

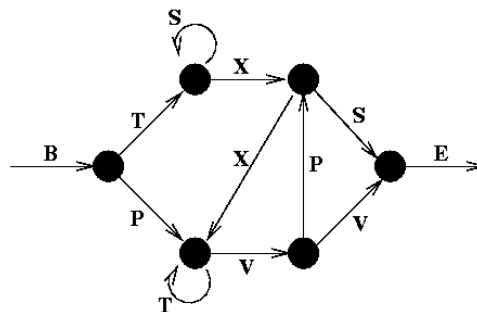
(for groups who focus on the MLP)

Exercises to the module: Algorithmisches Lernen
SS 2012 Sheet 8 (for groups who focus on the MLP)
 Due: 06.06.2012

Task 8.1 Elman Network

In this exercise you will gather some experiences with an Elman Network. To this end you find in the MinCommSy the program `time_reber.py` that generates data according to a Reber grammar; furthermore, two data sets taken from <http://sites.stat.psu.edu/~rho/forecast/>.

- (a) What are differences between the implementations of an MLP and an Elman network? Which parameters does the recurrent network have that the MLP does not have?
- (b) Using the Python script `time_reber.py` you produce a random sequence according to the following Reber grammar.



Train an Elman network on the generated sequence, i.e. train it to predict the next symbol! For every symbol in the Reber grammar, the network has one output neuron, which is active for its corresponding symbol, while the other neurons are inactive. Hence, the `softmax` function is a suitable activation function for the output layer.

Repeat the experiment, each time with a newly generated Reber sequence. When testing, only the first symbol is given to the input; the network output is then fed back to the input for the next time step. Plot the generated network outputs, compared with the target outputs. Can the network learn every sequence? Vary the network parameters, such as the number of hidden neurons. Characterize those sequences that are hard for the network to learn.

- (c) Train the network on one of the time series `time_erie.dat` or `time_river1.dat`. Again, during training only the value of the next time step will be learnt, while for testing, the network output is fed back to the input. Which characteristics of the data does the network capture, which does it fail to capture? What can be improved on the learning rule?

Note: Now a linear transfer function is appropriate for the output unit. Nevertheless, it is advisable to scale the data; for example to subtract the mean.

- (d) Find another time series data set for prediction. Considering the characteristics of the data, how do you expect the Elman network to cope with these data? Train the Elman network on the data and report your observations.



Note: the link to the UCI Machine Learning Repository is also on the web site of the Marsland book: <http://archive.ics.uci.edu/ml/datasets.html>. However, you may also search elsewhere for suitable time series data.

Note: there is an example Python implementation of an Elman network in the MinCommSy: `mlprec.py` (class), and `mlprec_eg.py` (usage of the class). In the methods `mlppplot` and `mlppplotreber` the network feeds back its outputs to its inputs (from time step `t_given` onwards). In these methods you also find suitable `gnuplot` commands, which you may use for visualisation like in `Aufgabenblatt_7_SOM`.