A stochastic short-term memory using a pRAM neuron and its potential applications.

Guido Bugmann*, John G. Taylor

Centre for Neural Networks, King's College London
London WC2R 2LS, United Kingdom
email: gbugmann@soc.plym.ac.uk

22 June 1993

Contents

---

*Now at: Neurodynamics Research Group, School of Computing, University of Plymouth, Plymouth PL4 8AA, United Kingdom, http://www.tech.plym.ac.uk/soc/staff/guidbugm/bugmann1.html

Abstract

We consider here the type of short-term memory which manifests itself in the form of sustained activity of a neuron or a cluster of neurons. We assume that, due to the stochastic release of neurotransmitter at synapses, the sustained activity has a limited lifetime. The parameters controlling the average lifetime are analysed mathematically using the pRAM model of a neuron receiving its own spikes as feedback. Beside giving short-term memory properties, self-feedback loops increase the gain of feature detecting neurons. The biological counterpart of the pRAM based model of short-term memory is then discussed.

As an example of how to exploit the combined properties of feature detection and short-term memory, we propose a multilayer neural network where feature detecting nodes maintain their activity until detectors of more complex features in higher layers have used their information. The lower layer activity is then shut down via inhibitory feedback. Such a network is self timed and can realize parallel and serial pattern recognition tasks. Short-term memory properties can also be used to realize a neuronal AND-function, a sequence detector and a neuronal timer. That timer can generate the gated signals assumed by Grossberg and Schmajuk [1989] in their model of classical conditioning.

1   Introduction

In physiological terms, short-term memory (STM) or "working memory" refers to the temporary storage of information needed to perform a later mental or physical operation. STM is for instance tested in delayed matching experiments where the presentation of a cue determines the response to a stimulus given at a later time. In such experiments, the delays between cue and stimulus can be as long as 32 sec [Fuster et al., 1981]. At this time scale, short-term memory is electrophysiologically correlated with the sustained firing of specific neurons. These so-called short-term memory neurons have been observed in various parts of the cortex, in the motor cortex [Smyrnis et al., 1992], in the supplementary motor area [Mushiake et al., 1990], in the prefrontal cortex [Funahashi et al, 1989] and in areas IT and V1 in the visual system [Fuster, 1990].

At shorter time scales, up to 2-4 sec, information may also be stored in the form of a temporary modification of membrane properties. Such a mechanism has been assumed in models of the articulatory loop [Burgess and Hitch, 1992] and of temporal discrimination in classical conditioning [Grossberg and Schmajuk, 1989]. It is possible that both storage mechanisms are used by the same neural system.

The sustained activity of STM-neurons is assumed to arise within a group of mutually excitatory neurons. Such an activity, called "reverberation", was assumed to be able to last for a very long time. Caianiello [1961] even proposed that the role of sleep was to extinguish the excess of reverberations. Reverberating circuits, although not using spiking neurons, are extensively described in monographs by Rashevsky [1960] and Grossberg [1982].

An important aspect of the information transmission between biological neurons is the probabilistic nature of the release of neurotransmitter at synapses [Taylor, 1972; Pun et al., 1986]. As a consequence, the duration of reverberations is limited and governed by statistical laws. In section 2 we analyse the statistical properties of the simplest possible reverberating circuit, consisting of a pRAM model of a neuron [Gorse

and Taylor, 1990] with a self-feedback excitatory connection. Properties such as decay-time, rise-time, frequency gain and methods to control the reverberation are presented.

We then relate the properties of the pRAM based model to the characteristics of biological short-term memory neurons in section 3. We discuss the possible biological "self-feedback" mechanisms, the control mechanism of STM neurons and the stochasticity of synapses.

In section 4 we describe a number of applications of STM neurons. We describe how STM allows the realization of a neuronal AND-function. We propose a self-timed neural network capable of parallel and sequential pattern recognition. Some properties of this network are discussed, in particular concerning the conditions under which short-term memory neurons may be observed in biological systems and the role of inhibitory feedback. Also, we present a modification giving that network the capability of performing sequence recognition. Finally, we address the question of reliable time estimation by using stochastic short-term memories and propose a neuronal timing circuit. Such a neuronal timer has the properties required to explain temporal characteristics of classical conditioning [Grossberg and Schmajuk, 1989].
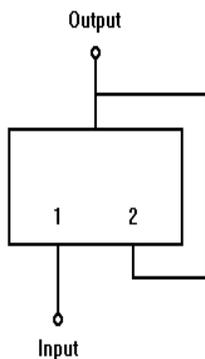
Figure 1: A 2-pRAM neuron with self-feedback. Each spike produced by the pRAM has a probability α_{01} of initiating a new output spike. Each input spike, on input 1, has a probability α_{10} of initiating an output spike.

