

# Planning in a Rule Learning System

Armin Duff<sup>1</sup> and Paul F.M.J. Verschure<sup>1,2</sup>

<sup>1</sup>Laboratory of Synthetic Perceptive, Emotive and Cognitive Systems, Institute of Audiovisual Studies, Universitat Pompeu Fabra, <sup>2</sup>Catalan Institute of Advanced Studies (ICREA)

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. One of the core elements of intelligent behavior is planning. On the neuronal level, planning has been shown to involve the pre-frontal cortex (PFC). Two important properties of the PFC critical for planning are 1) sustained activity through lateral connectivity and 2) synaptic modulation mediated through reward. In the context of the Distributed Adaptive Control (DAC) Architecture [1], we investigate to what extent these two mechanisms combined with a rule learning system is sufficient to support real-time planning in behaving systems.

The Distributed Adaptive Control (DAC) architecture comprises three tightly coupled layers; reactive, adaptive and contextual. The reactive and adaptive layers provide the robot with basic behaviors to support simple tasks. The contextual layer coordinates the ongoing behavior in the context of higher level behavioral goals. Recently we extended the contextual layer following the organizational principles of the PFC, accounting for rule learning and rule switching. With this extension we can now incorporate planning capabilities in the contextual layer. The core of the contextual layer is a group of laterally interconnected neurons. The sustained activity of these neurons is driven by perceptual inputs and modulated by lateral coupling and a dopamine like reward mechanism. A competition at the contextual level selects the neurons with the highest activity, i.e. relevance given the current input state and the context, that in turn trigger the Bayes optimal behavioral action [2]. The lateral connections of the contextual layer can be seen as a restricted world model and facilitate planning through propagation of activity. The validity of this acquired world model is assessed in both formal tasks and open-environment foraging.

Our experiments show that sustained activity and activity modulation, as found in the PFC, can be used to learn and express sequential rules in real time, and that these can be further extended to support planning. The behavioral results show that the DAC framework is suitable for dynamical environments and constitutes a promising basis towards solving complex real world tasks.

1. Verschure, P.F.M.J., T. Voegtlin, and R.J. Douglas, *Environmentally mediated synergy between perception and behaviour in mobile robots*. Nature, 2003. **425**: p. 620-624.
2. Verschure, P.F.M.J. and P. Althaus, *A real-world rational agent: Unifying old and new AI*. Cognitive Science, 2003. **27**(4): p. 561-590.