Noname manuscript No.
(will be inserted by the editor)

Hot Coffee: Associative Memory with Bump Attractor Cell Assemblies of Spiking Neurons

Christian Robert Huyck · Alberto Arturo Vergani

Received: date / Accepted: date

Abstract Networks of spiking neurons can have persistently firing stable bump attractors to represent continuous spaces (like temperature). This can be done with a topology with local excitatory synapses and local surround inhibitory synapses. Activating large ranges in the attractor can lead to multiple bumps, that show repeller and attractor dynamics; however, these bumps can be merged by overcoming the repeller dynamics. A simple associative memory can include these bump attractors, allowing the use of continuous variables in these memories, and these associations can be learned by Hebbian rules. These simulations are related to biological networks, showing that this is a step toward a more complete neural cognitive associative memory.

Keywords: Spiking Neurons, Associative Memory, Cell Assemblies, Bump Attractor, Hebbian Learning.

1 Introduction

How does the brain represent concepts that are continuously valued, like height, weight, and temperature? How can these be included in the brain's associative memory. For example, what is the neural basis of the representation of hot coffee? Moreover, how can coffee be considered cold at one temperature, and another drink, say a coke, be considered warm at the same temperature?

This paper proposes a neural topology for a particular associative memory involving hot coffee, but the general association between a continuous value and two binary semantic values is readily replicable. A problem of multiple bumps arises when the semantic input causes a large associated range of the continuous value to be activated, but extra topology has been added that merges bumps into one (see

C. Huyck
Middlesex University
Tel.: +44-208-411-5412
E-mail: c.huyck@mdx.ac.uk

A. Vergani
Middlesex University

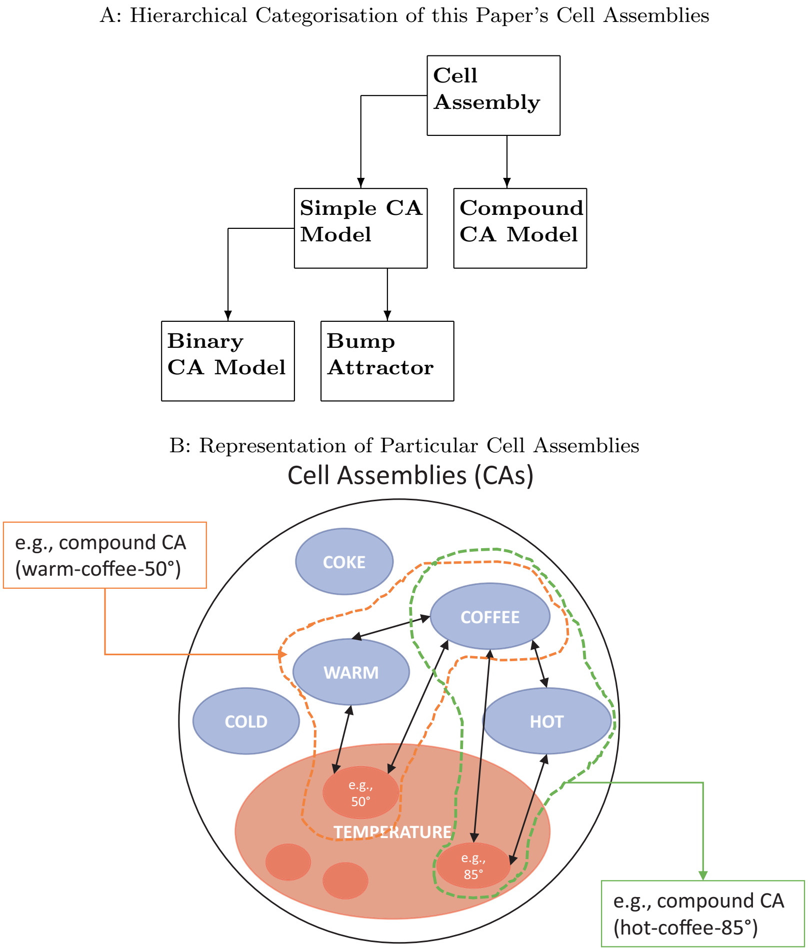

Fig. 1 Cell Assemblies from this Paper. 1A is a representation of categories of CAs used in this paper. All are simple models, with a novel form of bump attractor representing continously valued CAs. Compound CAs can be formed of simpler CAs. 1B shows binary CAs, such as Coffee, a series of CAs representing the continous value of temperature, and two compound CAs.

section 4). The associations can also be learned using Hebbian learning (see section 5). The resulting topology has the added benefit of merging in psychologically realistic times.

The authors firmly believe that concepts, like hot, coffee and particular temperatures, are represented in the brain by Cell Assemblies (CAs) (see section 2.3), and associations are represented either by synaptic connections or by CAs in the brain. In the simulations described in this paper, basic semantic concepts, like coffee, are represented by a simple version of CAs called binary CAs. Other concepts, like temperature, are represented by continuously valued networks that fall into a category called Winner Take All (WTA) networks, or more precisely, bump attractors (see section 2.1). These bump attractors support another simple version of CAs that differs from binary CAs. Finally, the associations are themselves CAs that are more than just their primitive components. All fire persistently, and thus can operate as short term memories. These can be seen in figure 1. Figure 1A shows the types of CA models used in this paper. The compound CAs, shown more fully in figure 1B, are formed by excitatory synapses from the neurons in the consituent CAs. While bump attractors are a common and long standing model, their use as components in associative memory has shown a problem that, as far as the authors are aware, is novel: large valued inputs lead to multiple bumps.

2 Literature Review

A great deal of literature relates to this paper, but this review concentrates on four bodies. The first is Winner Take All Networks and bump attractors in particular (section 2.1). The main cognitive component of the paper is the associative memory (section 2.2), and the bridge between the two is the Cell Assembly (section 2.3), which is the neural basis of concepts. The final literature review section (2.4) is on Hebbian learning, which is used in the simulations for learning associations.

2.1 Winner Take All Networks

The brain is a part of the central nervous system, which processes multi-modal information. Although there are several sources of stimuli (coming from inside and outside of the body), the brain selectively analyses this huge amount of information. The information is often ambiguous, so the brain must select one of the possible options. In simulated neural and connectionist systems, one model to select between options is a winner take all (WTA) system. One connectionist system that uses this model is the self organizing map (Kohonen, 1982). The map is made up of several nodes, and when an item is presented, the nearest node wins.

The stationary bump, which is another mechanism proposed for feature selectivity in the brain (Somers et al., 1995; Laing et al., 2001), has been used in spiking neural networks. The bump activity is an example of WTA neural behaviour because a group of co-firing neurons can be considered as winners of competition via inhibitory synapses. Adopting a general point of view, WTA networks in general and bump attractors in particular are a pattern formation process in a population with excitatory and inhibitory synapses; they work on the patterns of a stable grid (Wilson and Cowan, 1973), leading to an activity dependent neural group selection (Edelman, 1987).

In the typical stationary bump model, distance is considered either in one or two dimensions. There are local excitatory synapses, and more broad inhibitory synapses. The recurrent neural network is able to select particular neurons using inhibitory synapses that sustain the competition between neurons (e.g. Chen's surround inhibition (Chen, 2017)).

There is evidence that excitatory cells (i.e., principal neurons) are associated with specialized inhibitory cells (i.e., interneurons or secondary cells) that synapse to principal cells as well as other interneurons. The proper dynamics in the neural network can only be sustained if the excitatory behaviour of principal cells is modulated by the stopping function of secondary cells. If there were only excitatory neurons, their positive spikes could lead to an excitation that produces more excitation (an avalanche effect potentially leading to simulated epilepsy), and therefore, it would be difficult to observe transiently active groups of co-firing neurons such as, for example, CAs (see section 2.3).

This type of stationary bump is widely used in neural simulations. For example, it is used to manage a robot's direction in a path integration task (Kreiser et al., 2018).

2.2 Associative Memory

It is widely agreed that in the brain (and mind) concepts do not exist in isolation, but instead are associated with each other. These associations are part of the semantics of the underlying concepts, making an associative memory. This is a long standing psychological theory (Quillian, 1967).

This associative memory is the basis of priming effects (Collins and Loftus, 1975). If a concept is activated, it spreads its activation to associated concepts. The associative memory can be thought of as a symbolic Semantic Network, and Semantic Nets are widely used in AI for knowledge representation (Brachman and Schmolze, 1989).

Early versions of associative memories in simulated biological neural nets typically refer to associating vectors of firing neurons (Willshaw et al., 1969). An input vector of neurons, when fired once, causes an associated output vector to fire once. This work however is not particularly well suited for bridging the gap between biological neural behaviour and the emerging psychological behaviour of, in this case, associative memory, because individual concepts are not represented by a vector of neurons firing once, but by persistently firing Cell Assemblies.

There are more modern versions of associative memories. For example Chrysanthidis et al. (2019), like the simulations described below, use NEST, spiking neurons, and Hebbian learning to associate memories; this impressive system is aligned to biological spiking neurons and topology, but performs a rather weak cognitive associative task.

2.3 Cell Assemblies and Concept Representation in the Brain

Continuous concepts are mental representations of continuous phenomena like time, space, temperature, and force. These concepts share a common mathematical structure that is processed by specific anatomical regions. There is strong evidence that biological bump attractors are used to represent continuous phenomena, such as head direction in Drosophilia (Kim et al., 2017), and in rodents (Laurens and Angelaki, 2018; Gerstner et al., 2014).

It is less clear how associative memory is represented in the human brain, but the standard theory is that concepts are typically distributed over neurons from many different cortical (and perhaps subcortical) areas (Pulvermuller, 1999). For example, Martin (2007) collected evidence from functional neuroimaging studies about the storage in the cortex of salient properties of an object, like movement, shape and function; he found that those features are stored in separate sensory and motor systems, suggesting that concepts emerge from a weighted activity of property-based brain regions. Handjaras et al. (2016) investigated modality independent and category based organization of semantic knowledge in the human brain; they concluded that patterns of neural activity spread over a large semantic cortical network represent concepts independent from the modality of stimulus presentation.

