Dieterrich, Advances in neural information processing systems 14 vol 2

# Vol 2

Reinforcement learning with Long Short Term Memory, Bram Bakker, Leiden University

LSTM originally designed for supervised timeseries learning. Based on the analysis of the problems that conventional recurrent neural network learning algorithms e.g. BPTT or RTRL have when learning timeseries with long term dependencies – errors propagated back in time tend to vanish or blow up.

LSTMs solution is to enforce constant error flow in a number of specialized units called Constant Error Carrousels (CECs). CECs have linear activation functions that do not decay over time. To prevent CECs from filling up with useless info access is regulated with input gates. Input gates receive input from time series and other units in the network and they learn to open and close access to CECs at appropriate moments. Access from the activations of the CECs to the output units is regulated with output gates which also learn when to send an output. Also have forget gates which reset the activation when the info in CEC is no longer useful. Combination of CEC and gates is a memory cell.