Learning of sequential data

Dynamic Neural Networks (Part 2)

Hugo Leonardo Rufiner
lrufiner@fich.unl.edu.ar

Tópicos Selectos en Aprendizaje Maquinal
Doctorado en Ingeniería, FICH-UNL

15 de octubre de 2010

Outline

Outline

sinc(*i*)

Classification

Dynamic Neural Networks (DNN)

- Time Delay Neural Networks (TDNN)
- Recurrent Neural Networks (RNN)
  - Fully Recurrent Neural Networks
    - Hopfield Networks (associative memories)
    - Boltzman Networks (supervised)
    - Adaptive Resonance Theory (ART)
  - Partially Recurrent Neural Networks (PRNN)
    - Backpropagation Through Time (BPTT)
    - Elman Networks
    - Jordan Networks
- Other (Spiking Networks, Liquid State Machines, etc)

Outline

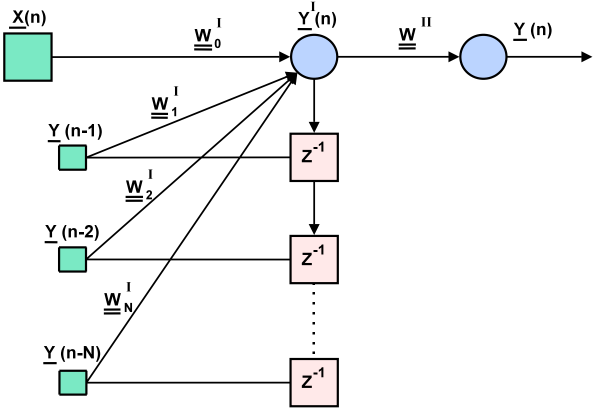
sinc(*i*)

Backpropagation Througth Time (BPTT)

- Total recurrent architecture
- Expansion in a pure feedforward network
- Truncate method (parcial recurrency)

Outline

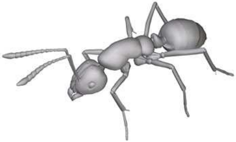
Elman Networks: architecture

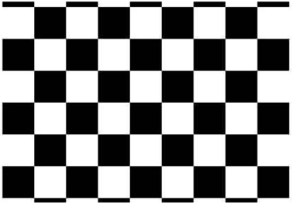


Elman Networks: generalities

- Are natural and simple extension to standard feedforward networks.
- Differ from conventional two-layer networks in that the first layer has a recurrent connection. (Context layer)
- The delay in this connection stores values from the previous time step that are treated as just another set of inputs.
- Standard back-propagation learning techniques can be used.
- Can store information for future reference, so it is able to learn temporal patterns as well as spatial patterns.
- Can be trained to respond to, and to generate, both kinds of patterns.

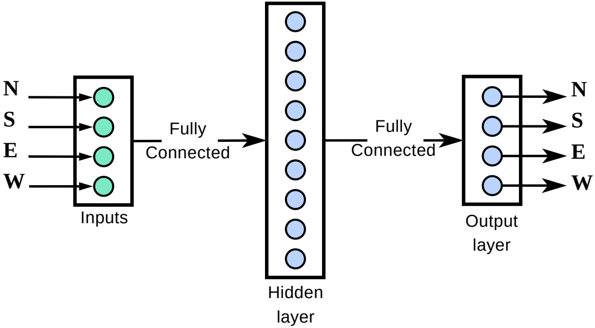
Elman Networks: example





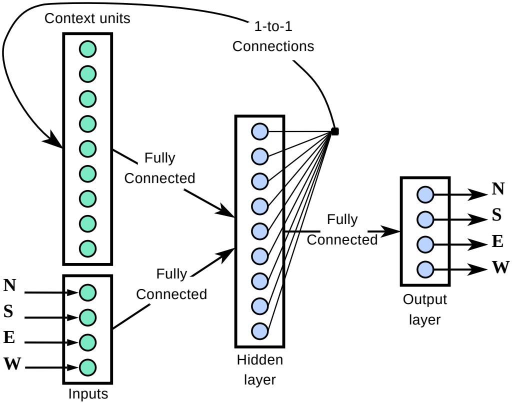
- An intelligent agent that is navigating a grid-world of cells.
- It can sense the cells and move to the north, south, east, and west.
- In order to know what is in the diagonally adjacent cells (i.e., north-east) then the agent will need to remember the values from its last position (which will give some of the missing information, but not all).

Elman Networks: example



- It can't remember but an Elman network can be devised properly.
- An Elman network can remember previous state by adding a new set of inputs which are fully connected by recurrent links to the hidden layer outputs (but delayed by one unit of time).