In componential theories of lexical semantics, particular concepts are represented by features. One study shows key features that are particularly salient (Binder et al., 2016), with temperature being one of the 65 features represented. There is evidence that temperature is represented in the insular cortex (Craig et al., 2000), but the authors are unaware of any neural level study that shows the form of that circuit. It does seem reasonable to propose that temperature, being continuously valued, is represented by a bump attractor like head direction.

In his book, Hebb (1949) developed his famous synaptic learning rule, and used that rule to propose the CA as the neural basis of concepts. That is, semantic concepts, such as coffee are represented by CAs. A CA is a relatively small group of neurons that have high mutual synaptic strength that is formed by Hebbian learning. When some of the neurons fire, that mutual synaptic weight supports firing in the other neurons in the CA, and this allows a cascade of firing so that the neurons in the CA can fire persistently for a considerable amount of time (seconds). This firing is the neural basis of short term psychological memory.

In the intervening 70 years, there has been significant and growing evidence of the existence of CAs in brains (Singer et al., 1997; Harris, 2005; Buzsaki, 2010), and there is evidence of CAs in all major cortical areas (Huyck and Passmore, 2013). Neurons in a CA fire persistently, once activated, and fire synchronously.

While the neurobiological evidence is accumulating, there are not very good neural simulations of CAs. The authors have spent considerable time simulating CAs, used them in many tasks, and recent work has made extensive use of binary CAs (Huyck and Mitchell, 2018). In these simulated CAs, the neurons are either mostly firing, or none are firing, so it is binary, either on or off. This can be implemented with a well connected topology of neurons. Once the neurons start firing, they fire persistently until some external source shuts them off. This is obviously a poor model of CAs because, among many reasons, in a normal case, CAs would stop firing on their own, just like normal short term memories stop on their own.

None the less, binary CAs have been used in simulations of associative memory (Huyck and Ji, 2018). In this case, three concepts are associated, and when two become active, the associated third comes on. That is a 2/3 associative memory. This is an example of work in attractor networks to form associative memory (Lansner, 2009). Using spiking point neural models starts to bring together more biologically accurate simulations to manage more complex, and psychologically accurate neuropsychological simulations.

The elevated firing of a CA is the neural correlate of a short term or working memory item. In the case of the 2/3 associative memory items, two CAs, instantiating two concepts, are firing at an elevated rate, and they cause the third to fire at an elevated rate, retrieving the concept associated with both of the first two. There is evidence that bump attractors instantiate CAs for continuously valued phenomena. For instance, recordings of pre-frontal cortical neurons of monkeys in oculomotor delay response tasks is consistent with bump attractors (Wimmer et al., 2014) representing position. Similarly, single cell recordings of grid cells, are consistent with 2-D continuous attractors (Yoon et al., 2013). Bump attractors have been used to model hippocampal place cells (Stringer et al., 2002), and head direction cells (Redish et al., 1996). While these models, and indeed the models used in this paper have no short term plasticity, it seems that there is a sound basis for supporting the use of this type of plasticity in continuous attractor neural models for short term memory (Seeholzer et al., 2019). This paper uses a relatively simple topology that represents continuously valued concepts, and uses them as components of an associative memory.

2.4 Learning

While spiking neural networks possess many benefits, such as parallelism, perhaps their main benefit is their ability to learn. Learning in biological neural nets can be divided into three categories: structural plasticity, long term plasticity, and short term plasticity.

Most synaptic modification is thought to be Hebbian. That is, when a neuron tends to cause another to fire, the synaptic weight tends to increase. A simple computational Hebbian learning rule (Oja, 1982) forces the synaptic weight toward a weighted value of the likelihood that the post-synaptic neuron fires when the pre-synaptic neuron fires. If the modifier reflects the total weight of the synapses leaving a neuron, the rule is a pre-compensatory rule (Huyck and Mitchell, 2014), and if it reflects the total weight entering the post-synaptic neuron it is a post-compensatory rule. These rules normalise the weights keeping total synaptic weight stable, so they support homeostasis. Learning in section 5 uses pre and post-compensatory Hebbian rules.

3 Bump Attractors

Simulations in this paper use bump attractors consisting of neurons with excitatory synapses to adjacent neurons and inhibitory synapses to neurons beyond those that are excited. For representing continuous linear values, such as temperature, distance in this topology is linear, so that a neuron has excitatory synapses to two neurons on either side, and inhibitory synapses to four beyond those.

The neuron model is a leaky integrate and fire model. It follows Fourcaud-Trocmé et al. (2003) and is described more fully in Appendix A. An exploration of parameter values shows the behaviour of these bump attractors (see Appenix B). When there is an appropriate balance between excitation and inhibition, a small number of neurons fire persistently indefinitely. When there is too much excitatory strength, the activation spreads throughout the attractor and all neurons fire persistently. Insufficient excitatory or too much inhibitory strength does not support persistent firing.

This is standard bump attractor behaviour, but these attractors can also support multiple bumps firing simultaneously. Appendices B.1 and B.2 show the emergence of multiple bumps. In particular, these occur when a large number of inputs are sent to the bump attractor.

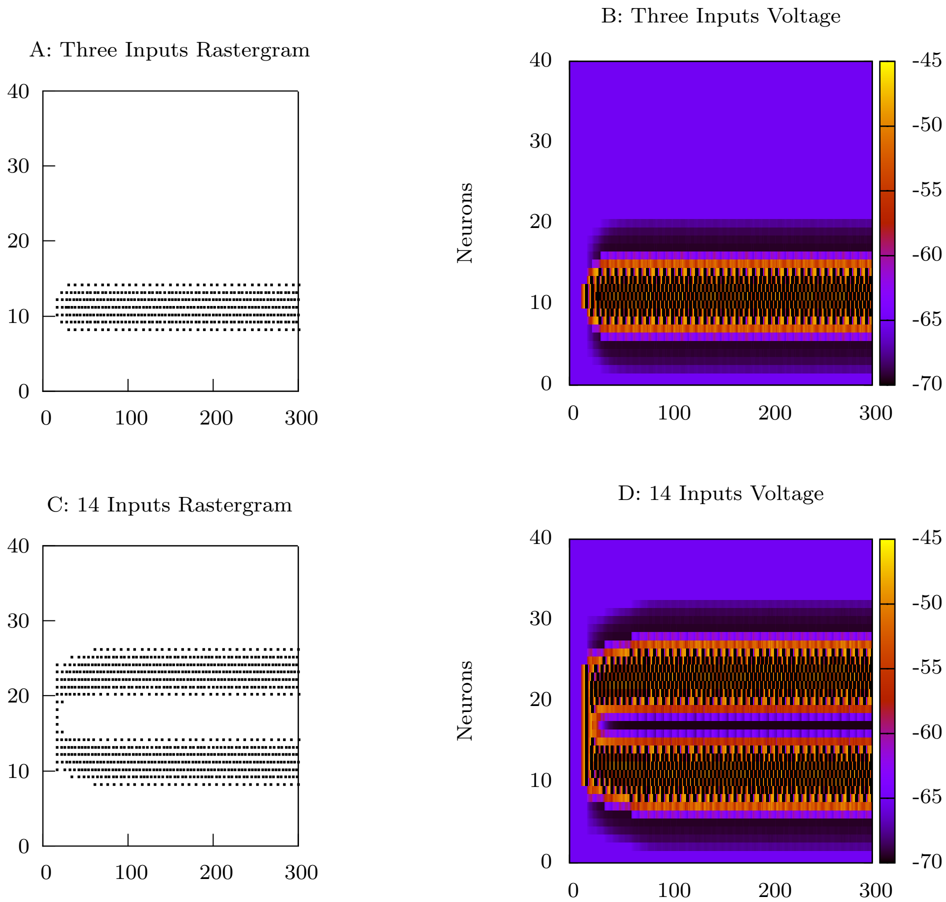

Fig. 2 A rastergram of a single bump (A), voltage signature of that bump (B), rastergram of two adjacent bumps (C), and the voltage signature of adjacent bumps(D). (A) shows that a small number of inputs leads to a single bump, while (C) shows that with a larger number of adjacent inputs two bumps emerge. The voltage signature in (B) shows the spiking neurons are osciallating between high and low voltage, while those on the edge are stimulated, but do not fire, and those further out have low voltage. (D) shows a similar effect to (B), but the neurons in between the bumps have a particularly low voltage.

An example of multiple emerging bumps is shown in figure 2. Figure 2A and 2B represent three neurons being sent initial spikes from outside the system to ignite the stable bump. The number of neurons that fire persistently remains largely the same as more neurons are initially spiked until 14 neurons are sent initial spikes. When 14 neurons are initially spiked, two bumps are generated. This behaviour is represented by figure 2C and 2D. Note how the initial neurons are not the only ones in the bump. The bumps have repelled each other and moved to include new neurons not initially stimulated.

4 Bump Attractor Repeller Dynamics

When bumps are at a distance, they do not influence each other, but when nearby they repel each other. From a cognitive standpoint, multiple bumps should be replaced with one so that only one value of the concept is active. Multiple bumps can be replaced by merging, and merging two bumps that arise from presenting a broad input to the stable bump can be broken into two subproblems: the first is overcoming the repeller effect around the two bumps (see Appendix C). This has been solved using a rather extreme topology with quite heavy excitatory and inhibitory weights. The second and easier problem, at least for these simulations, is moving distant bumps toward each other (Appendix C.1). Finally, once these two problems are addressed, they need to be combined into a single subnetwork.

A complete topology to move two bumps into one combines the initial bump attractor, the *Merge* subnet, to bring distant bumps close to each other, and the *Overcome* subnet, to replace the two adjacent bumps with one. The initial range of neurons fire, based on the input. If the range is long enough, two bumps of neurons on either end of the range fire persistently. The *Merge* subnet is activated, and causes the two bumps to move together. Once together, the merge subnet cannot overcome the inhibition from the bumps that are repelling each other, but the *Overcome* subnet fires, causing the two bumps to move into one stable state. As both the *Overcome* and *Merge* subnets require two bumps in the bump attractor to fire, none of their neurons fire, and the attractor is stable.

