World Academy of Science, Engineering and Technology
International Journal of Cognitive and Language Sciences
Vol:10, No:08, 2016

A Hebbian Neural Network Model of the Stroop Effect

Authors : Vadim Kulikov

Abstract : The classical Stroop effect is the phenomenon that it takes more time to name the ink color of a printed word if the word denotes a conflicting color than if it denotes the same color. Over the last 80 years, there have been many variations of the experiment revealing various mechanisms behind semantic, attentional, behavioral and perceptual processing. The Stroop task is known to exhibit asymmetry. Reading the words out loud is hardly dependent on the ink color, but naming the ink color is significantly influenced by the incongruent words. This asymmetry is reversed, if instead of naming the color, one has to point at a corresponding color patch. Another debated aspects are the notions of automaticity and how much of the effect is due to semantic and how much due to response stage interference. Is automaticity a continuous or an all-or-none phenomenon? There are many models and theories in the literature tackling these questions which will be discussed in the presentation. None of them, however, seems to capture all the findings at once. A computational model is proposed which is based on the philosophical idea developed by the author that the mind operates as a collection of different information processing modalities such as different sensory and descriptive modalities, which produce emergent phenomena through mutual interaction and coherence. This is the framework theory where 'framework' attempts to generalize the concepts of modality, perspective and 'point of view'. The architecture of this computational model consists of blocks of neurons, each block corresponding to one framework. In the simplest case there are four: visual color processing, text reading, speech production and attention selection modalities. In experiments where button pressing or pointing is required, a corresponding block is added. In the beginning, the weights of the neural connections are mostly set to zero. The network is trained using Hebbian learning to establish connections (corresponding to 'coherence' in framework theory) between these different modalities. The amount of data fed into the network is supposed to mimic the amount of practice a human encounters, in particular it is assumed that converting written text into spoken words is a more practiced skill than converting visually perceived colors to spoken color-names. After the training, the network performs the Stroop task. The RT's are measured in a canonical way, as these are continuous time recurrent neural networks (CTRNN). The above-described aspects of the Stroop phenomenon along with many others are replicated. The model is similar to some existing connectionist models but as will be discussed in the presentation, has many advantages: it predicts more data, the architecture is simpler and biologically more plausible.

Keywords : connectionism, Hebbian learning, artificial neural networks, philosophy of mind, Stroop
Conference Title : ICCSP 2016 : International Conference on Cognitive Science and Psychology
Conference Location : London, United Kingdom
Conference Dates : August 25-26, 2016