Procedia Computer Science
Volume 88, 2016, Pages 139–144
7th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2016

The Virtual Reality of the Mind

John F. Sowa
Kyndi, Inc
sowa@bestweb.net

Abstract
In evolutionary terms, imagery developed hundreds of millions of years before symbolic or language-like systems of cognition. Even the most abstract reasoning in science and mathematics requires imagery: diagrams and written symbols supplement short-term memory, and richer imagery is essential for novel analogies and creative insights. A cognitive architecture must relate symbols to the perceptions and purposive actions of an embodied mind that interacts with the world and with other minds in it. This article reviews the evidence for an internal virtual reality as the foundation for the perception, action, and cognition of an embodied mind. Peirce's theory of signs is a unifying framework that relates all branches of cognitive science, including AI implementations. The result is a theory of virtual reality for cognitive architectures (VRCA) that spans the minds from fish to humans and perhaps beyond.

Keywords: Mental imagery, embodied cognition, virtual reality, cerebellum, semiotics, cognitive architecture

1 Symbols and Imagery

For years, the mainstream in AI ignored mental imagery or considered it a side effect of perception that is irrelevant to cognition. Good Old Fashioned AI (GOFAI) is based on symbols organized in language, logic, networks, rules, frames, or chunks. But the emphasis on symbols created more problems than it solved: the Chinese room (Searle 1980), symbol grounding (Harnad 1990), and the way animals relate their bodies to perception of and action upon the world.

After millions of years of leaping and swinging through trees, primates developed three-dimensional cognition with excellent hand-eye coordination. Modern humans have not lost those abilities. Note the feats of Olympic gymnasts or basketball players who can score three points while running through interference by the opposing team. To support that ability, their visual system must process two-dimensional snapshots of a dynamically changing 3-D world, anticipate likely changes, and respond appropriately. Since humans and apes can perform similar kinds of gymnastics, their brains must process the same kind of dynamic 3-D geometry. Either the apes have symbolic systems as advanced as humans, or both humans and apes use similar analog methods.

During the six million years from apes to humans, a modest increase in brain size came with Homo habilis about two million years ago. A significant increase came with Homo erectus about one mya.

Selection and peer-review under responsibility of the Scientific Programme Committee of BICA 2016

doi:10.1016/j.procs.2016.07.417

Deakin (1997) claimed that the need to extend and enhance a protolanguage stimulated "the co-evolution of brain and language." The greatest increase in modern humans is in the huge cerebral cortex, but the cerebellum and brain stem are similar to the apes'. Figure 1 shows the human cortex overlaid with a neurocognitive network by the linguist Sydney Lamb (2016). The areas in pink are highly active in fMRI or PET scans for tasks that involve language semantics; the gray areas are less active for those tasks (Binder et al. 2009).

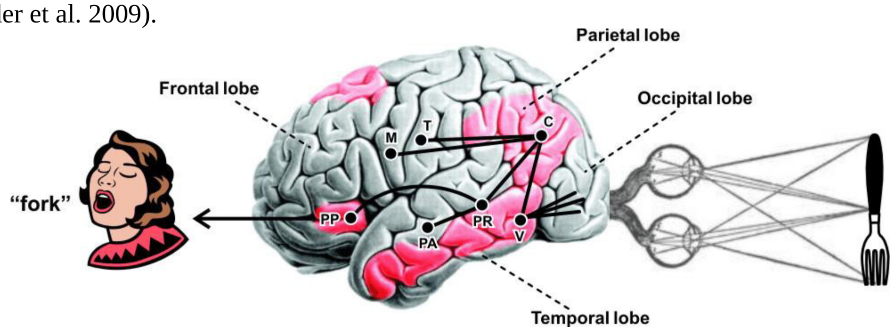

Figure 1. Areas of the left hemisphere that are active in language

The network in Figure 1 shows links from the image of a fork in the primary visual cortex to Broca's area for pronouncing the word fork. According to Lamb (2010), each labeled node represents a cortical column. Node C is a column for the concept of a fork. He placed it in the parietal lobe, which has links to the primary projection areas for all sensory and motor modalities. For the image of a fork, C has a link to node V, which connects to percept nodes in the occipital lobe. For the tactile sensation of a fork, C links to node T in the sensory area for the hand. For the motor patterns for manipulating a fork, C links to node M in the motor area for the hand. For the word fork, C links to node PR in Wernicke's area. Then PR links to node PA for recognizing the sound and to node PP in Broca's area for pronouncing the phonemes.

The primary sensorimotor areas are among the gray areas in Figure 1. For each body part, the sensory area contains a topographic (point-to-point) map from the skin, and the motor areas map to the muscles that control the body parts. The parietal lobes are among the association areas that expanded rapidly in the evolution from primitive mammals to apes and humans. To explain "the nature and development of imagery and verbal symbolic processes," Allan Paivio (1971) proposed a dual-coding theory (DCT) with a symbolic verbal system that maps to and from nonverbal imagery.

If there are two codes, the next question is whether they are processed by the same methods. In ACT-R (Anderson et al. 2004), production rules are symbolic if-then rules. Images must be mapped to symbols before they can be processed by those rules. In DCT, logogens (symbols) and imagens (percepts or larger images) may be stored and processed by the same mechanisms (Paivio 2007). Marvin Minsky's Society of Mind (1986) supports an a open-ended variety of modules and forms of representation. To connect different modules with different representations, Minsky proposed a system of K-lines (knowledge links), which allow modules at opposite ends of a K-line to use different representations. A module may interpret messages received via K-lines without any information about the internal representations of the sending modules. In his Emotion Engine, Minsky (2006) proposed emotions as the driving forces that motivate the modules and determine the goals to be achieved.

Since these systems address different aspects of cognition, they could be related as components of a larger framework. ACT-R, for example, might be extended to support both codes of DCT. The K-lines of Minsky's Society of Mind might represent the same nerve fibers as the links in Lamb's networks. In Figure 1, for example, the concept node C has long links across different lobes. Node T has links from the hand (afferent nerves); node M has links to the hand (efferent nerves); node V links to nodes for visual percepts; and node PP links to nodes that control muscles for producing phonemes. More research is needed to relate the details, but the experimental evidence for each of these systems could be compatible with a larger framework that addresses the role of mental imagery.