

# Internal Simulation of an Agent's Intentions

Magnus Johnsson and Miriam Buonomante

**Abstract** We present the Associative Self-Organizing Map (A-SOM) and propose that it could be used to predict an agents intentions by internally simulating the behaviour likely to follow initial movements. The A-SOM is a neural network that develops a representation of its input space without supervision, while simultaneously learning to associate its activity with an arbitrary number of additional (possibly delayed) inputs. We argue that the A-SOM would be suitable for the prediction of the likely continuation of the perceived behaviour of an agent by learning to associate activity patterns over time, and thus a way to read its intentions.

To interact means to understand what others want and to foresee their intentions. Verbal communication is not enough, but intentions must be inferred from behaviour. Humans are able to do this, but an interesting question is how this ability could be implemented in artificial agents. One way to approach this problem would be to look closely to how this ability could arise in humans.

It has been argued that humans and animals can simulate perceptions, different actions and evaluate their likely consequences [1, 2] by eliciting similar activity in perceptual and motor parts of the brain as if the actions and their likely perceptual consequences actually happened [3]. Among other things, internal simulation could also explain the appearance of an inner world [4]. We propose that humans

---

Magnus Johnsson  
Lund University Cognitive Science, Kungshuset, Lundagård, 22222 Lund, Sweden, e-mail: magnus@magnusjohnsson.se

Miriam Buonomante  
Department of Chemical, Management, Computer, Mechanical Engineering of the University of Palermo, Palermo, Italy e-mail: miriambuonamente@gmail.com

understand others' intentions by simulating what they would have done in a similar situation.

To provide an artificial agent with an ability for internal simulation we propose the Associative Self-Organizing Map (A-SOM) [5, 7], which is related to the Self-Organizing Map (SOM) [9]. The A-SOM is able to learn to associate its activity with the, possibly time delayed, activity of other neural networks.

By associating the activity of the A-SOM with its own earlier activity the A-SOM becomes able to remember perceptual sequences. This provides the ability to receive some initial sensory input and to continue to elicit the activity likely to follow in the nearest future even though no further input is received. This could be considered as sequence completion of perceptual activity over time. This has been tested in several simulations [6, 7, 8].

We propose that the A-SOM could be used in a cognitive architecture to predict the intentions of an observed agent.

This could be done because the A-SOM should be able to learn to associate its own activity over time elicited by the observed agent's behaviour. For example, the behaviour of the observed agent could be its particular way of moving towards certain goals. After enough observation of the agent, the A-SOM should have learnt the typical perceptual sequences associated with the agent's typical behaviours when it is moving towards different goals.

Then, due to the ability of the A-SOM for sequence completion of perceptual activity over time, it should be possible for it to internally simulate the sequence of activity likely to follow the activity elicited by the agent's initial behaviour. In this way the A-SOM should be able to reach the activity corresponding to the likely goal intended by the observed agent by internal simulation alone.

## References

1. Barsalou, L. W. (1999). Perceptual symbol systems, *Behavioral and Brain Sciences*, 577-609.
2. Grush, R. (2004). The emulator theory of representation, *Behavioral and Brain Sciences*, 377-442.
3. Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception, *Trends Cogn Sci*, 6, 242247.
4. Hesslow, G. and Jirenhed, D.-A. (2007). The inner world of a simple robot, *J. Consc. Stud.*, 14, 8596.
5. Johnsson, M., Balkenius, C., and Hesslow, G. (2009). Associative Self-Organizing Map, *Proceedings of IJCCI 2009*, Funchal, Madeira, Portugal, 363-370.
6. Johnsson, M., Gil, D., Balkenius, C. and Hesslow, G. (2010). Supervised Architectures for Internal Simulation of Perceptions and Actions, *Proceedings of BICS 2010*, Madrid, Spain.
7. Johnsson, M., Martinsson, M., Gil, D. and Hesslow, G. (2011). Associative Self-Organizing Map, In Mwasiagi, J.I. (Ed), *Self Organising Maps - Applications and Novel Algorithm Design*, INTECH, 603-626.
8. Johnsson, M., Gil, D., Hesslow, G. and Balkenius C. (2011). Internal Simulation in a Bimodal System, *Proceedings of SCAI2011*, Trondheim, Norway, 173-182.
9. Kohonen, T. (1990). The self-organizing map, *Proceedings of the IEEE*, 78, 9, 1464-1480.