This is shown in figure 3. There is no spiking in either the *Merge* or *Overcome* subnet once the bumps are merged. Figure 3D shows the voltage of the *Bump* subnet. Initially, it follows the change of firing from the merge mechanism; once the bumps are adjacent, the repelling force of the two bumps is overcome by the *Overcome* subnet, and the resulting bump is stable. As in all simulations described in this paper, this happens in parallel, with all neurons updating each cycle.

5 Associative Memory

The bump attractor, and attendant *Merge* and *Overcome* subnets can be used to represent continuous values, and these can be used in an associative memory. An example is an associative memory for the temperatures of beverages, and the gross topology of the spiking network that implements this memory is shown in figure 4. There are two beverages, *Coffee* and *Coke*, and three temperature values *Hot*, *Warm* and *Cold*. These are all represented by binary CAs.

The *Input Temperature* subnet is a bump attractor, and the *Inhibition* subnet is used to reduce the possibility of both beverage temperature subnets becoming active.

The basic temperature associations are described in table 1. The temperatures and associated labels seem about right to the authors, and include a variety of ranges, which is good for expository purposes.

The idea is a two of three (2/3) associative memory. If two of the concepts are firing, they should ignite the third. Appendix D.1 describes a simple version of this system. The beverage temperature subnets have no internal connectivity, and the network shows largely the expected results for the 18 basic inputs. For instance, an input temperature of 36°-38° and *Coffee* causes the *Warm* CA to ignite and fire

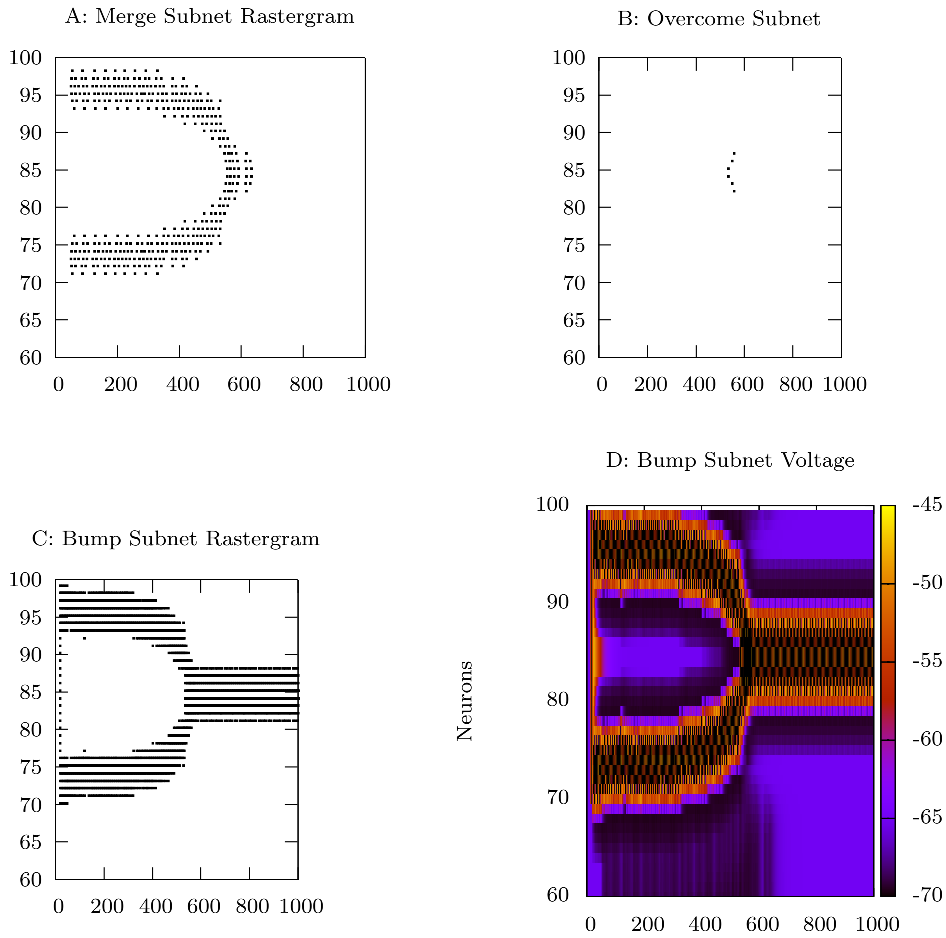

Fig. 3 Results from the full merging network, initially spiking neurons 70 through 99. All neurons below 60 behave identically. A rastergram of spiking behaviour for the Merge subnetwork (A), rastergram of spiking behaviour for the Overcome subnetwork (B), rastergram of spiking behaviour for the Bump Attractor subnetwork (C), and the voltage signature of the Bump Attractor (B). The merge and bump subnets' rastergrams show the initial two bumps, and them moving towards each other. Once close enough, (B) shows the Overcome subnet spiking, causing the two bumps in to merge. There is no subsequent firing in either Merge or Overcome subnets.

persistently. In this case, the *Input Temperature* and *Coffee* CA cause the associated *Coffee Temperature* neurons to fire, which in turn leads to the ignition of *Warm*.

However, this shows the problem of multiple bumps. When two semantic inputs are presented, for example, *Hot* and *Coffee*, the full temperature range is activated. As there is no internal connectivity in the *Coffee Temperature* subnet, they all fire persistently, but at a low rate. Making the beverage temperature subnets bump attractors leads to two bumps forming and persisting.

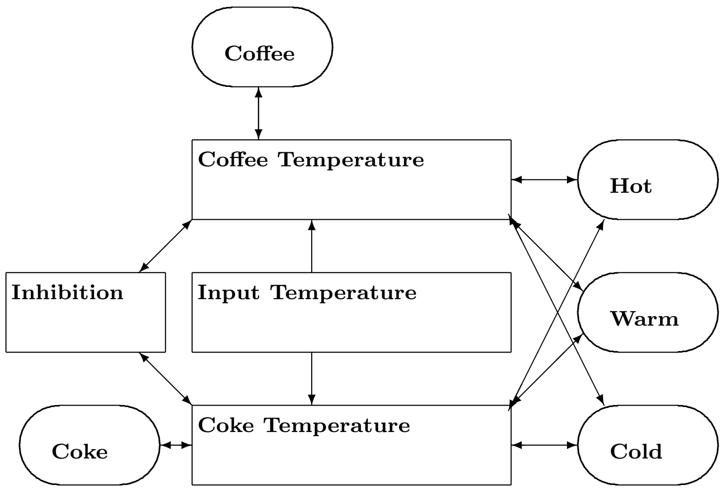

Fig. 4 Hot Coffee Gross Topology. Ovals represent binary CAs of 10 neurons that represent simple versions of semantic concepts. Boxes represent a large group of neurons. The *Input Temperature* is a bump attractor. The *Coffee Temperature* and *Coke Temperature* have no internal connectivity in the simulations of appendix D.1, but are bump attractors in appendices D.2 and D.3. Arrows represent synapses between the neurons in different boxes and ovals.

Semantic Pair	Temperature Range	Final Value	Time to Converge	Learned Result
Cold Coffee	0-24	10-14	1080 ms	10-14
Warm Coffee	25-69	45-49	7040 ms	47-51
Hot Coffee	70-99	82-87	463 ms	81-86
Cold Coke	0-9	2-7	90 ms (one bump)	2-6
Warm Coke	10-24	14-19	100 ms	10-14
Hot Coke	25-99	60-64	13920 ms	61-65

Table 1 Semantic Pair Input and Temperature Output. Time to converge refers to the time for the system with static synapses to converge. Note how larger ranges take significantly longer. The learned results column refers to plastic synapses learned from randomly distributed input.

Appendix D.2 describes the full associative memory with static synapses. The beverage temperatures each have a bump attractor and the attendant *Merge* and *Overcome subnets*. All results are correct, and the time to converge to a correct answer is shown in table 1. Note how wider temperature ranges take longer to converge, in the case of Hot Coke a cognitively implausible time of almost 14 seconds.

Fortunately learning the associations can lead to a solution that answers in a psychologically plausible time. Appendix D.3 describes the Hebbian learning mechanism used and several presentation regimes. When random temperatures for each associative pair are presented during training, the system learns all pairs, as shown in the final column of table 1. All answers converge in under 300 ms, which is a cognitively plausible time.

6 Discussion and Conclusion

This paper has focused on a particular stable bump attractor, and extending an associative memory topology around it to account for associating two binary concepts and one continuous concept. If two of three associated concepts are presented to the network, the third is retrieved. When a large range is presented to a bump attractor, as in the case when the temperature range associated with two semantic values is large, extra neural topology causes the two bumps to merge into one, with those neurons firing in a self-sustaining persistent manner. As the retrieval time is relatively brief, this could easily be used in a neural cognitive model.

This basic attractor network should be readily extendible to different local connectivity, different neural models and different time steps. Moreover, the overall associative memory will be generalisable. For instance, during development of the system while trying to resolve the merge problem, a 1-10 attractor was developed; unlike the seven neuron bump that emerged from the 2-4 topology, the 1-10 topology had an 11 neuron bump. In general, larger basins have larger bumps.

The particular associative network could be readily extended to more beverages. One could add, for instance, tea, hot chocolate, and orange juice by simply adding three sets of temperature attractors, the three semantic concepts, and the relationships. Perhaps more interesting and neuropsychologically plausible is the use of hierarchy in the concept structure. The authors earlier work on associative memory (Huyck and Ji, 2018) made use of hierarchical relations. In this case, there could be higher level categories, like fruit drinks, that had default values, that could later be refined. It should also be noted that association in these simulations have been done by synaptic connectivity; an independent concept is linked to two others by excitatory synapses; in the brain, particularly salient associations might also be CAs.

Similarly, other continuously valued concepts like time, weight and height could be used. Another reasonable extension would to combine two continuous attractors with binary semantic categories. For instance, one might combine weight and height to determine when someone is skinny or chubby. Interestingly, this might also apply in vowel recognition in the auditory cortex (Peterson and Barney, 1952) with the primary and secondary formants producing the vowel.

