Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Cognitive Systems Research 6 (2005) 333–363
Cognitive Systems RESEARCH
www.elsevier.com/locate/cogsys

A cognitive model in which representations are images

Action editor: John Barnden

Janet Aisbett, Greg Gibbon *

Faculty of Science and Information Technology, The University of Newcastle, University Drive, Callaghan 2308, Australia

Received 22 March 2004; accepted 18 February 2005
Available online 23 May 2005

Abstract

Interpretations of images of the brain are starting to reveal the conceptual tasks in which the person was engaged at the time of imaging. Existing mathematical models can explain the patterns of activity observed in such images in terms of the coherent activity of large populations of neurons, but not in terms of cognition. This paper is an early investigation into how such patterns might provide the internal representations for a cognitive system. Probes, working memories and memories are all represented as images. The accompanying process model describes how attention is set according to the contents of working memory, how attention determines what parts of the probe are memorised, how memories are activated according to similarity to the probe in areas in attention, and how working memory is managed. The model is demonstrated on re-creations of classic simulations of recognition memory and categorisation.

Keywords: Internal representation; Cortical maps; Process models; Attention; Recognition memory; Categorisation

1. Introduction

Neuroimages provide an intriguing but baffling window into human cognition (Poldrack, 2000). Cortical activity captured in such images is not randomly distributed. It is spatially and temporally continuous, at least at the resolution of current imaging. The spatial organisation reflects underlying structure in perceptual input, such as the pitch of a sound, and hints at how higher-level concepts are represented. The temporal organisation hints at how perceptual stimuli are processed and how cognitive tasks are performed. As advances in imaging techniques improve temporal and spatial resolution and reduce noise, there is reason to believe that cortical images will be key to understanding how we think.

Images taken while subjects are engaged in a perceptual or conceptual task have been used to assign a range of cognitive responsibilities to parts

* Corresponding author.
E-mail address: greg.gibbon@newcastle.edu.au (G. Gibbon).

1389-0417/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogsys.2005.02.004

of the brain, in the burgeoning cognitive neuroimaging literature (e.g., Courtney, Petit, Haxby, & Ungerleider, 1998; Duncan & Owen, 2000; Engel & Singer, 2001; Haxby et al., 2001). Attempts have been made to predict from cortical maps what tasks the subjects were engaged in at the time of imaging (e.g., Wang, Hutchinson, & Mitchell, 2003). Models of coherent activity of large populations of neurons have been developed to explain cortical patterns of activity in terms of average neuron firing rates and the like (e.g., Erwin, Obermayer, & Schulten, 1992; Liley, Cadusch1, & Wright, 1999). But there have been few attempts to build these large-scale patterns of activity in the brain into computational models of cognition. Biological computational models – those that attempt to describe how processing is implemented in the brain – almost invariably employ neural nets (Norman, 2003). Such models have been very successful at predicting human performance on memorisation and conceptual tasks (e.g., O'Reilly, Norman, & McClelland, 1998; Kruschke, 1992). However, they are constructed from what are – in biological terms – tiny numbers of rudimentary models of neurons.

This paper develops a formal mathematical model of a cognitive system (Aisbett & Gibbon, 2001) into a computational model of cognition in which information is represented using continuous spatial functions, or images. The system is demonstrated by simulating two classic cognition experiments into recognition memory and categorisation that have previously been explained using attribute vectors as internal representations. The new model is called CIM, for Cortical Image Manipulation.

Many physiological functions can be considered to be points in an infinite dimensional space in which the dimensions are the spatial positions on a cortical layer. The two characteristics of *spatial organisation* and *infinite dimensionality* distinguish representations based on images in general, and cortical maps in particular, from traditional representational forms.

The link between the images in CIM and fields of biological activity is at this stage only by analogy. We do not attempt to map locations in the image plane to locations on cortical layers. Nor do we specify whether the images represent local average neuronal firing rates, or the electromagnetic fields proposed by McFadden (2002) as the source of consciousness, or any other physiological function definable on a surface. CIM is, however, intended to provide the groundwork for future use of such physiological images, or sequences of images. Although CIM is not connectionist, it could be biologically related to connectionist approaches because it is based on coherent activity of large neuron populations.

Before giving an overview of the CIM model, the next section presents terminology and constructs from cognitive modelling in order to motivate our later use of terms. The overview and the formal mathematical definitions follow in Sections 3 and 4.

The question of the type of image intended to be used as input in the modelling is addressed in Section 5: this is important since we do not propose at present to use actual cortical images. In brief, the images are spatially organised, they are used to represent abstract as well as concrete concepts, and they are analogical in the sense of having structure which carries information about the concept represented (Sloman, 1978). For example, any concept associated with magnitude – height for instance – might be represented by a family of radial functions in which the distance from the centre is related to magnitude and the absolute intensity is related to certainty. Representation of magnitude through the standard deviation of a circular Gaussian is illustrated for two values in Fig. 1(a). Fig. 1(b) is a possible representation of a plasticine cube (pictured to the left); the four quadrants going clockwise from the left top, respectively, contain descriptors of size and weight, hue, texture and feel, and shade (Aisbett & Gibbon, 2003). These examples and various properties of infinite dimensional and spatially organised representations are discussed further in Section 5.

Demonstrating how the CIM model performs any cognitive task requires simulation of stimuli and memories. Section 6 describes CIM application to word recognition memory. Shiffrin and Steyvers (1997) simulated words as random vectors, so we simulate them as images composed from randomly located circular Gaussians. Section 7 describes categorisation using the CIM model,