A Robot Based Model of the Pre-Frontal Cortex: The Mechanisms of Rule Learning and Rule Switching

Armin Duff[1] and Paul F.M.J. Verschure[1]

[1]University Pompeu Fabra

Humans and animals are able to make near optimal use of their knowledge to achieve their goals. This ability is seen as a hallmark of intelligent behavior [1]. Rule learning describes a main subset of intelligent behavior. On a neuronal level rule learning has been shown to critically depend on the pre-frontal cortex (PFC) [2]. Two main properties of the PFC critical for the expression of rules are sustained activity and activity modulation [2]. How this organization facilitates rule learning and switching in real time is however not yet known.

In our work we investigate how rule learning can be modeled using sustained activity and activity modulation following a conceptual approach using realistic robot simulations. Our model is based on the Distributed Adaptive Control (DAC) [3] architecture for behavioral control. DAC comprises of three tightly coupled layers of behavioral control; reactive, adaptive and contextual. The reactive and adaptive layers provide the robot with basic behavior supporting simple tasks. The contextual layer coordinates ongoing behavior in the context of behavioral plans. Here we develop a new contextual control layer able to express and learn rules. The contextual layer is based on a group of laterally connected model neurons. The sustained activity of these neurons is driven by perceptual inputs and modulated by lateral connectivity and rewards. A competition mechanism selects the neurons with the highest activity which subsequently trigger one the possible actions. The model is validated in a T-maze rule learning task analog to delayed conditional discrimination [4]. The robot has to choose one arm of the T-Maze according to a cue color patch situated at the start of the maze. After the completion of a predefined number of successful trials the reward contingency changes. This paradigm allows to validate the rule learning and rule switching capabilities of the model.

Our experiments show that sustained activity and activity modulation are enough to learn and express sequential rules in real time. A reliable performance could be achieved by including a mechanism to ensure the stability of the reward function. This stability is explicitly needed in real world environments where the model has to cope with high noise levels. In addition we showed that our model can efficiently adapt to changing reward contingencies which means that rule switching is significantly faster than learning the rules initially. A detailed analysis of the reward function shows that fast forgetting is essential for fast rule switching. The behavioral results show that the DAC framework is applicable for changing environments and forms a promising basis to accomplish complex real world tasks.

References

[1] Gallistel, C.R., *The Organization of Learning*. 1990, Cambridge: MIT Press.
[2] Miller, E.K., Freedman, D.J. and Wallis, J.D., *The prefrontal cortex: categories, concepts and cognition.* Philos Trans R Soc Lond B Biol Sci, 2002. **357**(1424): p. 1123-1136.
[3] Verschure, P.F.M.J., Voegtlin, T. and Douglas, R.J., *Environmentally mediated synergy between perception and behaviour in mobile robots.* Nature, 2003. **425**(6958): p. 620-624.
[4] Herremans, A.H., Hijzen, T.H. and Slangen, J.L., *Validity of a delayed conditional discrimination task as a model for working memory in the rat.* Physiol Behav, 1994. **56**(5): p. 869-75.