The authors' prior work with stable bump attractors (Nadh and Huyck, 2010) shows that they can readily be extended from the one dimensional attractors discussed above to two dimensions, and they can be further extended to higher dimensions. The bump attractors here represent a large number of simple CAs. These CAs can act as short term memories when firing, but CAs should do some sort of calculation, and once stable these do no calculation (Tetzlaff et al., 2015). However, the composite associations are also CAs, so *Hot Coffee* and 82° is a CA composed of three other CAs. When the full hot coffee network is presented with the semantic terms *Hot* and *Coffee*, the merge topology does the initial calculation of completing the association, but then does a calculation to convert the broad temperature range to a single coffee temperature CA. That is a relatively complex calculation, though once in that state the CA does no further calculation.

The literature about bump attractors and WTA functionality also investigates the thermodynamics of the network (see for example (Tkačik et al., 2015; Hahn et al., 2017; Pena et al., 2018)). Using the framework of dynamical system theory and related concepts (Meiss, 2007), this work could be extended by focusing on properties of attractors and repellers, stability and instability of the bumps, dynamics of pattern formation, and forking behaviour as bifurcation phenomena.

Attention should be drawn to weaknesses in the simulations. The first is that the underlying CAs (the binary CA, the bump CA, and the associative CAs) do not make good models of short term memory because they persist indefinitely. The authors are currently working on improving the basic CA model including neural models with adaptation and synaptic models using short term plasticity. A second weakness is learning. As bump attractors are found in the brain, they must be learnable, but it is not clear how the merge and overcome topologies could be learned; they could be thought of as a proof of concept. A third weakness is the lack of calculation after stability in the associative memory; a reasonable associative memory should spread activation, dynamically priming and moving to new concepts over time.

A note should be made about the terms winner take all networks and bump attractors. It seems the literature often equates the two, but as commonly used, winner take all refers to a single winner amongst multiple competitors. Take for instance the winner in a British Parliamentary election; there may be many competitors, but the one with the most votes wins. This is not what is being modelled in our bump attractor. If several neurons in different areas were given different stimulation, the one with the largest would not typically win.

The simulations in this paper have centred on stable bumps to represent large valued properties, beverage temperatures in particular. These have been included as components in associative memories, leading to an, as far as the authors are aware, novel problem of large valued inputs to the bump. While bump attractors are a common and long standing model, these large valued inputs force a reduction in the number of inhibitory connections so that the bump can have a small number of persistently firing neurons. These models are not good neuron for neuron models of the biological hot coffee representation. However, the point neural models are commonly used, albeit simple, models of biological neurons. Similarly, the topologies are relatively sparse, so that they could reasonably be subsets of the actual biological topology. However, by using the combination of simple bump attractors for the linear phenomenon with binary cell assemblies for semantic primitives, the new issue of broad input to the bump has been raised. It is hoped that this will inform future work that moves beyond these relatively simple neural network models and topologies to more biologically and psychologically informative models.

Developing simulations of simple neural circuits furthers understanding of their behaviour, as is the case with persistent bump attractors. These can be combined with other simple circuits to further understanding of more complex behaviour, such as associative memory. Hopefully, this will lead to a deeper understanding of more complex circuits, such as CAs, and eventually to the full brain.

Finally, the paper (and associated code[1]) should provide topologies, and an underlying theory for new topologies for improved associative memories in, for example, neural agents.

[1] Code can be found at http://www.cwa.mdx.ac.uk/NEAL/wta.html and at http://modeldb.yale.edu/266507.

Appendix A Neural and Synaptic Models

The biophysical neuron model used in the simulations described in this paper is a leaky integrate and fire model with a fixed threshold. Synaptic conductance is transmitted at a decaying-exponential rate from the pre to post-synaptic neurons (Gerstner et al., 2014). The simulations are coded in PyNN (Davison et al., 2007) to specify the topology, flow of inputs, and recording. The neurons themselves are simulated using NEST (Gewaltig and Diesmann, 2007).

The model used in this paper follows Fourcaud-Trocmé et al. (2003) (but also see (Richardson and Gerstner, 2003)). The activation is the current voltage V_M. Equation 1 describes the change in voltage; V_M is the membrane potential and C_M is the membrane capacity. The four currents are the leak current, the currents from excitatory and inhibitory synapses, and the input current (from some external source). The variable currents are governed by equations 2, 3 and 4. In equations 2 and 3 E_{Ex}^{rev} and E_{In}^{rev} are the reversal potentials; excitation (and inhibition) changes slow as the voltage approaches these reversal potentials. In equation 4, V_{rest} is the resting potential of the neuron, and τ_M is the leak constant.

$$\frac{dV_M}{dt} = \frac{(-I_{Leak} - I_{Ex}^{syn} - I_{In}^{syn} + I_{Ext})}{C_M} \tag{1}$$

$$I_{Exc}^{syn} = G_{Ex} \times (V_M - E_{Ex}^{rev}) \tag{2}$$

$$I_{Inh}^{syn} = G_{In} \times (V_M - E_{In}^{rev}) \tag{3}$$

$$I_{Leak} = \frac{C_M(V_M - V_{rest})}{\tau_M} \tag{4}$$

$$G_{Ex}(t) = k_{Ex} \times t \times e^{-\frac{t}{\tau_{Ex}^{syn}}} \tag{5}$$

$$G_{In}(t) = k_{In} \times t \times e^{-\frac{t}{\tau_{In}^{syn}}} \tag{6}$$

In equations 5 and 6, G_{Ex} and G_{In} are the conductance in mS/cm^2 to scale the post-synaptic potential amplitudes used in equation 2, and 3. t is the time step. The constant k_{Ex} and k_{In} are chosen so that $G_{Ex}(\tau_{Ex}^{syn}) = 1$ and $G_{In}(\tau_{In}^{syn}) = 1$. The τ_{Ex}^{syn} and the τ_{In}^{syn} are the decay rate of excitatory and inhibitory synaptic current.

When the voltage reaches the threshold, there is a spike and the voltage is reset. No current is transferred during the refractory period $\tau_{refract}$. In these simulations $v_{thresh} = -48.0$mV, $\tau_{refract} = 2.0$ ms. The time step t is 1ms. $C_M = 1.0$nF, $v_{reset} = -70.0$mV, $v_{rest} = -65.0$mV, $E_{Ex}^{rev} = 0.0$mV, $E_{In}^{rev} = -70$mV, $\tau_{Ex}^{syn} =$ 5.0ms, $\tau_{In}^{syn} = 5.0$ms and $\tau_M = 20.0$ms. The particular parameters v_{thresh},$\tau_{refract}$, and t, were selected as the authors have used them for prior simulations; they are the parameters used in the binary CAs for the semantic portion of the associative memory[2]. The remaining parameters are default values.

[2] Note that the model expressed in equation 1 about the exponential integrate-and-fire neuron is a particular case of the AdEx model by removing the adaptation current (w) (Gerstner and Brette, 2009).

Appendix B The Winner Take All (WTA) Model: Stationary Bumps

This paper describes work on a linear WTA model, instead of a planar model or hyper-planar model. A line of neurons is connected with local excitatory synapses, and a surround of inhibitory synapses. (If the first and last neurons are also adjacent, it is a ring attractor.) There are many synaptic matrices that lead to persistent behaviour once the initial neurons are stimulated (see section B.2). What is needed is sufficient local synaptic excitatory strength to allow the neurons within the winning group to fire persistently. This needs to be balanced with sufficient inhibitory synaptic strength to prevent spread beyond the initial group.

One example is a network with each neuron having excitatory synapses to the nearest two neurons on both sides (distance(d)$d <= 2$) and inhibitory synapses to the next nearest 4 neurons on both sides ($3 <= d <= 6$). This is called a 2-4 bump attractor in this paper. The model is implemented in a neural network with 100 spiking neurons. This is also called a stationary bump (Laing et al., 2001), or a bump attractor. In the associative memory, the bump attractor approximates the continuous value activation of a temperature scale in which every neuron represents a single degree from 0° to 99° C.

This paper explores a range of parameters instead of developing a mathematical model. Mathematical models of bump attractors usually take advantage of simplifications, such as firing rates (e.g. (Carroll et al., 2014)), or they take advantage of statistical mechanics (e.g. (Tkačik et al., 2015)). Unfortunately, statistical mechanical analysis of bump attractors usually makes use of a large number of probabilistic neurons. Here we are using a relatively small number of deterministic neurons.

B.1 Persistent Bumps, Divergent Bumps, and Multiple Bumps

The behaviour of the 2-4 bump attractor varies based on the weight of the excitatory and inhibitory synapses. Table 2 describes the behaviour of this network as the inhibitory and excitatory synaptic weights vary in steps of .01 μS (microsiemens). In the Table, D means firing spreads, diverging so all of the neurons fire persistently, and a number reflects how many bumps are firing persistently after 1000ms. One means that all of the neurons that are firing persistently are adjacent to each other; they are one bump.

In the top part of the table, three adjacent neurons are forced to spike, representing input from the environment. Three neurons are chosen as it is typically thought that several neurons are needed to cause another to spike (Churchland and Sejnowski, 1999), so this is the minimum input needed to (ignite) start a CA persistently firing. After the initial stimulation, each simulation is run for 1000ms. The value in the cells of the Table is the number of neurons that are firing at the end of the simulation.

In the first row of Table 2, there is insufficient excitatory synaptic strength to enable the neurons to continue to fire each other. In the excitatory 0.06 row, there is enough spread of activation to enable the neurons to fire persistently. In the first columns (e.g. cell .06 -.02) there is insufficient inhibition to prevent the neural activation spreading, and all of the neurons fire. On the right however (e.g. cell .08 -.09), a small reverberating population fires throughout the simulation.

This shows that persistent stable firing occurs after three adjacent neurons are spiked. This persistent spiking is similar to the behaviour of a binary CA. However, when a larger range of neurons are initially spiked (as may be the case in an associative memory), there is further interesting behaviour.

The bottom part of the Table 2 refers to 75 neurons being initially spiked. As in the top portion, some excitatory inhibitory weight pairs lead to no persistence, and some lead to all of the neurons firing.

