Papers from the AAAI Fall Symposium (FS-09-01)

Neural Network Architecture for Crossmodal Activation and Perceptual Sequences

Magnus Johnsson[1], Christian Balkenius[1] and Germund Hesslow[2]

[1]Lund University Cognitive Science, Sweden
[2]Department of Experimental Medical Science, Lund, Sweden

Abstract

A self-organizing neural network is described that can associate between different modalities and also has the ability to learn perceptual sequences. This architecture is a step towards the development of a complete agent containing simplified versions of all major neural subsystems in a mammal. It aims at exploring as well as takes inspiration from the idea that cognitive function involves an internal simulation of perception and movement. We have tested the architecture in simulations as well as together with real sensors with very encouraging results.

Perceptul activity in the brain is normally elicited from the sense organs, but it has recently been argued that a crucial aspect of cognitive function in biological organisms is the ability to simulate perception, that is, to elicit perceptual activity from other brain areas rather than via sensory afferents (Hesslow, 2002; Grush, 2004). Among other things, perceptual simulation could explain the appearance of an inner world (Hesslow and Jirenhed, 2007) and also how an organism could try out various courses of action 'in the mind'. One plausible source of simulated perceptual activity might be activity in another sensory modality. A familiar example is when the texture of an object that is felt in the pocket can evoke visual images/expectations of the object. A more dramatic biological illustration is the McGurk-MacDonald effect. If you hear a person making the sound /ba/ but the sound is superimposed on a video recording on which you do not see the lips closing, you may hear the sound /da/ instead (McGurk and MacDonald, 1976).

Accordingly, a bio-inspired autonomous robot should be able to learn on its own and to automatically develop sensory representations of its different sensory modalities. These representations should also co-develop so that suitable activity in some modalities elicits appropriate activity in other sensory modalities as well. It would also be desirable with an ability to remember perceptual sequences, and that a corresponding sequence of activity evokes a suitable sequence of activity in the other modalities of the robot.

With these traits a robot learning to navigate through an environment could remember a sequence of e.g. visual landmarks. Imagine that the robot has learned appropriate motor actions associated with these visual perceptions. When learning sequences of visual landmarks, the robot has simultaneously learnt sequences of perceptions in other modalities, e.g. tactile landmarks. The memory of a sequences might be of different quality, e.g. because of noise or sensor limitations, but with the ability for cross-modal activation the sequential memories are reinforced in each modality. Moreover, if the robot suddenly lacks input to some modality, e.g. the visual modality when the environment turns dark, then the robot will still be able to elicit associated motor actions. This is so because appropriate activity in its visual representation will be evoked by the activity in its other modalities. With the architecture proposed by us, the robot would actually be able to extensively act without any sensory input at all, i.e. by internally simulating sequences of perceptions likely to follow (Hesslow, 2002). Some experiments with implementation of simulation in a robot have been done (Jirenhed et al., 2001; Ziemke et al., 2005).

To provide a robot with the abilities sketched above we propose a novel neural network architecture, the Associative Self-Organizing Map (A-SOM), uppermost Fig. 1. The A-SOM is based on the ordinary Self-Organizing Map (SOM) (Kohonen, 1988) but also learns to associate its activity with (possibly delayed) additional inputs, e.g. the activities of a number of external SOMs or A-SOMs. It consists of a grid of neurons with a fixed number of neurons.

Each neuron has multiple sets of weights, one for main input (which is similar to the input of an ordinary SOM) and one for each ancillary input. All neurons receive both main input (e.g. from a sensor), and ancillary inputs, i.e. inputs from the associated representations (the other modalities) or the A-SOMs activity from previous iterations.

Each neuron calculates activities for its main input and for each ancillary input. The main input activity is calculated in way similar to the ordinary SOM, with dot product as the similarity measure. Also the adaptation of the weights corresponding to main input are calculated as in an ordinary SOM, i.e. so that the neuron with the most similar weight vector and the neurons in its vicinity are adjusted.The ancillary activities of a neuron are calculated using dot product and are adjusted by the delta rule to approach the main activity. The total activity of a neuron is calculated by averaging the main activity and the ancillary activities.

By connecting the total activity of the A-SOM back to itself as an ancillary input with a time delay the A-SOM is turned into a memory of perceptual sequences. This is so because then the ancillary weights will have learned to evoke activity based on the previous activity in the A-SOM.

We have tested the A-SOM in several simulations. In one experiment we connected an A-SOM to two ancillary SOMs and trained all three neural networks with a set of randomly generated samples (Johnsson et al., 2009a). We tested the system with the training samples as well as with a new set of randomly generated samples in all possible constellations of inputs, i.e. one, two or all three neural networks received input. The ability of the A-SOM proved to be good, with 100% accuracy with the training set and about 80-90% accuracy in the generalization tests, depending on which constellation of inputs which was provided to the system.

