BMC Neuroscience

Poster presentation

Operant behavior controlled by position of a moving object – a reinforcement learning model

Cyril Brom*[1], Daniel Klement[2] and Michal Preuss[1]

Address: [1]Dept. of Software and Computer Science Education, Faculty of Mathematics and Physics, Charles University in Prague, 118 00, Czech Republic and [2]Dept. of Neurophysiology of Memory and Computational Neuroscience, Institute of Physiology, Academy of Sciences of the Czech Republic, Prague, 142 00, Czech Republic

Email: Cyril Brom* - brom@ksvi.mff.cuni.cz

* Corresponding author

Description

It has been demonstrated that operant behavior can be controlled by spatial stimuli. In one of our experiment, rats were conditioned to press a lever for reward when a moving object was passing through a particular region of the experimental room (unpublished data). Although the stimulus was changing smoothly, the transitions between rewarded and non-rewarded condition were sudden. Consequently the animals anticipated the arrival to the rewarded zone by responding in its vicinity.

We developed a reinforcement learning model to simulate this anticipatory behavior and to study its spatial and temporal components. An output neuron integrated inputs from four classes of sensory neurons: (1) neurons detecting the position of the object, (2) neurons indicating the time elapsed since the last reward and (3) since the last operant response, and (4) a neuron signaling the presence/absence of the reward. While the output neuron was a leaky-integrator with a binary activation function, a manner for sending a motor signal to press the lever, the sensory neurons were simple nodes lacking the time dynamic component that signaled the presence of a stimulus in their receptive field in a rate-coded manner. The synapses between the sensory neurons and the output neuron were modified according to a rule based on the Rescorla-Wagner rule [1]. The overall model resembles the spectral-timing model of Grossberg and Schmajuk [2] extended to the spatial domain.

Depending on the set up of learning parameters related to the different classes of sensory neurons, the network can learn the spatial and/or temporal features of the task resulting in spatial and/or temporal anticipation of the reward. The network well approximates data observed in real animals.

Acknowledgements

References

1. Rescorla RA, Wagner AR: **A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement.** In *Classical Conditioning II: Current Research and Theory New-York: Appleton-Century-Crofts*; 1972:64-69.
2. Grosberg S, Schmajuk NA: **Neural dynamics of adaptive timing and temporal discrimination during associative learning.** *Neural Networks* 1989, **2**:79-102.