In the bottom half of the Table, those with one in the cell have more than 75 neurons persistently firing, but not all 100. Most of the table cells show two bumps firing. These two bumps are always on the edge. The edge neurons inhibit the interior neurons, as do the interior neurons themselves so that they do not fire a second time. The edge neurons have fewer incoming inhibitory connections, so they can persistently fire. After the initial burst, the interior neurons do not fire, the two bumps do not influence each other, and they fire persistently as if they were ignited by two individual sets of inputs. It is also interesting to note that several of these cells have more than two bumps; this Table shows four, six and seven bumps. Again these bumps are all quite thin, with approximately seven persistently firing neurons, and they have a relatively small number of non-firing neurons in between them.

				3 Input				
	-0.03	-0.04	-0.05	-0.06	-0.07	-0.08	-0.09	-0.10
0.05	0	0	0	0	0	0	0	0
0.06	D	1	1	1	1	1	1	1
0.07	D	D	1	1	1	1	1	1
0.08	D	D	D	D	1	1	1	1
0.09	D	D	D	D	D	1	1	1
0.10	D	D	D	D	D	1	1	1
				75 Input				
0.05	1	2	2	2	2	0	0	0
0.06	D	1	4	2	2	2	2	2
0.07	D	D	1	7	2	2	2	2
0.08	D	D	D	D	7	2	2	2
0.09	D	D	D	D	D	6	2	2
0.10	D	D	D	D	D	D	D	4

Table 2 Table of persistently firing bumps of neurons for a 2-4 stable bump topology. The top refers to input of three adjacent neurons, and the bottom to an input of 75 contiguous neurons. The value in the cell represents the number of persistent bumps; D (divergent) refers to all of the neurons persistently firing.

Note that it is possible to have local excitation with inhibition to all other neurons. When there are a small number of inputs, this performs largely the same as, for instance, a 2-4 stationary bump. However, with a larger number of inputs, say 75, the inhibition from the initial firing prevents all the neurons from firing. In Table 2, the 75 input cells would all be 0. Let us call this topology with inhibition to all other neurons a 2-n topology. It is possible to set the synaptic weights so that a 2-n topology leads to persistent firing from 75 inputs, but the width of the bump would be very large. For instance, a 2-n topology with 0.08 excitation and

0.005 inhibition has a persistent bump 68 neurons wide when 75 neurons initially spiked, and 56 neurons wide when 3 neurons are initially spiked.

B.2 Exploring Input Variation

The stationary bump attractor's spiking behaviour differs as the excitatory and inhibitory weights vary. They also vary as the number of initial neurons spiked changes.

The number of input sources were varied from 1 to 40, while the synaptic weights were also varied, from 0.05 to 0.10 (step 0.01). All the simulations run in NEST (Gewaltig and Diesmann, 2007)[3].

	-0.05	**-0.06**	**-0.07**	**-0.08**	**-0.09**	**-0.1**
0.05	4	4	4	4*	4*	/
0.06	2	2	2	2	2	2
0.07	2	2	2	2	2	2
0.08	2(D)	2(D)	2	2	2	2
0.09	2(D)	2(D)	2(D)	2	2	2
0.1	1(D)	1(D)	1(D)	1	1	1

Table 3 Minimal number of spikes sources to ignite the 2-4 bump-attractor network. The "/" means absence of any spikes, the $*$ means firing without persistence and (D) means diverging behaviour.

Table 3 shows the number of inputs needed to support ignition of the bump attractor network. The ignition of the bump attractor network depends on both the number of inputs and the specific weight combination. In most cases, only two spike sources are needed to ignite the network, but there are cases where four sources are needed or only one. It is important to note that there are some cases where there is ignition, without continued persistence (represented by a * in the Table); for example 0.05 - 0.08 only fires until 88 ms. There are also cases in which the bump ignites but with divergence, those in the lower left triangle of Table 3 indicated with (D). Like Table 2, Table 3 confirms that there needs to be sufficient excitatory strength to support persistence, and sufficient excitatory strength to overcome inhibitory strength. Similarly, too much excitatory strength in relation to inhibition leads to divergence. When there is less excitatory strength, more input sources can support ignition, as shown when excitation is 0.05, and fewer sources are needed for more excitation, as shown when excitation is 0.1.

The splitting behaviours are presented in Tables 4 and 5. Table 4 shows the smallest number of sources needed for two bumps to emerge. It also shows when there is diverging behaviour indicated with the letter D. Again, Table 4 shows that even as the number of sources increases, the stable, non divergent bumps are focused around a balance between excitation and inhibition. When there is less inhibition (toward the left of the Table), more sources are required for splitting to emerge.

[3] The same exploration procedure has been done with SpiNNaker (Furber et al., 2013) neuromorphic hardware (see (Vergani and Huyck, 2020))

	-0.05	-0.06	-0.07	-0.08	-0.09	-0.1
0.05	13	12	11	/	/	/
0.06	14	13	13	12	11	11
0.07	17	14	13	13	12	12
0.08	D	D	17	14	13	12
0.09	D	D	D	15	14	13
0.1	D	D	D	D	D	14

Table 4 Number of inputs that determine the splitting behaviour into 2 bumps of the 2-4 bump attractor network. *D* means divergent behaviour, the "/" means no persistent firing.

Table 5 shows a special subset of weight combinations in which three and four bumps emerge. They are the subset belonging to the central diagonal in Table 5 where excitatory weights are one step (0.01) greater than inhibitory, except the case with 0.1-0.1 has equal weights.

E-I Weights	3S	4S
0.06-0.05	24	36
0.07-0.06	24	37
0.08-0.07	26	40
0.09-0.08	23	37
0.10-0.10	25	38

Table 5 Combination of excitatory E and inhibitory I weights that determine the splitting behaviour with 3 bumps or 4 bumps of the 2-4 bump attractor network.

This section has shown parameter sweeps over a 2-4 bump attractor, with a particular leaky integrate and fire model, a particular synaptic transmission model, and particular neural parameters. The same parameters are unlikely to work with different models or even different neural parameters. While the same weight and number of input parameters are unlikely to work on different variants, the general idea of balancing excitatory and inhibitory weight will. Moreover, it will generalise to different sized bump attractors, for instance a bump attractor with excitatory synapses at a distance of three, and inhibitory synapses eight beyond, a 3-8 attractor.

The emergence of multiple bumps is caused by similar, and, in the case of this section, equal inputs to a large contiguous range of the bump attractor. Bumps may influence each other.

Appendix C Replacing Two Bumps with One

The bump attractor is both an attractor and a repeller. Adjacent neurons support each making a local attractor, but adjacent bumps repel each other. A rastergram is insufficient to show this, but figure 2 (in section **??**) includes the voltage and rastergram of a single bump, and of two bumps that are near to each other.

With a single bump, note that the neurons in the centre of the bump spike more frequently than those on the outside (figure 2A). However, the rastergram does not show the effect on non-spiking neurons. The voltage diagrams (figures

2B and 2D) show the voltage of the neurons at each step. The neurons adjacent to the bump are excited but not to the firing threshold, and those beyond are inhibited. Those in the bump are, of course, excited; their voltage changes from the reset voltage $-70mV$ toward the firing threshold $-48mV$. Those further away are unaffected and remain at the base level of activation, $-65mV$ as are all neurons before the initial input.

The inhibition in the neurons between the bumps is quite high even when compared with the inhibition on the neurons on the outside of the bump. So, what is needed is a burst of excitation into those central neurons with inhibition to the outside, but only in the case where the two bumps are quite close together.

The inhibition can be overcome and the two bumps replaced by one using an extra set of neurons that only fires when two nearby bumps are firing. There are the same number of neurons in the new *Overcome* population as in the original stable bump population, and they are aligned. The neurons in the *Overcome* population get excitatory input from the corresponding neurons below in a small window (10 neurons on either side with a weight of 0.0013). Where the two windows overlap, the neurons fire. If only one bump is firing in the subnet, or if the bumps are quite distant, no neurons in the *Overcome* subnet fire.

When the *Overcome* neurons fire they send excitation to the *Bump* neurons directly below, and inhibition more distantly. In this case, the *Overcome* neurons excite in a window of three about themselves (with a weight of 0.0013), and inhibit the next eight (with a weight of -0.4). This is quite similar to a bump topology with three excitatory and eight inhibitory synapses, a 3-8 bump attractor.

The result of this is that the interior neurons fire, and then remain persistently firing. Figure 5 shows this. The bottom figures (C and D) are the raster and voltage plots of the bump attractor. Note that the initial firing behaviour leads to a split into two bumps in figure 5C. Voltage slowly builds in the appropriate neurons in the *Overcome* subnetwork 5B, causing a single set of neural firing that shifts the behaviour in the stable bump attractor.

C.1 Merging Bumps

When the temperature range is quite large, the two initial bumps do not influence each other. For example, in Table 2 when there are 75 inputs (from neurons 25 to 99), the bumps are from neurons 25 to 31 and 92 to 98.

As there is no theoretical limit to the range of a continuous value for a specific semantic category, the 3-8 bump attractor approach, or indeed any X-Y bump attractor including all to all inhibition, will not work; larger bump attractors can account for larger ranges of inputs, but the bumps themselves also get larger. This can however be solved by an additional inhibitory topology increasing with distance, activated by the two firing bumps.

As in the overcome case, there is an extra population of neurons of the same size, which is called the *Merge* subnet. It is important that these neurons do not fire unless there are two bumps firing, so each of the *Bump* neurons has a small excitatory connection to each of the merge neurons (weight 0.0006). There are also direct one to one excitatory synapses (weight 0.01) from the *Bump* subnet to the *Merge* subnet so that the neurons associated with the bumps fire; so, these subnets are also aligned.