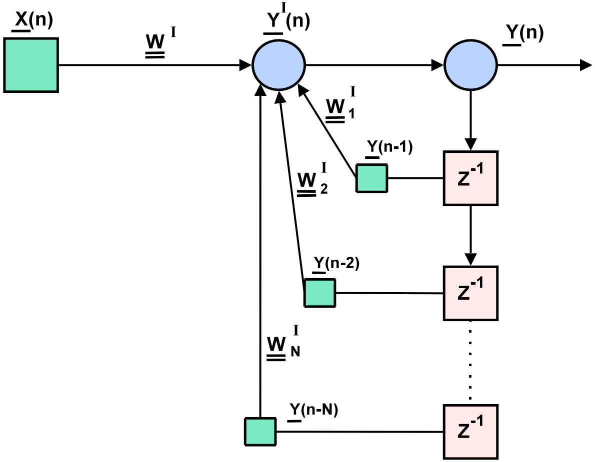
Elman Networks: example



It should be able to learn a suitable function from inputs and stored state to action.

Outline

Jordan Networks: architecture



sinc(i)

Jordan Networks: generalities

- Use the output historial
- The network state is defined by:

$$\mathbf{s}(t) = \tfrac{1}{2}\mathbf{s}(t-1) + \mathbf{y}(t-1)$$

- An output infinity historial is stored, then the network state is:

$$f_{\mathbf{s}}(\mathbf{s}(t-1), \mathbf{x}(t), \mathit{\Gamma}) \equiv \sum_{i=1}^{t} (\tfrac{1}{2})^{i-1} \mathbf{y}(t-i)$$

- Modifications have been made to facilitate its applications in dynamic systems identification.

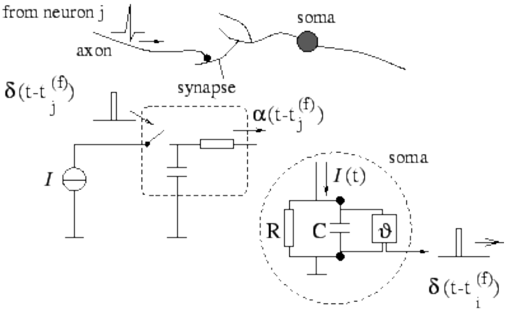
sinc(i)

Outline

sinc(*i*)

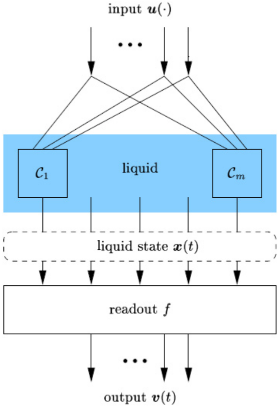
Other networks that model temporal dynamics

- Spiking Neural Networks (also "Pulsed", Maass and Bishop, 2001) Architecture, simulation and training



Other networks that model temporal dynamics

- Liquid State Machines (Mass and Natschläger, 2001),



- More ...

sinc(i)

Outline

Simon Haykin.
Neural networks, A Comprehensive Foundation.
Macmillan College Publishing Company, New York, 1994. Caps. 11-15.

Raúl Rojas.
Neural networks: A Systematic Introduction,
Springer-Verlag New York, Inc., New York, NY, 1996.
http://page.mi.fu-berlin.de/rojas/neural/

Jeffrey L. Elman.
Finding structure in time.
Cognitive Science, 14(2):179–211, 1990.

sinc(*i*)

Michael I. Jordan.
Attractor dynamics and parallelism in a connectionist sequential machine.
Proceedings of the Eighth Annual Meeting of the Cognitive Science Society, 531–546. Erlbaum, Hillsdale, NJ, 1986.

Ernest Haselsteiner.
What Elman networks cannot do
Neural Networks Transactions, 6(4):1000–1004, 1995.

J. Ludik; W. Prins; K. Meert; T. Catfolis.
A comparative study of fully and partially recurrent networks
International Conference on Neural Networks,1997., 1:292–297, 1997.

D.T. Pham; D. Karaboga.
Training Elman and Jordan networks for system identification using genetic algorithms
Artificial Intelligence in Engineering, 13(11):107-117, 1999.

Wolfgang Maass; Christopher M. Bishop.
Pulsed Neural Networks (Bradford Books)
The MIT Press, 2001.

W. Maass, T. Natschlaeger, and H. Markram.
Real-time computing without stable states: A new framework for neural computation based on perturbations
Neural Computation, 14(11):2531-2560, 2002.

sinc(*i*)