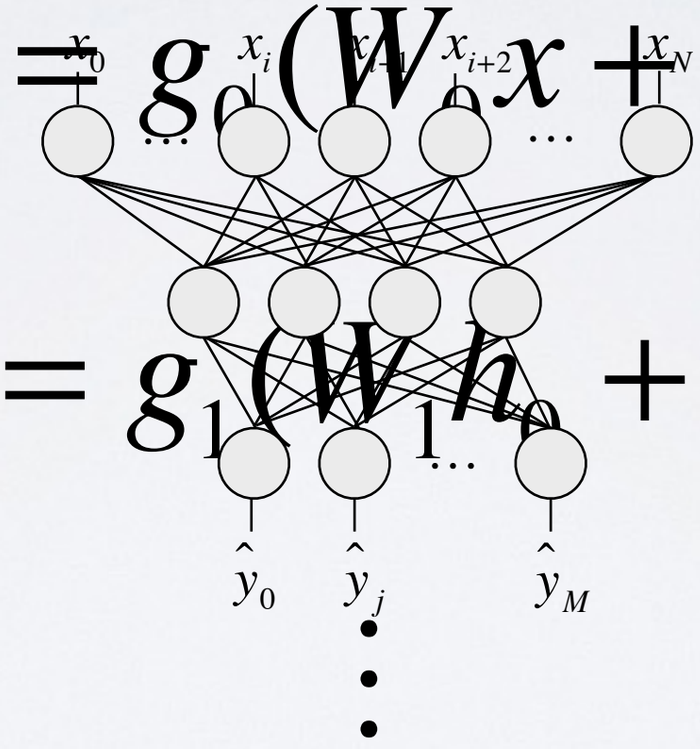
$$c = l(y, \hat{y})$$

anders@watty.se

Neural Networks

…Or any *differentiable parameterised function*?

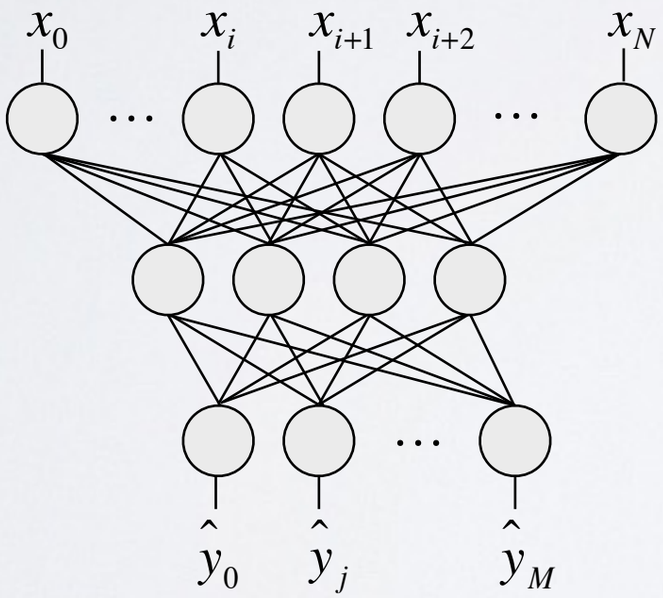
$$h_0 = g_0(W_0 x + b_0)$$



anders@watty.se

Neural Networks

Basically any function
- parameterised
- *differentiable*



$$\hat{y} = f(x, \theta)$$

anders@watty.se

Neural Networks

What's needed to train a NN

- Compute gradients
- Update rule

$$\nabla f = \left(\frac{\partial f(\theta, D)}{\partial \theta_1}, \ldots, \frac{\partial f(\theta, D)}{\partial \theta_N} \right)$$

$$\theta^{i+1} = (\theta^i, \nabla f(\theta^i, D))$$

anders@watty.se

Overview

- ✓ Neural Networks (*the very basics*)
- **Theano** (*basics*)
  - ⋆ notebook tutorial
- Blocks and Fuel
  - ⋆ notebook tutorial
- Recurrent Neural Networks
- RNNs in Blocks
  - ⋆ notebook tutorial
- Discussion/Questions

anders@watty.se

Theano

- Symbolic math
  - Built in standard operators (Ops)
- Symbolic differentiation
- Compiles code for us to efficiently carry out “updates” of parameters of a function (gradient descent)

anders@watty.se

https://github.com/andhus/DLMeetupSept2015/blob/master/notebook_tutorials/theano_basics_tutorial.ipynb

