Robotic Models of Spatial Biases in Categorization

Anthony F. Morse, Tony Belpaeme, Angelo Cangelosi
Centre for Robotics and Neural Systems, University of Plymouth

At the heart of all sensorimotor theories of cognition is the claim that perception is to a large degree based upon the use of sensorimotor knowledge in predicting the future sensory consequences of an action or actions, either overtly executed or covertly simulated. Such an embodied account of perception is supported by a large number of psychology experiments exposing various bodily biases in categorization and recall. For example in a series of experiments conducted by Linda Smith and Larissa Samuelson (REF in press) children between 18 and 24 months of age are repeatedly shown two different objects in turn, one consistently presented on the left, and the other consistently presented on the right. After several presentations of the objects, the child's attention is drawn to one of the presentation locations and the linguistic label "MODI" is spoken in the absence of either object. Finally both objects are presented, overlapping, in a new location and the children are asked to find the modi. The significant majority of children select the spatially correlated object even though the name was presented in the absence of the object named. The results of this and similar experiments challenge the hypothesis that names are linked to the thing being attended to at the time the name is encountered. Moreover, in these experiments changes in posture from sitting to standing eradicate the effect, while other visual or auditory distracters do not.

We have developed a neural robotic model of sensorimotor learning displaying this and realted phenomenon with the iCub humanoid robot. The model uses the body posture of the robot as an association 'hub', making a strong connection to the sensorimotor literature as actions, here interpreted as changes in body posture, also have the ability to prime all the information associated with that new position, such as what the agent would expect to see, similarly multimodal sensory input primes actions as subsets of body posture changes. At the current stage of modeling information from each modality is processed in a variety of ways and mapped using self-organizing maps (SOM). For example, the body map of the iCub robot has 54 inputs, each being the joint angle of a single joint, while auditory input analysed for voice processing by the open source Sphinx library, the output of which is similarly mapped into another SOM. The average RGB value of the central region of the visual input (fovea) is also projected into a SOM generating a colour map. The SOMs are initialized using random values in the appropriate ranges until the SOM's have stabilized. An additional mechanism scans whole images for connected regions of change and directs the robot to orient to the largest change (moving eyes, head, and torso). We then replicated several experiments including the experimental set up used by Smith and Samuelson (in press) and linked the activity of the colour map and the auditory words to the body map using positive Hebbian connectivity where the connections are only modifiable from the current active body posture node. As the maps are linked together in real time based on the experiences of the robot, strong connections between objects typically encountered in particular spatial locations, and hence in similar body postures / orientations build up. Similarly when the word 'modi' is heard, it is also associated with one such body posture. Finally at the end of the experiment, when the robot is asked to 'find the modi' , activity in the 'modi' node spreads to the associated posture and on to the colour map node(s) associated with that posture. The result is to prime particular nodes in the colour map which then filter the whole input image and the robot adjusts its posture (to a new location) centered on the region of the image most closely matching this colour profile. This is achieved using the same mechanism that detects and moves to look at regions of change in the image. This model is able to replicate human data from several experiments including Smith and Samuelson's experiments described here using the iCub robot. This model represents preliminary work investigating spatial biases in object categorization. Further work developing and extending this model as a model of sensorimotor learning is currently underway.

This work has been supported by the ITALK project, funded by the EU PFP Cognitive Systems and Robotics Unit (ICT Integrating Project 214668)

Smith, L.B., & Samuelson, L., (In Press) Objects in Space and Mind: From Reaching to Words. In Mix, K., Smith, L. B., & Gasser, M. (Eds). Thinking through Space: Spatial Foundations of Language and Cognition, Oxford, UK: Oxford University Press.