Neuroscience and Biobehavioral Reviews 48 (2015) 92–147

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review

A computational theory of hippocampal function, and tests of the theory: New developments

Raymond P. Kesner[a], Edmund T. Rolls[b,c,*]

[a] University of Utah, Department of Psychology, Salt Lake City, UT 84112, USA
[b] Oxford Centre for Computational Neuroscience, Oxford, UK
[c] University of Warwick, Department of Computer Science, Coventry CV4 7AL, UK

ARTICLE INFO

Article history:
Received 5 July 2014
Received in revised form 24 October 2014
Accepted 12 November 2014
Available online 20 November 2014

Keywords:
Hippocampus
Episodic memory
Recall
Autoassociation
Cortical Backprojections
CA3
CA1
Dentate granule cells

ABSTRACT

The aims of the paper are to update Rolls' quantitative computational theory of hippocampal function and the predictions it makes about the different subregions (dentate gyrus, CA3 and CA1), and to examine behavioral and electrophysiological data that address the functions of the hippocampus and particularly its subregions. Based on the computational proposal that the dentate gyrus produces sparse representations by competitive learning and via the mossy fiber pathway forces new representations on the CA3 during learning (encoding), it has been shown behaviorally that the dentate gyrus supports spatial pattern separation during learning. Based on the computational proposal that CA3–CA3 autoassociative networks are important for episodic memory, it has been shown behaviorally that the CA3 supports spatial rapid one-trial learning, learning of arbitrary associations where space is a component, pattern completion, spatial short-term memory, and spatial sequence learning by associations formed between successive items. The concept that the CA1 recodes information from CA3 and sets up associatively learned backprojections to neocortex to allow subsequent retrieval of information to neocortex, is consistent with findings on consolidation. Behaviorally, the CA1 is implicated in processing temporal information as shown by investigations requiring temporal order pattern separation and associations across time; and computationally this could involve associations in CA1 between object and timing information that have their origins in the lateral and medial entorhinal cortex respectively. The perforant path input from the entorhinal cortex to DG is implicated in learning, to CA3 in retrieval from CA3, and to CA1 in retrieval after longer time intervals ("intermediate-term memory") and in the temporal sequence memory for objects.

Contents

---

[*] Corresponding author at: Oxford Centre for Computational Neuroscience, Oxford, UK.
E-mail address: Edmund.Rolls@oxcns.org (E.T. Rolls).
URL: www.oxcns.org (E.T. Rolls).

http://dx.doi.org/10.1016/j.neubiorev.2014.11.009
0149-7634/

1. Introduction

In this paper a computational theory of hippocampal function developed by Rolls (1987, 1989a,b,c, 1996b, 2010b, 2013b,c), Treves and Rolls (1992, 1994) and with other colleagues (Rolls and Stringer, 2005; Rolls et al., 2002) is refined and further developed, and tests of the theory based especially on subregion analyses of the hippocampal system are described. The relation of this theory to other computational approaches to hippocampal function is described. This paper follows the general format of the earlier joint paper by Rolls and Kesner (2006), but updates the theory and the tests of the theory. The aims of this paper are thus to update a particular computational theory of hippocampal function (which remains the only quantitative theory of hippocampal function in memory and its recall to the neocortex), and the predictions it makes about the different hippocampal subregions (dentate gyrus, CA3 and CA1); and to update the empirical tests of these predictions by especially subregion analysis of hippocampal function. The paper is intended to provide an updated position or landmark description of the theory, and how it has been tested, and this combination makes this paper a unique contribution.

The theory was originally developed as described next, and was preceded by work of Marr (1971) who developed a mathematical model, which although not applied to particular networks within the hippocampus and dealing with binary neurons and binary synapses which utilized heavily the properties of the binomial distribution, was important in utilizing computational concepts.[1] The model was assessed by Willshaw and Buckingham (1990). Early work of Gardner-Medwin (1976) showed how progressive recall could operate in a network of binary neurons with binary synapses. Rolls (1987) described a theory of the hippocampus presented to the Dahlem conference in 1985 on the Neural and Molecular Bases of Learning in which the CA3 neurons operated as an autoassociation memory to store episodic memories including object and place memories, and the dentate granule cells operated as a preprocessing stage for this by performing pattern separation so that the mossy fibers could act to set up different representations for each memory to be stored in the CA3 cells.[2] He suggested that the CA1 cells operate as a recoder for the information recalled from the CA3 cells to a partial memory cue, so that the recalled information would be represented more efficiently to enable recall, via the backprojection synapses, of activity in the neocortical areas similar to that which had been present during the original episode. This

---

[1] Marr (1971) showed how a network with recurrent collaterals could complete a memory using a partial retrieval cue, and how sparse representations could increase the number of memories stored (see also Willshaw and Buckingham, 1990). The analysis of these autoassociation or attractor networks was developed by Kohonen (1997, 1984) and Hopfield (1982), and the value of sparse representations was quantified by Treves and Rolls (1991). Marr (1971) did not specify the functions of the dentate granule cells vs the CA3 cells vs the CA1 cells (which were addressed in the Rolls (1989a,b) papers and by Treves and Rolls (1992, 1994)), nor how retrieval to the neocortex of hippocampal memories could be produced, for which a quantitative theory was developed by Treves and Rolls (1994). In addition, Treves and Rolls (1994) and Rolls and Treves (1998) have argued that approaches to neurocomputation which base their calculations on what would happen in the tail of an exponential, Poisson, or binomial distribution are very fragile.

[2] McNaughton and Morris (1987) at about the same time suggested that the CA3 network might be an autoassociation network, and that the mossy fiber to CA3 connections might implement 'detonator' synapses. However, the concepts that the diluted mossy fiber connectivity might implement selection of a new random set of CA3 cells for each new memory, and that a direct perforant path input to CA3 was needed to initiate retrieval, were introduced by Treves and Rolls (1992). Contributions by Levy (e.g. 1989); McNaughton (1991); Hasselmo; Lisman; McClelland et al. (1995), and many others, are described below.