2   The pRAM as a stochastic short-term memory

2.1   The pRAM with self-feedback

The pRAM [Gorse and Taylor, 1990] is a device operating in discrete time steps and produces spikes according to the following rule. At each time step, it observes the spike activity at its N inputs, which is represented by a bit string μ containing N bits which are either "0" or "1". This bit string is the address of a memory where a firing probability α_μ is stored. By comparing this value α_μ with the number produced by a random number generator, an output spike (a "1") is then produced with the probability α_μ.

Let us consider a pRAM with 2 inputs (figure 1). The input 1 is connected to an input line. The input 2 receives a feedback connection from the output. This device behaves as a stochastic short-term memory when its parameters are set in such a way that a previous output spike can generate a new output spike and cause a self-sustained activity.

As an example of such parameters, let us first set the firing probability α_{00} corresponding to the input configuration [0,0] (no spike present at any of the two inputs) to zero, such that no spontaneous activity be permitted. Let us then set the firing probabilities corresponding to the input configurations [1,0] and [1,1] to $\alpha_{10} = 1$ and $\alpha_{11} = 1$. With these values, each input spike on input 1 will produce an output spike. Now let us set the firing probabilities corresponding to the input configurations [0,1] to $\alpha_{01} = 0.98$. In this way, each output spike will cause a spike with 98% probability when the input 1 is silent (configuration [0,1]) and will cause a spike with 100% probability when the input 1 is also active ($\alpha_{11} = 1$). The probabilistic action of feedback spikes will eventually cause the sustained activity to stop. This gives the stochastic characteristic to this device.

2.2 Decay time

When the input 1 is stopped, the pRAM will continue to fire with, at each time step, a probability of 2% to stop. The duration t of the output spike train can therefore only be characterized by a probability. The probability $Prob(n\Delta t)$ of the activity in the loop to survive n simulation time steps decreases following the law:

$$
\begin{aligned}
Prob(t = n\Delta t) &= (\alpha_{01})^n = e^{-\frac{t}{\tau}} \\
\tau &= \frac{\Delta t}{ln(1/\alpha_{01})}
\end{aligned}
\tag{1}
$$

This decreasing probability is illustrated by the peri-stimulus histogram shown on figure 2A where $\alpha_{01} = 0.98$. Assuming a time-step of $\Delta t = 2ms$ this value of α_{01} would correspond to a time constant $\tau = 99ms$. Due to the averaging over many runs, the probability to observe a spike after a given number of time-steps after the end of the stimulation appear as an exponentially decaying curve. This is similar to the decay of the potential in a leaky capacitor. However, while the potential can be measured at any time, the survival probability is not an observable. The activity of the pRAM during each single run is maximal during a certain number of time steps and then stops. The smooth decay is the effect of averaging over many runs. Examples of individual spike trains are shown in figure 2B.

2.3 Rise time

When an input spike does not systematically cause an output spike, in the case of $\alpha_{10} < 1$, then there is also a rise time in the onset firing probability. The probability $Prob(k)$ for an input spike to induce an output spike (and start the feedback loop) increases with the number k of input spikes as:

$$
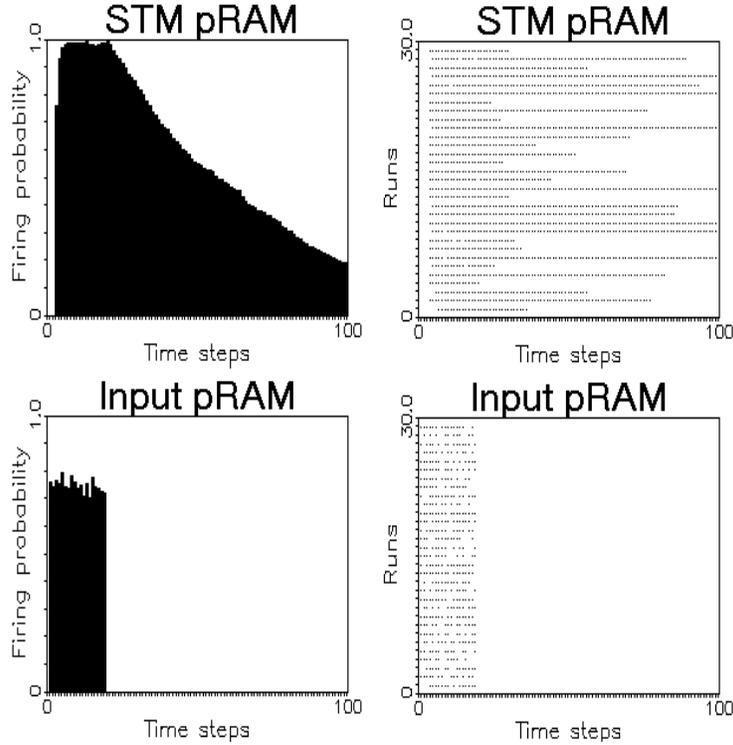Prob(k) = 1 - (1 - \alpha_{10})^k
\tag{2}
$$

