620.6

UNSUPERVISED AND REINFORCEMENT LEARNING BY PYRAMIDAL NEURONS IN A LAYERED MODEL OF NEOCORTEX

M. Eisele* and T. Sejnowski, Computational Neurobiology Laboratory, Salk Institute, San Diego CA 92186-5800

We previously demonstrated that computational models of pyramidal neurons are capable of a simplified version of indirect, or model-based, reinforcement learning. They can detect temporal correlations in their excitatory input, use them to predict future inputs, infer the impact of their present firing on future inputs, detect which excitatory inputs are followed by a reward signal like dopamine, and change their firing patterns so that they have a positive impact on the occurence of rewards. A major limitation of this learning algorithm is that neurons can not easily detect delayed impacts, like those produced by the motor output of the neocortex on its sensory input. Here we demonstrate computationally how this problem can be solved by a layered network of similar, but more specialized neurons, whose connections resemble those of layers 2&3, 5, and 6 of the neocortex. Layers 2&3 in our model detect temporal invariances on which predictions of the sensory input can be based. Neurons of layer 6 learn to predict the sensory input to the thalamus. This works best if the network goes through a sleep-wake cylcle, with layer 6 driving the thalamus during sleep. Neurons of layer 5 learn to detect how their motor output affects the predictions made by layer 6 neurons. The learning rule in basal dendrites is similar to temporal difference learning, but the apical learning rule is quite different, as it is specialized in detecting the feedback from other, nearby neurons. Together, these unsupervised learning processes produce a network that can perform efficient, model-based reinforcement learning in layer 5. A few reward signals are enough to change the motor output in the right way, even if its impact on sensory input is quite delayed.

Supported by the german DFG grant EI 373/1-2