A neural network model of operant behavior controlled by a moving object

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

[1]Faculty of Mathematics and Physics, Charles University in Prague, CZ 118 00
[2]Institute of Physiology, Academy of Sciences of the Czech Republic, Prague, CZ 142 00
brom@ksvi.mff.cuni.cz

Abstract

We have simulated behaviour of a rodent solving a spatial task using a simple artificial neural network. The model behaves similarly to the rodent; in particular it anticipates the reward accordingly.

Description

Recently non-spatial operant behavior controlled by spatial stimuli has been employed in the research of spatial cognition in rodents (Klement and Bures, 2000; Pastalkova et al., 2003; Nekovarova and Klement, 2006). In one of our experiments, rats were conditioned to press a lever for food reward when a light rectangle was passing through a particular region on a computer screen (Klement et al. 2008). The rewarded region could be determined only by its spatial relation with respect to the edges of the monitor. Thus the rats were able to estimate the transitions between the non-rewarded and rewarded periods only approximately. The effort to receive maximum amount of reward resulted in anticipatory operant responses emitted before the rectangle entered the rewarded zone.

We developed a reinforcement learning model to simulate this anticipatory behavior and to study its spatial and temporal components. The model comprises an output neuron, which is a leaky integrator with a binary activation function, and several classes of sensory neurons, rate-coded detectors with simple receptive fields lacking the temporal component. The output neuron signals whether to press the lever or not. The sensory neurons are of for classes: (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. The synapses between the sensory neurons and the output neuron are modified according to a modified Rescorla-Wagner rule (Rescorla and Wagner, 1982). The overall model resembles the spectral-timing model of Grossberg and Schmajuk (1989) extended to the spatial domain and for an operant behavior task.

The network can learn the spatial and/or temporal features of the task, thereby anticipating the reward based on spatial/temporal information. The network well approximates data observed in real animals (Fig. 1). We have also scaled the original spectral-timing model for this task and compared the data of the scaled model with data of our model. Both models produce similar results; however, our model is conceptually simpler than the original spectral-timing.
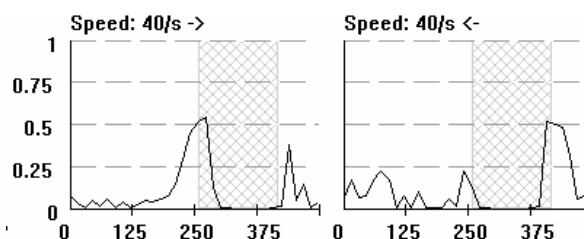

Fig. 1. Frequency of pressing the lever after the model has been learned (over one modeled trial). Y-axis: the frequency (normalised). X-axis: the position at the screen (in hypothetical units: 500 units = the horizontal width of the screen). Left: the rectangle is moving from right to left at a constant speed. Right: the rectangle is moving from left to right at a constant speed. The rewarded zone is shaded and asymmetrically positioned. Notice the anticipation of the reward when the rectangle is approaching the rewarding zone. At the left part, there is also the "phantom pressing" after the rewarded zone visible, an attempt to receive the second reward.

References

Grosberg, S., Schmajuk, N. A. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. *Neural Networks*, 2:79-102.

Klement, D., Bures, J. (2000). Place recognition monitored by location-driven operant responding during passive transport of the rat over a circular trajectory. *Proc. Natl. Acad. Sci. USA*, 97(6):2946-51.

Klement, D., Duskova, L., Levcik, D., Nekovarova, T. (2008). Is the ability to recognize actual position of a moving object hippocampal dependent? [In Czech]. 50th Czech-Slovak Psychofarmacological Coference, Jesenik, Czech Republic.

Nekovarova, T., Klement, D. (2006) Operant behavior of the rat can be controlled by the configuration of objects in an animated scene displayed on a computer screen. *Physiol. Res.* 55(1):105-13.

Rescorla R., A., Wagner A. R. (1972). A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. *Classical Conditioning II: Current Research and Theory*, pages 64-69. New-York: Appleton-Century-Crofts

Pastalkova E, Kelemen E, Bures J. (2003) Operant behavior can be triggered by the position of the rat relative to objects rotating on an inaccessible platform. *Proc Natl Acad. Sci USA*, 100(4):2094-9.