

An associative model of cortical language and action processing

Andreas Knoblauch, Heiner Markert, Günther Palm

Department of Neural Information Processing, University of Ulm, D-89069 Ulm
knoblauch@neuro.informatik.uni-ulm.de

When words referring to actions or visual scenes are presented to humans, distributed networks including areas of the motor and visual systems of the cortex become active [3]. The brain correlates of words and their referent actions and objects appear to be strongly coupled neuron ensembles in defined cortical areas. Being one of the most promising theoretical frameworks for modeling and understanding the brain, the theory of cell assemblies [1, 2] suggests that entities of the outside world (and also internal states) are coded in overlapping neuron assemblies rather than in single ("grandmother") cells, and that such cell assemblies are generated by Hebbian coincidence or correlation learning.

One of our long-term goals is to build a multimodal internal representation using several cortical areas or neuronal maps, which will serve as a basis for the emergence of action semantics, and to compare simulations of these areas to physiological activation of real cortical areas. In this work we have developed a cell assembly-based model of several visual, language, planning, and motor areas to enable a robot to understand and react to simple spoken commands. The essential idea is that different cortical areas represent different aspects (and correspondingly different notions of similarity) of the same entity (e.g., visual, auditory language, semantical, syntactical, grasping related aspects of an apple) and that the (mostly bidirectional) long-range cortico-cortical projections represent hetero-associative memories that translate between these aspects or representations. This system is used in a robotics context to enable a robot to respond to spoken commands like "bot show plum" or "bot put apple to yellow cup". The scenario for this is a robot close to one or two tables carrying certain kinds of fruit and/or other simple objects. We can demonstrate part of this scenario where the task is to find certain fruits in a complex visual scene according to spoken or typed commands. This involves parsing and understanding of simple sentences, relating the nouns to concrete objects sensed by the camera, and coordinating motor output with planning and sensory processing.

References

- [1] D.O. Hebb. *The organization of behavior. A neuropsychological theory*. Wiley, New York, 1949.
- [2] G. Palm. Cell assemblies as a guideline for brain research. *Concepts in Neuroscience*, 1:133–148, 1990.
- [3] F. Pulvermüller. Words in the brain's language. *Behavioral and Brain Sciences*, 22:253–336, 1999.