Associative latching dynamics vs. syntax

Publisher:
Springer

Published
DOI:10.1007/978-90-481-9695-1_17

Associative Latching Dynamics vs. Syntax

Eleonora Russo[1], Sahar Pirmoradian[1], Alessandro Treves[1]

[1] SISSA Cognitive Neuroscience, via Beirut 2-4,
34151 Trieste, Italy
{russo, pirmorad, ale}@sissa.com

Abstract. We model the cortical dynamics underlying a free association between two memories. Computationally, this process may be realized as the spontaneous retrieval of a second memory after the recall of the first one by an external cue, what we call a *latching transition*. As a global cortical model, we study an associative memory Potts network with adaptive threshold, showing latching transitions. With many correlated stored patterns this unstable dynamics can proceed indefinitely, producing a sequence of spontaneously retrieved patterns. This paper describes the informational properties of latching sequences expressed by the Potts network, and compares them with those of the sentences comprising the corpus of a simple artificial language we are developing, BLISS. Potts network dynamics, unlike BLISS sentences, appear to have the memory properties of a second-order Markov chain.

Keywords: memory, attractor dynamics, information, artificial language.

1 Introduction

Cortical networks have been thought to retrieve memories associatively [1], both at the local and global level [2]. The simple Hopfield neural network model [3] has stimulated the study of content-addressed, auto-associative retrieval in terms of attractor dynamics. Once a memory has been retrieved by an external cue, however, if the corresponding attractor state is made unstable, it may serve itself as an internal cue for a second memory. Based on this simple observation, many authors have explored recurrent networks that model processes of the free association between two memories. These proposals differ in the ingredients introduced in order to destabilize the first attractor state, so as to produce a spontaneous sequential retrieval of several stored memories.

In 1986 Sompolinsky and Kanter [4] proposed a simple network that, with a specific set of connection weights, could retrieve an equally specific sequence of memories. In 1987 Tsuda [5, 6] proposed a model with two coupled networks, a stable and an unstable one, interacting with each other and generating oscillating retrieval. Hermann et al [7], in 1993, obtained "episodic" and "semantic" transitions among memories with a dynamic threshold influencing active neurons.

Similarly to the work of Hermann, we study here a network where transitions among patterns are due to an adaptive threshold, but with a different kind of units.

2 Potts Model

The network, described in detail in [8], is a recurrent network with Potts units. Instead of a single neuron each unit represents a local cortical network, and can take S different uncorrelated active states (plus one inactive state). The active states represent memories at the local network level, realized as local attractor states through a process the model does not describe, as it focuses on the global cortical level. Within each active state the activation rate is taken to be threshold-linear. We study the global attractor dynamics, after storing on the tensor connections among Potts units p global activity patterns. Each unit receives C connections.

To have spontaneous transitions among attractors we introduce a fatigue model: both an overall and a state-specific adaptive threshold which, with a characteristic time constant, tracks the mean activation value of each Potts unit in that state. After retrieving a global pattern the system is in a quasi-stable attractor, as the threshold continues to increase and to weaken each active state, until the unit changes state or becomes inactive. The free-energy landscape of configurations is then dynamically modified, leading the system to abruptly jump to a new attractor, often nearby.

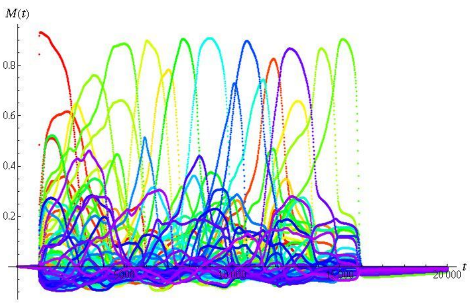
In order to have structured latching transitions, however, the introduction of a dynamic threshold is not enough: correlations among patterns are also needed. In [9] we have studied the detailed dynamics of transitions between two correlated patterns. In this simplified condition we have identified three types of latching transition, each characterized by a range of correlations: quasi-random transitions between weakly correlated attractors, history-dependent transitions between attractors with stronger correlations, oscillatory transitions between pairs of closely overlapping attractors.

3 Extended Potts Dynamics

With a generative algorithm, as explained in [8], we produce a set of correlated patterns and we store them into the network. In this more natural and complex condition, transitions among all pairs of patterns may occur, and they still seem to cluster in the three types above. Starting with an external cue, that induces the retrieval of a pattern, a series of transitions follows, as in Fig.1, until activity dies out or we stop the simulation. Its duration increases with the correlation among patterns and with p, the total number of patterns. Latching dynamics can be quite disorderly, but we can extract the sequence of patterns that, at each time t, have the highest overlap with the activity of the Potts network. We can then study the properties of such discrete sequences, neglecting the originally continuous nature of the dynamics.

4 BLISS Sentences

In a separate project, we have designed an artificial Basic Language Incorporating Syntax and Semantics, BLISS, in order to test the language acquisition capability of the Potts and of other networks. BLISS is intended to be of intermediate complexity, and in its current provisional form it includes a stochastic regular grammar with 30 production rules but no semantics yet. The associated probabilities (e.g. for transitive vs. intransitive verbs) are fine-tuned to the statistics of the one-million-word Wall Street Journal (WSJ) corpus. The 170 terminal symbols belong to different lexical categories such as verb, noun, adjective, determiner, demonstrative, …, whose relative frequencies are also tuned to the WSJ corpus. With a Perl code we have generated 50,000 BLISS sentences, of length between 5 and 12 words and with maximum 3 levels of embedding.