4

Figure 2: **A)** Peristimulus histogram showing the time-dependent probability of the pRAM with self-feedback to be firing. The curve is an average over 500 runs of 100 time-steps. **B)** Spike raster representation of a series of individual runs of a pRAM with self-feedback. The bottom figures show the input spike trains. The parameters were $\alpha_{10} = 0.5$ and $\alpha_{01} = 0.98$. There is a transmission delay between the input-pRAM producing random spike trains and the STM-pRAM. Other information in the text.

An example can be seen in figure 2 where $\alpha_{10} = 0.5$. Again, the smooth average curve hides an all-or-none behaviour. There are cases where even a large number of successive input spikes will not initiate the loop. The smaller the value of α_{10}, the longer the input spike train must be to have a good chance to start a loop.

2.4   Stopping the loop

Let us assume that, after the end of the actual input, the persistent activity of a loop is required only during a given time, for instance until it has been processed by a later neuronal stage. We will also assume that the loop life-time is much longer than the range of practically required retention times. How can the activity in the loop be stopped at will ?

This can done by using a pRAM with a third input. The memory contents α_{ij1} must then all be set to zero. In this way, as soon as a spike arrives at the input 3, no firing is possible and the loop is disrupted. This third input plays the role of an

inhibitory input.

This reseting technique is different from the one used in the recurrent short-term memory proposed by [Zipser, 1991]. In his system, a zero input has to be memorized in order to set the output to zero, and this state lasts only during the coherence time of the system.

2.5 Gain

A feedback loop does not necessary imply an all-or-none behaviour of the system. Let us consider an input made of a random spike train with a frequency x. In the general case the output frequency y is given by [Bugmann and Taylor, 1993a]:

$$y(x) = \frac{\alpha_{00} + (\alpha_{10} - \alpha_{00}) * x}{1 - (\alpha_{01} - \alpha_{00}) - (\alpha_{00} - \alpha_{10} - \alpha_{01} + \alpha_{11}) * x} \tag{3}$$

With the restrictions we have made on the values of the memory contents in section 2.1 this becomes:

$$y(x) = \frac{x}{1 - \alpha_{01} - (1 - \alpha_{01}) * x} \tag{4}$$

which is the equation of a non-linear amplifier with gain $1/(1 - \alpha_{01})$ near to $x = 0$. The amplifying properties of self-feedback loops are well known in analog electronics. The difference is that here it is not a continuous current which is fedback but a train of discrete spikes. This system is not a current amplifier but a spike frequency amplifier.

3 Self-sustained activity in biological systems

3.1 Underlying connectivity

The simplest neural network allowing a self-sustained firing is formed of an axon making synapses in the dendritic tree of its own neuron. Such synapses, called "autapses", have been described by Van der Loos and Glaser [1972] in some pyramidal cells in the striate cortex and have been studied electrophysiologically in the olfactory bulb by Nicoll and Jahr [1982]. Autapses have also been found in the motor cortex [Meyer, 1987]. There are typically less than 10 autapses per neuron [Van der Loos and Glaser, 1972], which represents a very small fraction of the total number of synapses. Most cells in layer 4 of the visual cortex receive less than this number of synapse from each LGN cell to which they are connected [Freund et al., 1985]. Therefore, for excitatory postsynaptic potentials (EPSP) of comparable size, the influence of self-feedback may be rather moderate, comparable to that of any other input. Self-feedback can increase the gain of a neuron (see section 2.5), a function needed to preserve the output frequency level of neurons with non-linear multiplicative transfer functions [Bugmann, 1991b] (see also section 4.1). It is therefore possible that more autapses will be found in areas like MST, where some neurons have highly non-linear transfer functions [Tanaka et al., 1989], possibly realizing AND-functions [Verri et al., 1992].