Overview

✓ Neural Networks (*the very basics*)
✓ Theano (*basics*)
- **Blocks and Fuel**
- Recurrent Neural Networks
- RNNs in Blocks
  - ⋆ notebook tutorial
- Discussion/Questions

anders@watty.se

Blocks

- Library and framework for putting together parameterised theano expressions
- Bricks
  - functions (parameterised)
  - Annotated graph
  - Update Algorithms
  - Main Loop
    - Datastreams - Fuel
    - Extensions

anders@watty.se

Theano
Blocks
Fuel

Theano symbolic variables

inputs
targets

MainBrick (NN)
f(inputs) → outputs

CostBrick

cost

ComputationGraph

parameters

Training Alg.

Grad Descent
Step Rules

coupled by name

inputs
targets

DataStream(s)
sources

MainLoop

Extensions
Plotting → **Bokeh Server**
Monitoring
Saving

https://github.com/andhus/DLMeetupSept2015/blob/master/notebook_tutorials/blocks_fuel_basics_tutorial.ipynb

Overview

✓ Neural Networks (*the very basics*)
✓ Theano (*basics*)
✓ Blocks and Fuel

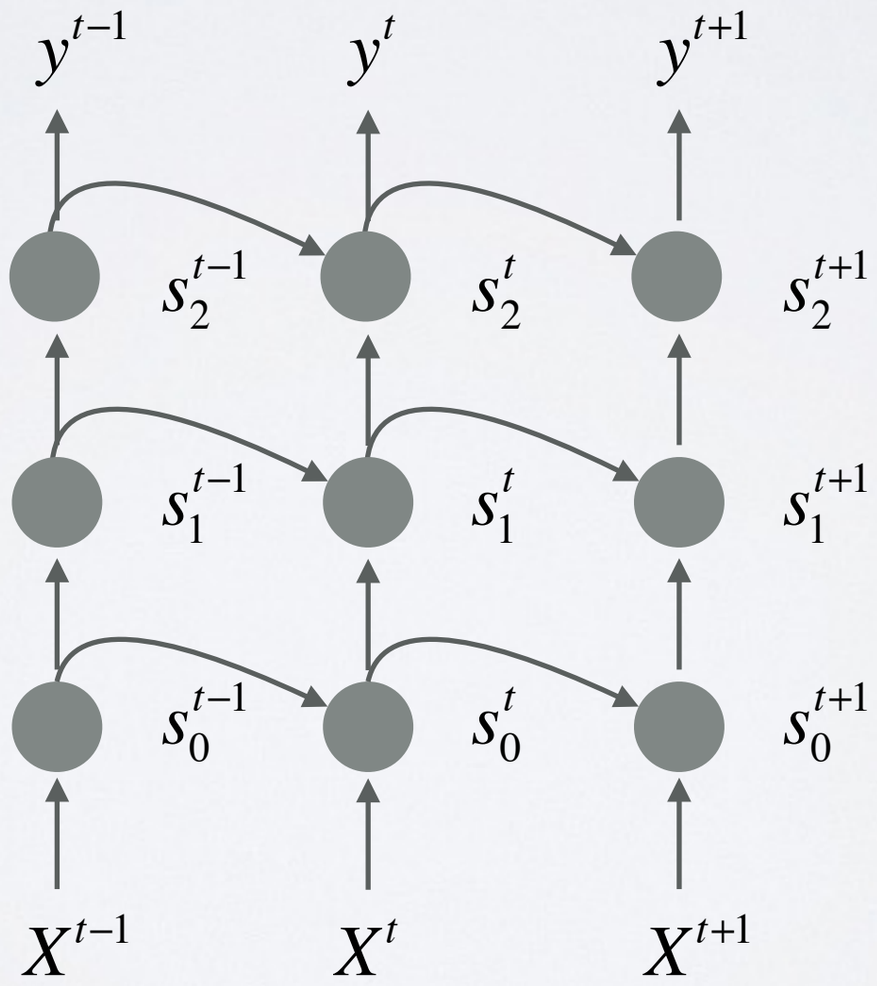
- **Recurrent Neural Networks**
- RNNs in Blocks
  - ⋆ notebook tutorial
- Discussion/Questions