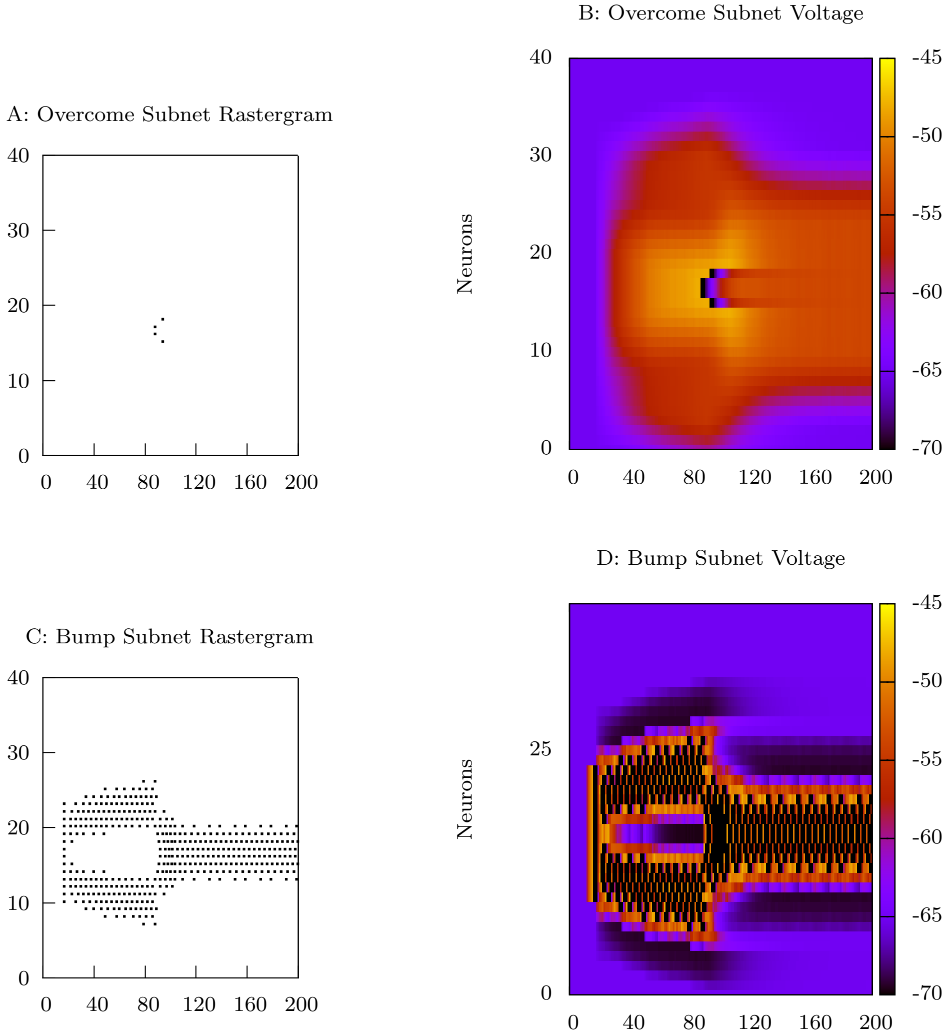

Fig. 5 Results from the Overcome topology initially spiking neurons 10 through 23. A rastergram of the inhibitory subnetwork for the Overcome simulation(A), voltage signature of the inhibitory subnetwork for Overcome (B), rastergram of the bump subnetwork for Overcome (C), and voltage signature of the bump subnetwork for Overcome(D). This shows an initial input that splits into two bumps. Neurons in the inhibtory subnetwork fire, causing the bumps to merge.

Inhibitory synapses from the *Merge* neurons to the *Bump* neurons are distance biased with more distant neurons being more inhibited (weight = $(1.02^{\delta}) * 0.04$ where δ is distance). Thus the outside neuron of the neurons in the opposing bump is more inhibited than the inside neuron, and this, metaphorically, pushes the bumps together. There are no synapses to nearby neurons (range of 15).

There is a difficulty that as the distance increases, the inhibitory strength becomes too large and the bumps stop each other. The inhibition described in this paper, increases exponentially with distance (with a base of 1.02). It is sufficient to cope with 75 neuron input, but will not work with much larger differences.

Figure 6 shows the behaviour of this system. Figure 6A shows the spiking behaviour of the inhibitory *Merge* subnet. Its firing is sparser than the bump attractor 6C, but follows it. Note that once the inside edge of the bump stops being inhibited, because it is near enough to have no more inhibitory synapses, it quickly moves because the inside is not inhibited, but the outside is. Also note that the voltage in the *Merge* subnet is high throughout the run after initial stimulation in figure 6B. This is below firing threshold, but is due to the all to all synapses.

This then moves the two bumps toward each other. It does not however push them together due to them repelling each other. Fortunately, this extra merge topology is compatible with the overcome topology. An example of the full bump merging topology, combining a bump, a *Merge* subnet, and an *Overcome* subnet, is shown in section 4 in figure 3.

Note that this topology has some stability across varying bumps. Changing the bump weights from .08 excitatory and -.08 inhibitory, to .09 -.09 does require a change, but it is only the inhibitory weight in the *Merge* subnet from -0.4 to -0.32. Similarly, changing the bump weights to .07 -.07 requires the inhibitory *Merge* weight to -0.29. With lower weights (e.g. -0.4) , pairs of bumps from the larger ranges are extinguished, and with higher weights, the larger ranges do not merge sufficiently close. It is not a linear relation (both greater and smaller bump weights require less inhibition) as the excitatory inhibitory balance is not an entirely linear relation.

Appendix D Associative Memory

Appendix D.1 introduces an associative memory with an input bump attractor. The goal of the system is to produce the correct associated output from two inputs. A refined system using bump attractors for each beverage, with the additional merge and overcome subnets (appendix D.2) shows the problem of broad input in context, and shows a solution to this problem. Finally, Hebbian learning is used to learn the associations (appendix D.3), which leads to more efficient runtime retrieval with psychologically plausible retrieval times.

D.1 The Hot Coffee Network: An Associative Memory Example

The basic simulation that has driven the development of this paper is an associative memory of beverages with a semantic value for temperature, and an underlying temperature in celsius.

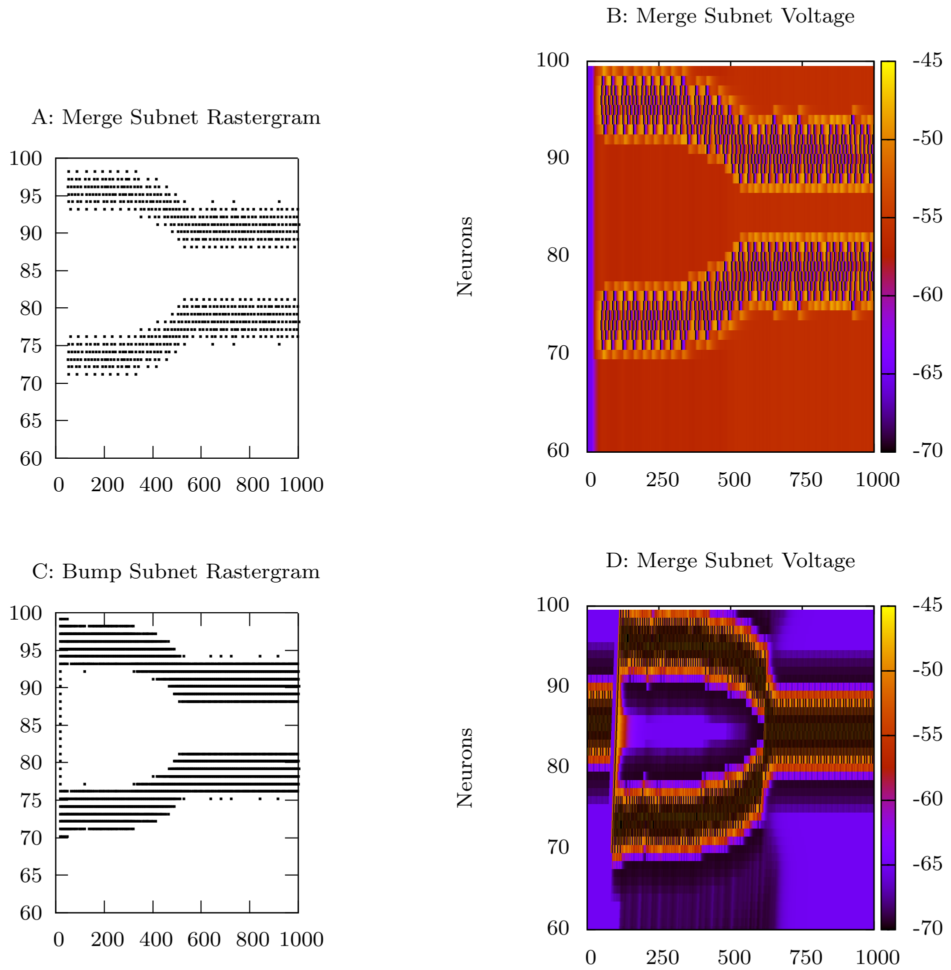

Fig. 6 Results from the Merging topology initially spiking neurons 70 through 99. A rastergram of spiking behaviour for the inhibitory subnetwork (A), voltage signature of the inhibitory subnetwork (B), rastergram of spiking behaviour for the bump attractor subnetwork (C), and voltage signature of the bump attractor (D). This shows how the networks collectively move the bumps towards each other.

The gross topology of the network is shown in figure 4 in section 5. The input temperature from the environment is represented by a stable bump subnetwork of 100 neurons. The topology used in the remainder of the paper is a 2-4 topology with excitatory connections from a given neuron to the two adjacent neurons on either side, then four inhibitory neurons beyond. The weights are .08 excitatory and -.08 inhibitory. This represents the temperature values between 0 and 99.

There are also neurons that represent the temperature of the individual beverages. These subnetworks are used in the associations. The 10 neurons in the

Inhibition subnetwork take input from the beverage temperature neurons, and in return inhibit them; each beverage neuron synapses to each inhibitory neuron with a weight of 0.04, and in return receives an inhibitory synapse with a weight of 0.016. The neurons in the *Inhibition* subnet synapse to two adjoining neurons with a weight of .03 providing extra inhibition via increased firing. The inhibition prevents the spread of activation from one beverage to another. For instance, if the input temperature is 75-77°, and *Coke* is queried, the associated neurons in the *Coke Temperature* subnet will spike, igniting *Hot*; now that the *Input Temperature* and *Hot* are firing, they ignite *Coffee Temperature*, which would then ignite *Coffee* without the *Inhibition* subnet.