In another experiment we focused on the use of the A-SOM as a memory for perceptual sequences (Johnsson et al., 2009b). This is useful for a robot because if its sensory input is interupted it can continue anyway by anticipating the sequence of perceptions likely to follow. This experiment was accomplished by using the total activity of the A-SOM as time-delayed ancillary input to itself (lowermost Fig. 2). In this way we showed that a system of two A-SOMs (one with recurrent connections) was able to produce appropriate sequences of activation in both A-SOMs even when receiving no more input. This seemingly continued indefinitely, with only cyclical deterioration.

We have also tested the A-SOM together with a couple of real sensors (texture/hardness) (Johnsson and Balkenius, 2008). This system developed representations for texture as well as hardness. It was able to discriminate individual objects based on input from each modality and to discriminate hard from soft objects. In addition input to one modality could trigger an activation pattern in the other modality, which resembled the pattern of activity the object would yield if explored with the sensor for this other modality.

The presented neural network architecture is a step in our exploration of the simulation hypothesis of cognitive function (Hesslow, 2002). It is also aimed as a step towards the implementation of a complete agent containing simplified versions of all major neural subsystems in a mammal, which we intend to carry out using the brain modelling infrastructure Ikaros (Balkenius et al., 2008).

In the near future we intend to test the A-SOM with several sets of recurrent ancillary input connections with different time delays, which might improve the capacity for remembering perceptual sequences. We will also try to extend the presented ideas to include motor neural networks.

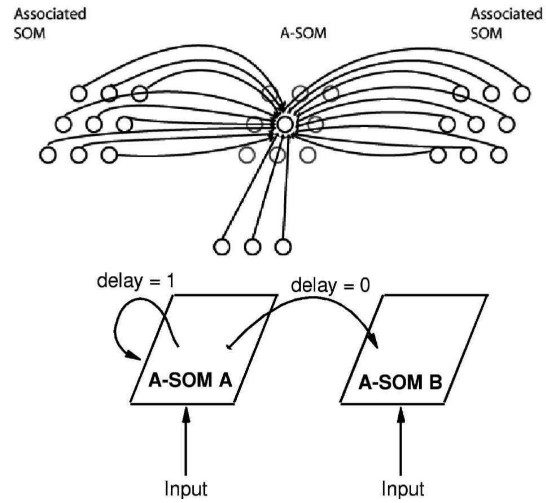

Figure 1: Uppermost: The connectivity of the A-SOM. Each neuron receives two kinds of input, main input and ancillary input. Activity can be triggered by main input or by activity in associated SOMs. Lowermost: A system with two A-SOMs, whereof one with recurrent connections. This system develop representations of two input spaces. The A-SOM B learns to associate its activity with the activity of A-SOM A, i.e. proper activity can be evoked in the A-SOM B even if it does not receive any ordinary input. This is similar to cross-modal activation in humans. The A-SOM A learns the sequence of activity presented during training.

References

Balkenius, C., Morén, J., Johansson, B., and Johnsson, M. (2008). Ikaros: Building cognitive models for robots. In Hülse, M. and Hild, M., editors, *Workshop on current software frameworks in cognitive robotics integrating different computational paradigms (in conjunction with IROS 2008), Nice, France*, pages 47–54.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery and perception. *Behav. Brain Sci*, 27:377–442.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends Cogn Sci*, 6:242–247.

Hesslow, G. and Jirenhed, D.-A. (2007). The inner world of a simple robot. *J. Consc. Stud.*, 14:85–96.

Jirenhed, D.-A., Hesslow, G., and Ziemke, T. (2001). Exploring internal simulation of perception in a mobile robot. In Arras, Baerveldt, Balkenius, Burgard, and Siegwart, editors, *Fourth European Workshop on Advanced Mobile Robots (Eurobot 01)*, pages 107–113.

Johnsson, M. and Balkenius, C. (2008). Associating SOM representations of haptic submodalities. In Ramamoorthy, S. and Hayes, G. M., editors, *Towards Autonomous Robotic Systems 2008*, pages 124–129.

Johnsson, M., Balkenius, C., and Hesslow, G. (2009a). Associative self-organizing map. In *To appear in the proceedings of the ICNC 2009.*

Johnsson, M., Balkenius, C., and Hesslow, G. (2009b). Learning perceptual sequences with associative self-organizing maps. *Neurocomputing (submitted).*

Kohonen, T. (1988). *Self-Organization and Associative Memory*. Springer Verlag.

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264:746–748.

Ziemke, T., Jirenhed, D.-A., and Hesslow, G. (2005). Internal simulation of perception: A minimal neuro-robotic model. *Neurocomputing*, 28:85–104.