Mental representation and cognitive behaviour – a recurrent neural network approach

SIMONE KÜHN & HOLK CRUSE, Dep. of Biological Cybernetics, Bielefeld, Germany, Simone.Kuehn@uni-bielefeld.de

A necessary precondition for the control of behaviour of a higher cognitive level is the ability to build up mental representations of external situations. To cope with this task we propose a way of using and training a special type of recurrent neural networks.

Each organism must have the ability to move around in the environment in order to find food to survive and mates to reproduce, thus faces the problem of *sensorimotor control*. Instead of reacting to the external environmental information directly (considering a technical system as, e.g., Braitenberg-vehicles) the organism can use it to build up internal representations[1] of the external situation which can still be available after the sensory input has already vanished. By means of these internal representations the behaviour can be decoupled from direct control of the environment which enables the organisms, e.g., to react to features of the world that are not present at the moment or to plan ahead [3]; these are amongst others prerequisites for any kind of higher cognitive abilities as goal-oriented sequential behaviour, which is not genetically pre-programmed.

Different models for cell assemblies which can be taken to realise such internal representations have been proposed [e.g. 4,6,7]. Here, we present a recurrent neural network based on the principle of MMC nets [2] which were primarily used to solve motor tasks as arm movements or six-legged walking. In these tasks the nets cope with the problem of sensorimotor integration. A step towards a more representational level is to decouple the motor output: then the nets can serve as internal body models. Up to now the synaptic weights within the net have been fixed according to the equations forming the basis of the network. Hence, this body model can be called "innate".

In contrast, when an organism faces a new situation comprising different objects, the mental model must emerge starting with a "naïve" recurrent network with small random weights. Therefore, we propose a learning mechanism that leads to weight distributions which produce fixed-point attractors and thus represent external situations, a first step towards controlling adaptive cognitive behaviour, as e.g. language sequences. As the nets build up representations of current situations, these internal models are directly grounded in the external world [1]. Additionally, subpopulation of neurons which are embedded in a larger recurrent network and which code for the features of a given situation could be activated without a spread of activation throughout the whole net. This is due to a new learning algorithm called *dynamic delta rule* that comprises an extension of the classical delta-rule.

As a further step a way is shown how this static information can be transformed into sequential information. By slightly changing the structure of the recurrent network it is even possible to train it so that also *events* (as opposed to objects being part of a situation as considered so far), for example motion events, can be represented which are – concerning language – mediated by the verb of a sentence.

[1] CANGELOSI,A. (2001). Evolution of communication and language using signals, symbols and words. *IEEE Transactions in Evolution Computation 5,* 93-101.
[2] CRUSE,H., STEINKÜHLER,U., & BURKAMP,C. (1998). MMC - a recurrent neural network which can be used as manipulable body model. In: *From Animal to Animats 5*, R. Pfeifer, B. Blumberg, J.-A. Meyer & S.W. Wilson (eds.), Cambridge, MA: MIT Press, 381-389.
[3] CRUSE,H. (2003). The evolution of cognition - a hypothesis. *Cognitive Science 27*, 135-155.
[4] HOPFIELD,J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA 79*, 2554-2558.
[5] STEELS,L. (1995). Intelligence - Dynamics and Representations. In: *The Biology and Technology of Intelligent Autonomous Agents*, L. Steels (ed.), Berlin: Springer.
[6] TANI,J. & NOLFI,S. (1999). Learning to perceive the world articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Networks 12*, 1131-1141.
[7] WOLPERT,D.M. & KAWATO,M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks 11,* 1317-1329.

[1] The term *representation* is used here in the broad sense of STEELS [5] as being "physical structures (for example electro-chemical states) which have correlations with aspects of the environment".