The arrows represent several synapses from a given set of neurons to another set to support associations. Each *Input Temperature* neuron excites its associated *Coffee Temperature* neuron weight of 0.01. As these neurons are meant to associate CAs, in themselves, they are insufficient to cause neurons in another CA to fire; each beverage temperature stimulates the eight excitatory beverage semantic neurons and the appropriate temperature semantic neurons with a weight of 0.025.

Each of the eight semantic excitatory beverage neurons excites each associated temperature beverage neuron with a weight of 0.0015, and the semantic temperature neurons excite their beverage temperature neurons with a weight of 0.001.

Synapses inside simple CAs are not represented in the figure. Being binary CAs, the neurons in the semantic CAs have internal synapses as does the bump attractor *Input Temperature* subnet.

The individual CAs can be ignited by external stimulation. When this happens they all persist, and do not cause any other CAs to fire. There are the five semantic binary CAs and the *Input Temperature* bump CAs.

There are two types of atypical case to mention. The first case is the activation of a low input temperature ($< 10°$) along with *Cold* or a high input temperature ($\geq 70°$) with *Hot*. In this case, both beverages are activated, which is of course the correct result.

The second is the activation of two semantic CAs, for example, *Hot* and *Coffee*. Here the full range of beverage temperature neurons fire, but they fire at a low rate. In this case, the *Coffee Temperature* neurons from 70-99 fire. While this topology, using weakly connected beverage temperature neurons, in a sense solves the problem, a better result might be the prototypical beverage temperature neurons firing persistently. For instance, for *Hot* and *Coffee*, the neurons around 85° might fire. This would lead to actual retrieval when the semantic neurons are turned off, as the bump attractor neurons would continue to fire.

An obvious modification is to replace the beverage temperature subnets, which have no internal connections, with bump attractors. However, a straight forward switch elicits a flaw. If for example *Hot Coffee* is stimulated, all 30 neurons (70-99°) fire, but the stable state that the *Coffee Temperature* subnetwork settles into is two bumps of neurons, one from 70° and one to 99°, firing persistently with those in between silent. This is similar to the two bumps of Tables 2 and 4.

D.2 Full Hot Coffee Network

While it seems reasonable for the full range of temperature neurons to fire due to direct semantic information, it is somewhat inconsistent with the firing behaviour from direct temperature input. Moreover, the firing in the beverage temperature subnet is not persistent on its own. A better result would be to have prototypical or average temperatures fire persistently. So, in the case of *Hot Coffee* semantic input, the neurons that represent coffee at 81° to 88° might fire persistently.

Making the beverage temperature subnets bump attractors enables persistent activity. However, with large input ranges, there are two bumps of activity at the ends of the range. For example, in the *Hot Coffee* case, neurons 70° to 74° and 95° to 99° fire persistently. This is the problem that was solved in section 4.

The associative memory topology is modified to include the extra *Merge* and *Overcome* subnets. The full hot coffee network is still described by figure 4, but the *Coffee Temperature* and *Coke Temperature* boxes are now three populations each, a stable bump attractor, a parallel *Merge* subnet, and a parallel *Overcome* subnet; so the beverage temperatures are now each represented by 300 neurons. Note that the dynamics of these three subnets in isolation differ from those in the full topology because of the *Inhibition* subnetwork in the full topology; the neurons in the bump attractors fire at a lower rate. Extra excitatory (weight 0.002 from 0.0013) and inhibitory strength (weight -0.5 from -0.4) are needed in the *Overcome* subnetwork as the stable bump is firing at a lower rate due to the effect of the *Inhibition* subnetwork.

Now, as in section D.1, the basic one semantic feature and one temperature input work properly, and quickly. For instance, temperature input of 85° to 87° and the semantic value *Coffee*, turns on the semantic value *Hot* within 300 ms. The ambiguous inputs (the semantic value *Cold* with a temperature below 10°, and the semantic value *Hot* with a temperature above 70°) turn on both semantic beverages.

The additional merging topology now causes the double semantic input queries to generate the appropriate temperature outputs. Each of the six pairs (e.g. *Cold Coke*, or *Warm Coffee*) produce a persistently firing output. Unlike the simple topology it is self sustaining when the two semantic CAs stop firing. Table 1 shows the association temperature range and the output results. It also shows the time to converge, noting how wider ranges take significantly longer to converge. However, small ranges (like *Cold Coke*) converge almost immediately even when they would break into two bumps (*Warm Coke*).

D.3 Associative Memory with Learning

The mammalian brain constantly learns, and one of the real benefits of spiking neural systems is that they can be reasonable models of at least parts of mammalian brains, so it is important that these systems learn. Using Hebbian learning, the subnetwork is presented triplets of inputs (semantic beverage, semantic temperature, actual temperature), and the firing behaviour is stored. Compensatory learning rules are Hebbian, and based on a Oja's rule (Oja, 1982), where the synaptic weight W_{XY} from neuron X to neuron Y is represented by equation 7; it is the likelihood that the two neurons cofire when the presynaptic neuron X fires, modified by a constant.

$$W_{XY} = C * \frac{coFire(X,Y)}{fire(X)} \tag{7}$$

So, if Y fires 80% of the time along with X and $C = 1$, the weight is 0.8. If $C = 0.5$, $W_{XY} = 0.4$. Cofiring requires a time window and for this work a 10ms window has been used so that Y can fire in the same step as X or 10ms later to cofire. In a compensatory rule, C is calculated based on the total synaptic weight entering a neuron(a pre-compensatory rule), or leaving a neuron (post-compensatory). Thus weights are distributed based on the cofiring behaviour of all attached neurons.

From this behaviour, new synaptic weights are calculated using a compensatory Hebbian learning rule. Results are shown from learning based on presentation of a single triple for each of the semantic pairings with one temperature (section D.3.1). The full range of semantic pairings with all of their temperatures (section D.3.2) is presented in one test and random inputs distributed about a centroid in another.

D.3.1 One of Each Category

The first training mechanism presents the system with one of each category triplet. The triplets were: *Hot Coffee* 80° to 82°; *Warm Coffee* 40° to 42°; *Cold Coffee* 15° to 17°; *Hot Coke* 40° to 42°; *Warm Coke* 15° to 17°; and *Cold Coke* 4° to 6 °.

These were each presented for 300 ms., and firing was stopped via inhibition between epochs. Firing behaviour is recorded and a compensatory Hebbian learning rule is applied. Weights from the semantic temperatures to the beverage bump attractors are calculated using a post-compensatory rule, and weights from the beverage bump attractors to the semantic temperatures are calculated using a pre-compensatory rule (see section 2.4). The total target synaptic weight was 0.03 and 0.1 respectively.

After training, all 18 pairs (e.g. *Cold* and *Coke*, *Coke* and 4° to 6°, and *Cold* and 4° to 6°) are presented to the system as a test. In each case, the correct remaining third member was retrieved, and no spurious elements were retrieved. Unlike the tests from appendices D.1 and D.2, a low temperature and *Cold* only retrieves one beverage, because the beverages are associated with different temperature inputs. Indeed, many temperatures not presented during training are not associated at all.

D.3.2 The Full Range of Inputs

Next, the full range of temperatures is presented with their appropriate semantic temperatures and for both beverages. This included roughly 200 runs of the system, with intervening inhibition to stop firing between epochs.

The Hebbian learning rule is applied leading to a system with correct results. As in the simulations in section D.3.1, all the base double inputs lead to correct results. The cold and hot temperatures with semantic *Cold* and *Hot* lead to both beverages being retrieved.

Presenting *Coke* or *Coffee* with each temperature ignites the correct association, though *Warm* is incorrectly activated along with *Cold* for 10° and 25° respectively. Presentation of the semantic temperatures with each input temperature turns on the correct association. When the temperature is low (e.g. 3°) or high (e.g. 93°), both *Coke* and *Coffee* come on when *Cold* or *Hot* is presented respectively. *Warm* splits perfectly at 25°. When input temperatures are presented where no

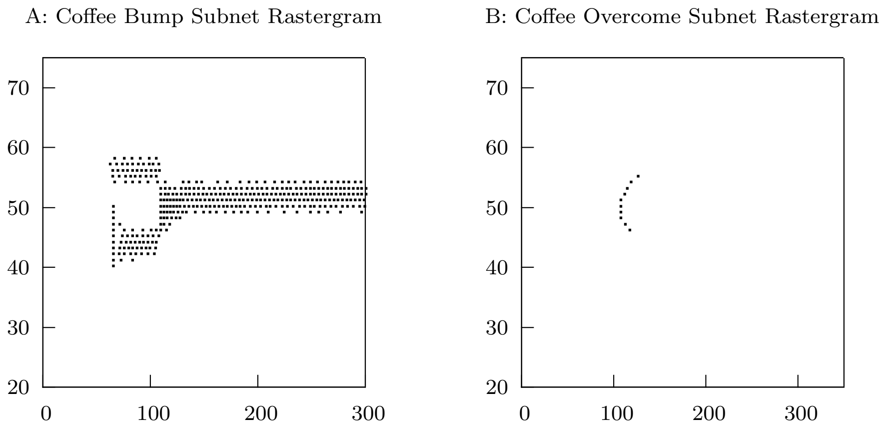

Fig. 7 Rastergram of the Coffee Temperature Bump Attractor subnetwork (A) and the Coffee Temperature Overcome subnetwork (B) from weights learned from presentation of temperatures with gaussian distribution about a mean. Initial activation of a broad range of neurons from the semantic pair leads to two streams that collapse into one when *Overcome* neurons spike.

beverage is associated with a semantic temperature, generally no beverage comes on. However, in the edge cases (e.g. *Cold* with 28°) the beverage temperature range expands (e.g. *Coffee*).

As a final test, random temperatures are presented during training. The same number of input epochs is generated as the full range of temperatures, with the same number for each semantic pair. The temperature presented is a Gaussian distribution about a particular temperature. After learning, when the semantic pair is presented, a bump attractor near the Gaussian centroid is retrieved. This is a prototypicality effect.

A particularly interesting case is the presentation of the semantic temperature and beverage (see the last column of table 1 for learned results). The interesting part is that all of these results converge by 300 ms. This is within psychologically realistic times. Decisions are made in this order of time, so this associative task could be used in a cognitive model.

