The Tensor Brain Hypothesis

Volker Tresp, Sahand Sharifzadeh and Dario Konopatzki
LMU Informatik, Oettingenstraße 67, 80538 München, Germany

Perception has evolved from simple stimulus-reaction in lower animals to the ability of a deep analysis of sensory input in humans. An important capability, for example, is the comparison to previous experiences: if a certain event is very similar to a past event, and that past event triggered a certain action, it makes sense that the current event should trigger the same action. Another important function is the identification of concepts and their relationships: “a child, located on a swing” will trigger very different actions than “a child, running in front of a car”. Clearly a more refined perception is tightly linked to an improved understanding of the world, its schema, objects and their relationships, or as Goethe put it: “you only see what you know”. In this paper we argue that episodic memory, i.e., the faculty to recall and restore past events, and semantic memory, i.e., knowledge about the world, are by-products of an evolving perceptual system which developed to deal with an increasingly complex world: our hypothesis is that episodic memory and semantic memory did not initially evolve as separate memory functions but instead repurposed faculties developed in perception for a semantic decoding of sensor stimuli. Furthermore, working memory might have evolved out of the need to store information to improve perceptual decoding.

The work in this paper is based on the tensor memory approach (Tresp et al., 2015; Tresp & Ma, 2016) which is an extension to the hippocampal memory indexing theory (Teyler & DiScenna, 1986). The key concepts of that approach are sparse index representations for entities, relationships and time instances. Each index has an associated distributed embedding, and memory and perception are based on an interplay between both. Perception, episodic memory and semantic memory might evoke sub-symbolic associations, but they are also declarative, indicated by the abilities of humans to report verbally about perception and memory contents. The semantic decoding in the tensor memory has exactly that declarative nature!

Here we significantly modify and extend that model. In the tensor memory model, the calculations of conditional probabilities required for decoding require marginalization operations which are costly and might be difficult to realize with biological wetware. Also, several indices and their embeddings needed to be active at the same time, which might not be biologically plausible (binding problem) and the approach required units to implement multiplication. Here, we propose a layered approach, where the sensory information is processed by a working memory layer, a representation layer and an index layer. The operations can be described as a single recurrent neural network where semantic memory evolves as an identifiable functional module.

Our experiments demonstrate that semantic memory can evolve from perception as a distinguishable functional module.

References

Teyler, T. J., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral neuroscience*.

Tresp, V., & Ma, Y. (2016). The tensor memory hypothesis. In *NIPS Workshop on Representation Learning*.

Tresp, V., et al. (2015). Learning with memory embeddings. *NIPS Workshop on Representation Learning*.