

# Readings in Neuroinformatics

Ephraim Seidenberg

Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. 1999.

## Abstract

Learning requires the prioritization of some information over other, which in turn needs training. The powerful learning model of Recurrent Neural Networks (RNNs) has recently been refined by the “Long Short-Term Memory” (LSTM) model. It allows to bridge time gaps between input events and target signals by the use of “constant error carousels” (CECs). It’s training however still relies on a clearly segmented input and is prone to failure when presented with a continual input stream. Here we present a solution to this by introducing “forget gates”. When activated, they reset memory which has become useless and in this way improve the information prioritization. To examine our model we used a more difficult, continual variant of the “Embedded Reber Grammar” problem (CERG) for training and testing, as well as another nonlinear long time task with its continual variant. We found that our LSTM model extended by forget gates was able to solve the continual tasks when the standard LSTM wasn’t. For the non-continual task, although taking a bit longer than the standard model, it was able to solve it as well. Our model renders RNNs more biologically plausible since real-world input does not usually present itself with inherent decomposition into subsequences or subtasks, but rather as a continuous stream. Thus, our model is most relevant to tasks as e.g. speech recognition where sequential processing is required.

226 words.