There is sufficient activation to cause the full range of neurons to fire, and bumps form at both ends. However, there is sufficient firing in the bumps to cause the associated *Overcome* subnetwork to fire. This is shown in figure 7. In this presentation of *Warm* and *Coffee*, the full temperature range can be seen to fire, 7A. This causes the coffee *Overcome* subnetwork to fire, 7B.

The final figure (8) shows a run of the system for 1900 ms. without the 300 neurons for the coke temperature, which do not fire. The system is initially presented with *Cold Coffee* at 25 ms.; firing begins rapidly in the appropriate range of the *Coffee Temperature* subnet, causing firing in the *Coffee Merge* and *Coffee Overcome* subnets, which quickly leads to a single stream. The *Inhibit* subnet fires in response. At 300 ms. the neurons are inhibited. Each subsequent input starts after 300 ms., so the next is at 325 ms. The second presentation is *Coffee* and 45-48°,

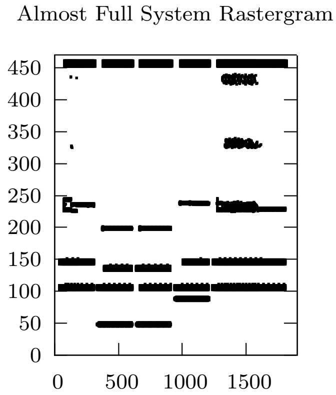

Fig. 8 Rastergram of the network run on five inputs. The 300 coke neurons are missing, though don't fire in this simulation. The first hundred neurons are *Input Temperature*; the next 50 are the semantics with *Coffee* as 100-109, *Coke* 110-119, *Cold* 120-129, *Warm* 130-139, and *Hot* 140-149; *Coffee Temperature* is 150-249, *Coffee Overcome* is 250-349, *Coffee Merge* is 350-449, and *Inhibit* is 450-459. There are five sets of input each followed by inhibition that shuts down the system. The inputs start every 300 ms. starting at 25 ms. The first presentations are: *Cold Coffee*; *Coffee* 45-48°; *Warm* 45-48°; *Coffee* 85-88°; and finally *Hot Coffee*.

which answers *Warm*; the third presentation is *Coffee* and 45-48°, which answers *Warm*. *Coffee* and 85-88° is the fourth presentation, which answers Hot. Finally *Hot* and *Coffee* are presented at 1425 ms. This is resolved by 1500 ms. at 84°, but it runs on here until 1800 ms. because it does not particularly look resolved as there is ongoing firing in the inhibitory coffee subnets. It does however resolve, exhibiting a fully functional hot coffee network. Thus the topology produces the correct associated output from any two inputs.

Funding

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 720270 (the Human Brain Project) and the UK EPSRC grants EP/D059720 and EP/P00542X/1.

References

Binder J, Conant L, Humphries C, Fernandino L, Simons S, Aguilar M, Desai R (2016) Toward a brain-based componential semantic representation. Cognitive neuropsychology 33:3-4:130–174

Brachman R, Schmolze J (1989) An overview of the KL-ONE knowledge representation system. In: Readings in artificial intelligence and databases, Morgan Kaufmann, pp 207–230

Buzsaki G (2010) Neural syntax: cell assemblies, synapsembles, and readers. Neuron 68:3:362–385

Carroll S, Josi K, Kilpatrick Z (2014) Encoding certainty in bump attractors. Journal of computational neuroscience 37:1:29–48

Chen Y (2017) Mechanisms of winner-take-all and group selection in neuronal spiking networks. Frontiers in computational neuroscience 11:20

Chrysanthidis N, Fiebig F, Lansner A (2019) Introducing double bouquet cells into a modular cortical associative memory model. Journal of computational neuroscience 47:2-3:223–230

Churchland P, Sejnowski T (1999) The Computational Brain. MIT Press

Collins A, Loftus E (1975) A spreading-activation theory of semantic processing. Psychological review 82:6:407–428

Craig A, Chen K, Bandy D, Reiman E (2000) Thermosensory activation of insular cortex. Nature neuroscience 3:2:184–190

Davison A, Yger P, Kremkow J, Perrinet L, Muller E (2007) PyNN: towards a universal neural simulator API in python. BMC neuroscience 8(S2):P2

Edelman G (1987) Neural Darwinism: The theory of neuronal group selection. Basic books

Fourcaud-Trocmé N, Hansel D, Van Vreeswijk C, Brunel N (2003) How spike generation mechanisms determine the neuronal response to fluctuating inputs. Journal of neuroscience 23(37):11628–11640

Furber S, Lester D, Plana L, Garside J, Painkras E, Temple S, Brown A (2013) Overview of the SpiNNaker system architecture. IEEE transactions on computers 62(12):2454–2467

Gerstner W, Brette R (2009) Adaptive exponential integrate-and-fire model. Scholarpedia 4(6):8427, DOI 10.4249/scholarpedia.8427, revision #90944

Gerstner W, Kistler W, Naud R, Paninski L (2014) Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press

Gewaltig M, Diesmann M (2007) Nest (neural simulation tool). Scholarpedia 2(4):1430

Hahn G, Ponce-Alvarez A, Monier C, Benvenuti G, Kumar A, Chavane F, Deco G, Frégnac Y (2017) Spontaneous cortical activity is transiently poised close to criticality. PLoS computational biology 13(5):e1005543

Handjaras G, Ricciardi E, Leo A, Lenci A, Cecchetti L, Cosottini M, Marotta G, Pietrini P (2016) How concepts are encoded in the human brain: a modality independent, category-based cortical organization of semantic knowledge. Neuroimage 135:232–242

Harris K (2005) Neural signatures of cell assembly organization. Nature reviews neuroscience 6:399–407

Hebb D (1949) The Organization of Behavior: A Neuropsychological Theory. J. Wiley & Sons

Huyck C, Ji Y (2018) Two simple neurocognitive associative memory models. In: Sixteenth International Conference on Cognitive Modelling, pp 31–36

Huyck C, Mitchell I (2014) Post and pre-compensatory Hebbian learning for categorisation. Cognitive neurodynamics 8:4:299–311

Huyck C, Mitchell I (2018) CABots and other neural agents. Frontiers in neurorobotics 12:79

Huyck C, Passmore P (2013) A review of cell assemblies. Biological cybernetics 107:3:263–288

Kim S, Rouault H, Druckmann S, Jayaraman V (2017) Ring attractor dynamics in the drosophila central brain. Science 356:6340:849–853

Kohonen T (1982) Self-organized formation of topologically correct feature maps. Biological cybernetics 43(1):59–69

Kreiser R, Cartiglia M, JMartel, Conradt J, Sandamirskaya Y (2018) A neuromorphic approach to path integration: a head-direction spiking neural network with vision-driven reset. In: IEEE International Symposium on Circuits and Systems (ISCAS), pp 1–5

Laing C, R, Chow, Carson C (2001) Stationary bumps in networks of spiking neurons. Neural computation 13(7):1473–1494

Lansner A (2009) Associative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. Trends in neuroscience 32:3:178–186

Laurens J, Angelaki D (2018) The brain compass: a perspective on how self-motion updates the head direction cell attractor. Neuron 97:2:275–289

Martin A (2007) The representation of object concepts in the brain. Annual review of pyschology 58:25–45

Meiss J (2007) Dynamical systems. Scholarpedia 2(2):1629, DOI 10.4249/scholarpedia.1629, revision #137210

Nadh K, Huyck C (2010) A pong playing agent modelled with massively overlapping cell assemblies. Neurocomputing 73:2928–2934

Oja E (1982) Simplified neuron model as a principal component analyzer. Journal of mathematical biology 15:3:267–273

Pena R, Zaks M, Roque A (2018) Dynamics of spontaneous activity in random networks with multiple neuron subtypes and synaptic noise. Journal of computational neuroscience 45(1):1–28

Peterson G, Barney H (1952) Control methods used in a study of the vowels. Journal of the acoustical society of america 24:2:175–184

Pulvermuller F (1999) Words in the brain's language. Behavioral and brain sciences 22:2:253–279

Quillian M (1967) Word concepts: a theory of simulation of some basic semantic capabilities. Behavioral science 12:410–30

Redish A, Elga A, , Touretzky D (1996) A coupled attractor model of the rodent head direction system. Network: computation in neural systems 7(4):671–685

Richardson M, Gerstner W (2003) Conductance versus current-based integrate-and-fire neurons: is there qualitatively new behaviour? Lausanne lecture

Seeholzer A, Deger M, Gerstner W (2019) Stability of working memory in continuous attractor networks under the control of short-term plasticity. PLoS computational biology 15:4:e1006928

Singer W, Engel A, Kreiter A, Munk M, Neuenschwander S, Roelfsema P (1997) Neuronal assemblies: necessity, signature and detectability. Trends in cognitive sciences 1:7:252–261

Somers D, Nelson S, Sur M (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. Journal of neuroscience 15(8):5448–5465

Stringer S, Rolls E, Trappenberg T, Araujo ID (2002) Self-organizing continuous attractor networks and path integration: two-dimensional models of place cells. Network: computation in neural systems 13:4:429–446

Tetzlaff C, Dasgupta S, Kulvicius T, Florentin W (2015) The use of Hebbian cell assemblies for nonlinear computation. Scientific reports 5:12866

Tkačik G, Mora T, Marre O, Amodei D, Palmer S, Berry M, Bialek W (2015) Thermodynamics and signatures of criticality in a network of neurons. Proceedings of the National Academy of Sciences 112(37):11508–11513

Vergani A, Huyck C (2020) Critical limits in a bump attractor network of spiking neurons. arXiv arXiv:2003.13365

Willshaw D, Buneman O, Longuet-Higgins H (1969) Non-holographic associative memory. Nature 222:960–962

Wilson H, Cowan J (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13(2):55–80

Wimmer K, Nykamp D, Constantinidis C, Compte A (2014) Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. Nature neuroscience 17(3):431

Yoon K, Buice M, Barry C, Hayman R, Burgess N, Fiete I (2013) Specific evidence of low-dimensional continuous attractor dynamics in grid cells. Nature neuroscience 16(8):1077