anders@watty.se

Recurrent NN

…Or just a certain type of weight sharing?

A *recurrent* nn is a *DAG*!



anders@watty.se

Overview

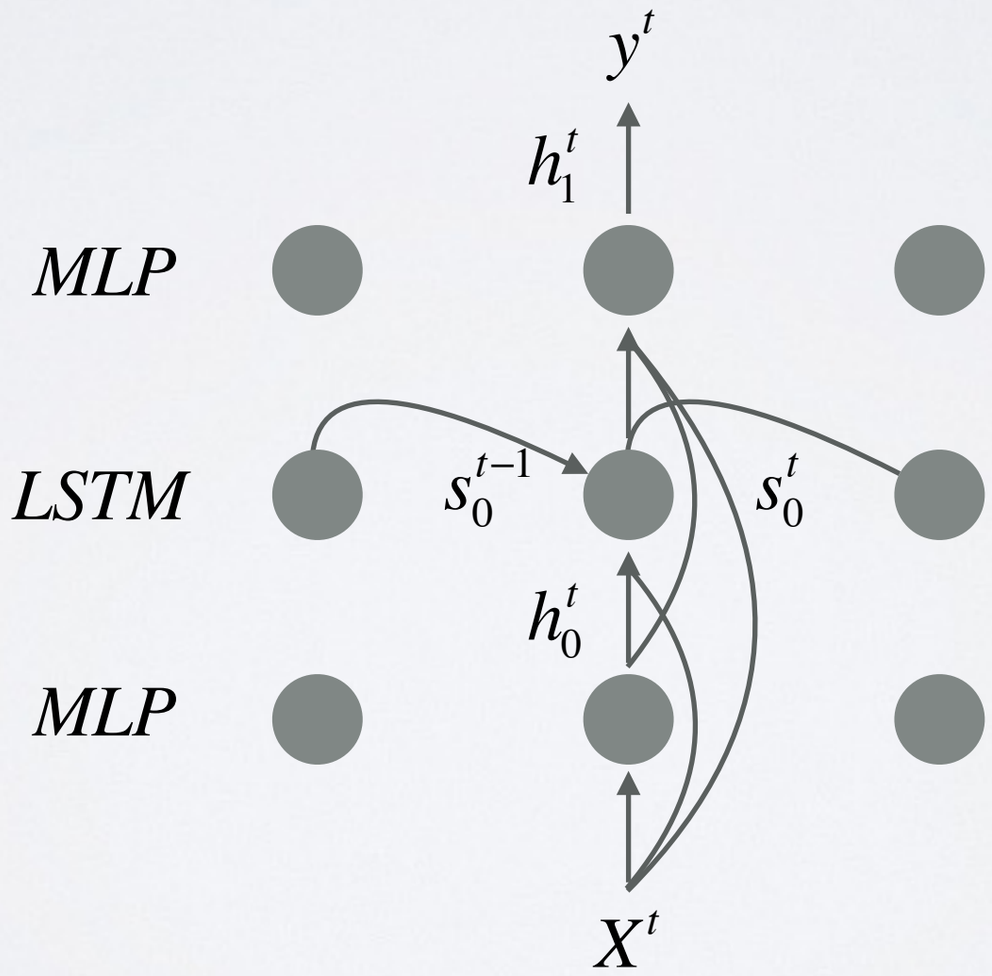
- ✓ Neural Networks (*the very basics*)
- ✓ Theano (*basics*)
- ✓ Blocks and Fuel
- • Recurrent Neural Networks
- • **RNNs in Blocks**
  - ⋆ notebook tutorial
- • Discussion/Questions

anders@watty.se

https://github.com/andhus/DLMeetupSept2015/blob/master/notebook_tutorials/sequence_to_sequence_rnn.ipynb

