A recurrent neural network acquires working memory properties by reward-dependent STDP

Cristina Savin and Jochen Triesch
Frankfurt Institute for Advanced Studies
Frankfurt am Main, Germany

Working memory (WM) is critical for many tasks in everyday life. However, the ability to maintain information in memory for brief periods of time is not a given. During development, WM improves significantly, in both capacity and maintenance time [1]. This improvement has been attributed to a variety of factors, from the maturation of cortical tissue, to the development of better strategies to encode and store information, e.g. chunking. Here, we ask if WM can develop by learning to perform tasks that require some temporary storage of information. We address this question in the context of general-purpose recurrent neural networks. Specifically, we ask if an initially unstructured neural network can acquire WM properties while learning to perform a delayed response task. Unlike classical WM models [2, 3, 4], we make no *a priori* assumptions on stimulus encoding, leaving the network to discover an appropriate respresentation by reward-dependent learning.

The network consists of linear threshold neurons with sparse recurrent connectivity [5, 6]. Stimulus specific inputs are delivered to small non-overlapping subpopulations within the network, with an additional nonspecific background input. The output layer reads the activity of excitatory units and, through a winner-take-all mechanism, selects the action to be taken, which yields a corresponding reward. The connectivity is modified through reward-dependent STDP [7], with an additional weight normalization. Network activity is stabilized by an intrinsic plasticity rule regulating the neuron's spike threshold to maintain a certain mean average firing rate. During the delayed response task, a stimulus, selected at random from a set of K, is presented to the network, leading to a one time step activation of the corresponding subpopulation within the recurrent network. After a delay —either fixed for all trials, or selected independently from a uniform distribution between 1 and a maximum $D_{\max}$—, a cue is presented, indicating that at the next time step the action will be selected and the corresponding reward (± 1 for correct and incorrect, respectively) will be delivered.

Our network is able to learn to correctly perform this task. Performance is influenced by the task difficulty, being better for fixed delay and decreasing with delay size. More interestingly, neurons within the recurrent network acquire stimulus specificity, as reported in various WM experiments. Additionally, neurons respond to inputs in a time-dependent way, suggesting that stimulus identity is encoded in a neuronal trajectory [8]. Such time-dependent responses have been recently discovered in various neurophysiological experiments (see [9] for list), but cannot be accounted for by traditional WM models, based on recurrent networks with attractor dynamics. Furthermore, the type of trajectory used varies for the fixed and variable delay task, indicating that the encoding of the same stimulus is task specific.

Our results suggest that reward-dependent learning can shape cortical connectivity to build a working memory. Additionally, as task demands influence the way information is represented within the system, our findings raise interesting questions related to neural coding. It is a challenge for future work to further investigate the context in which different types of encoding may arise within a general purpose recurrent network. Lastly, our network's performance is significantly higher than that of a similarly constructed liquid state machine [10]. From a more general computational perspective, this indicates that reward-dependent learning may be used for enhancing the computational abilities of recurrent neural networks.

Acknowledgments

The authors are supported by EC MEXT project PLICON grant number MEXT-CT-2006-042484 and the German Federal Ministry of Education and Research (BMBF) within the "Bernstein Focus: Neurotechnology" through research grant 01GQ0840.

References

[1] M. Luciana and C. Nelson. The functional emergence of prefrontally-guided working memory systems in four-to eight-year-old children. *Neuropsychologia*, 36(3), 1998.

[2] D. Amit and N. Brunel. Learning internal representations in an attractor neuron network with analogue neurons. *Network*, 6, 1995.

[3] N. Brunel and X.J. Wang. Effects of neuromodulation in a cortical model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, 11:63–85, 2001.

[4] G. Mongillo, O. Barak, and M. Tsodyks. Synaptic theory of working memory. *Science*, 319(5869):1543–1546, 2008.

[5] C. van Vreeswijk and H. Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724 – 1726, 1996.

[6] A. Lazar and J. Triesch. Predictive coding in cortical microcircuits. *Proc. ICANN*, 2008.

[7] E. Izhikevich. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17:2443–2452, 2007.

[8] W. Maass and D.V. Buonomano. State-dependent computations: spatiotemporal processing in cortical networks. *Nature Rev. Neurosc.*, 2009.

[9] Goldman. M. Memory without feedback in a neural network. *Neuron*, 61(4):621–634, 2009.

[10] W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14, 2002.