→ Ahriman takes Zarathustra

→ New lions build the house

→ The cow dances

→ The knife assumes that the worm sees AhuraMazda

→ Yasmin knows that Magi don't believe that the ox doesn't need the dog

→ Great black stones give the stone to Yasmin

Fig. 1. (left) Example of Potts dynamics: different colors show the overlap of the network with different patterns; (right) Examples of sentences generated by the artificial language BLISS.

5 Entropy and Information

We may now focus on the memory properties of both processes, Potts latching dynamics and BLISS sentences, described as discrete sequences of 'symbols' picked from among p=80 memory patterns and p=170 words, respectively.

In the Potts case, we have run $p \times p$ latching sequences stopping each sequence after 30 transitions, of which we have analysed the last 20. Averaging across all transitions, we count the occurrence of all possible single patterns, consecutive pairs and triplets of patterns, and compute e.g. the entropy $H(x_n)$ and the information among the symbols (= patterns) in position n and the ones in position $n+i$, $I(x_n\ ;\ x_{n+i})$. We proceed similarly with BLISS, using sentences with at least 7 symbols (= words).

Results show that both systems span an entropy lower than $\log_2 p$, naturally for BLISS and in the Potts case due to finite size effects. Both have favourite pairs and triplets: while many among all the possible combinations never occur, others are frequently present. Comparing the frequency of each pair with that of its inverse, BLISS is seen to be almost fully asymmetric (either one never occurs), while Potts dynamics are substantially symmetric, as expected. The most frequent Potts pairs are comprised of strongly correlated patterns. For BLISS, in Fig. 2 (right) each color denotes the entropy at a specific position and the information which that position conveys to the next positions. Over half of the total variability at each position is seen to be independent of the words in previous positions, while the rest is determined by a long history of all preceding words, pointing at extended syntactic dependences.

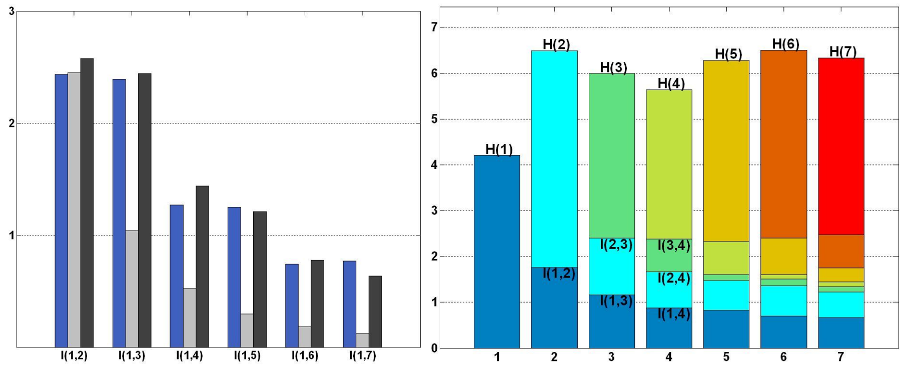


Fig. 2. (left) Information flow in the Potts latching sequence (blue), in a first-order Markov chain (gray) and in a second-order Markov chain (black); (right) Information flow in BLISS sentences, each color denoting the decaying influence of the word chosen at each position. Note the different *y*-scales (both in bits).

In order to better characterize information propagation along the Potts latching sequence, in Fig. 2 (left) we have compared our sequences with Markov chains of first and second order. To equalize the noise present in a finite sample, we have also produced *p x p* Markov chains of 20 steps, with the two transition matrices of the first and second order chains extracted from the original Potts latching sequences. We can see in Fig. 2 that, whereas in a first order Markov process information decays faster than in BLISS, but also monotonically, the surprising trend of the latching information is similar to the one of a second-order Markov chain.

References

1. Marr, D.: A theory for cerebral neocortex. Proc. Roy. Soc. London B 176 (1971) 161–234
2. Braitenberg, V., Schüz, A.: Anatomy of the cortex. Springer-Verlag, Berlin (1991)
3. Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Aca. Sci. USA 79 (1982) 2554-2558
4. Sompolinsky, H., Kanter, I.: Temporal association in asymmetric neural networks. Phys. Rev. Lett. 57 (1986) 2861–2864
5. Tsuda, I., Körner, E., Shimizu, H.: Memory dynamics in asynchronous neural networks. Progr. Theor. Phys. 78 (1987) 51-71
6. Tsuda, I.: Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. Behav. Brain Sci. 24 (2001) 793-847
7. Herrmann, M., Ruppin, E., Usher, M.: A neural model of the dynamic activation of memory Biol. Cybern. 68 (1993) 455-463
8. Treves, A.: Frontal latching networks: a possible neural basis for infinite recursion. Cogn. Neuropsy. 21 (2005) 276-291
9. Russo, E., Namboodiri, V.M.K., Treves, A., Kropff, E.: Free association transitions in models of cortical latching dynamics. New J. Phys. 10 (2008) 015008