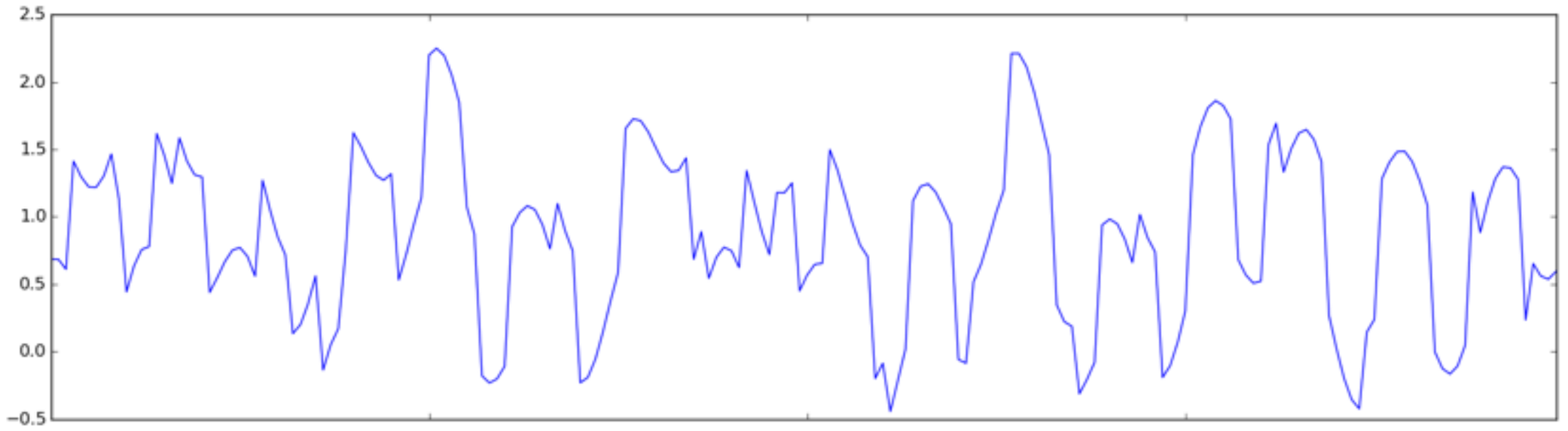
Recurrent NN

Cascading Structure with additional input and output transformations.

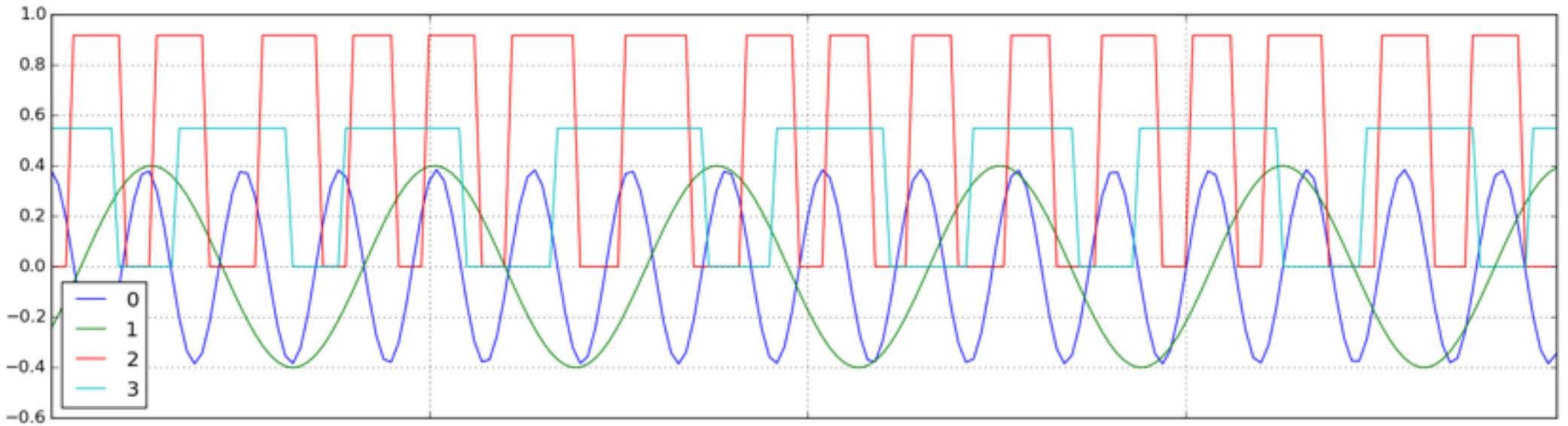


anders@watty.se

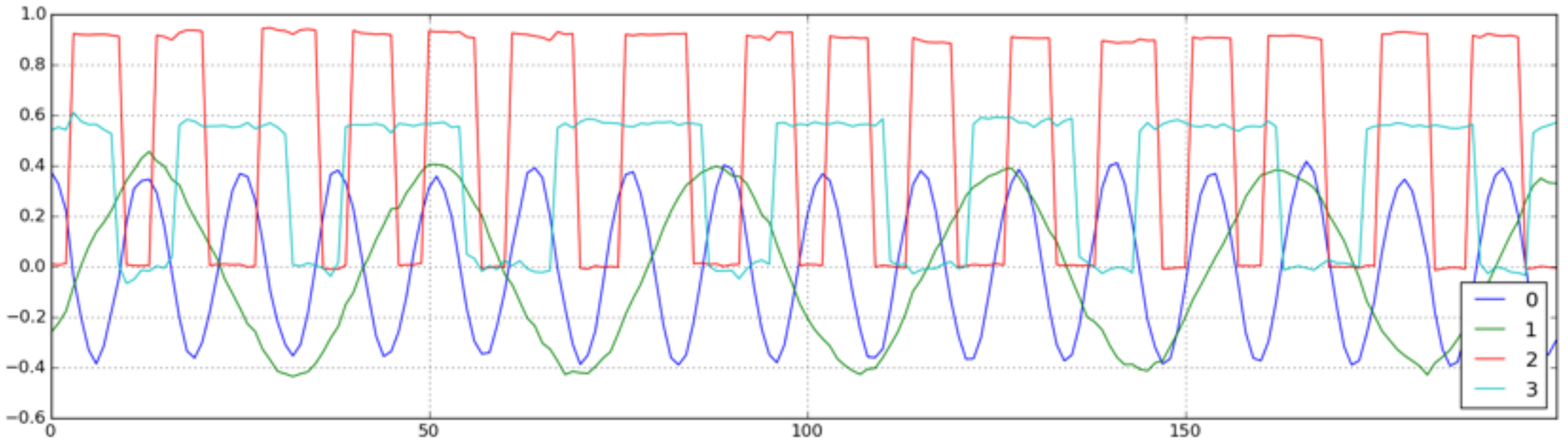
Input



Target



Prediction



Recurrent NN

So what makes them so powerful?

- Sequential Problems
  - handle inputs without fixed length
  - memory - *learning during prediction*

- *Time and causality* is something fundamental
  - Are all problems sequential?
  - Touring Complete
  - Concept of *Attention*

anders@watty.se

Recommended Reading

http://jiwonkim.org/awesome-rnn/

Alex Graves (2008)

- Supervised Sequence Labelling with Recurrent Neural Networks

Q&A

WATTY

anders@watty.se info@watty.se

© 2015 Watty AB. All Rights Reserved