As the number of feedforward synapses and autapses is small, the main excitatory input to a cortical neuron comes from neigbouring neurons. Such "lateral"" connections

have been infered in the striate cortex from secondary EPSP's following brief electrical stimulations of the LGN [Douglas and Martin, 1990]. While the primary EPSP's are monosynaptic, the secondary EPSP's are trisynaptic, indicating the involvement of an excitatory relay neuron in the cortex [Ferster, 1986]. The average latency between primary and secondary EPSPs is approximately 5ms.

Some aspects of the sustained activity of cortical neurons can not be explained by simple autapses and require a network of a certain complexity. For instance, in visual area IT, the frequency of the sustained firing is decreasing with time [Fuster, 1990]. In the prefrontal cortex, it is sometimes increasing [Funahashi et al., 1989]. It is not clear how this could be reproduced with a simple feedback loop which, as we have seen in section 2.2, would cause an all-or-none behaviour.

Not much is known about the details of the network causing the secondary EPSPs. In visual area IT, cross-correlation analysis shows that these arise mainly from neighbouring neurons selective to different stimuli and with non-shared inputs [Gochin et al., 1991]. Such a scheme would suggest an autoassociative function underlying the sustained activity, similar to that postulated for the CA3 network in the hippocampus [Rolls, 1990, Treves and Rolls, 1992], rather than a simple self-excitation.

However, self-feedback loops have been used in simulations as a simplified representation of multiple intracortical reexcitation and have allowed a good reproduction of experimental data. For instance the temporal response pattern of the intracellular potential of pyramidal neurons in V1 [Douglas and Martin, 1991] or the temporal low pass filtering properties of pyramidal neurons in V1 [Maex and Orban, 1992] seem to support the existence in the primary visual cortex of circuits functionally equivalent to self-excitation. However, these are not necessarily those underlying STM.

In some cases, the sustained firing rate does not exceed $25Hz$ [Funahashi et al., 1989]. If we take that as an indication of an average loop-time of 40ms we are faced with the problem that the local circuit, with typically a loop-time of 5ms, may not be the best candidate for explaining sustained firing. Another possibility would be that the activity is maintained using inhibitory interneurons and afterhyperpolarization rebounds [Steriade, 1984; Kirillov et. al, 1993]. Such circuits can exhibit quite long loop-times. This problem of low sustained firing frequencies has attracted the attention of some investigators analysing attractor neural networks [Amit and Tsodyks, 1991; Bressloff and Taylor, 1993]. These models however use deterministic synapses and propagate firing rates between neurons rather than spikes. Clearly, the details of the STM circuitry need further investigations.

3.2   Control of short-term memory

The sustained firing of STM neurons lasts only for the time during which specific information is to be retained. There must therefore exist mechanisms for enabling and reseting the loop at precise times. To our knowledge, there have been no experiments investigating this point. An involvement of the basal ganglia is possible [Chevalier and Deniau, 1990].

With regards to enabling, we can imagine various mechanisms. For instance, a barrage of small EPSPs produced at distant synapses, for instance those in layer I receiving feedback projections from other areas, could reduce the threshold and

facilitate firing in response to lateral or feed-forward input spikes. However, such an input would also reduce the apparent leak resistance of the membrane [Bernander et al., 1991] and reduce the temporal integration of other inputs. It is possible that there would be a facilitatory effect only for the detection of coincidence at proximal synapses. An alternative would be to remove the feedback inhibition observed in all neurons, via inhibition of interneurons, and allow the self-sustained firing.

With regards to stopping the loop, inhibition at the level of the axon hillock during a certain time is a possibility. The inhibition must last a certain time because, with multiple synapses, possibly located at different electrotonic distances on the dendritic tree, or multiple neurons in a cluster, the last EPSP fed back can occur a certain time after the end of firing. Simulations have shown that such an inhibition can increase the membrane leakage by a factor of 2 [Douglas and Martin, 1990]. This does not much reduce the amplitude of arriving EPSPs but it reduces the background potential on which arriving EPSPs are building up. This can be an effective way to increases the actual threshold for firing. Another, more speculative, way would be to reduce the synaptic transmission probability via an adequate neurotransmitter or neuromodulator. However, it is not clear which neuromodulator could play such a role.

It may also be possible that cortico-thalamo-cortical feedback effects may be involved, with additional control activity arising from the nucleus reticularis thalami [Taylor, 1992a, 1992b; Alavi and Taylor, 1992, 1993]. Such effects go beyond the scope of the level of the present discussion, however.

In all cases, there is evidence that external control signals are acting on STM neurons. One hypothesis is that the local inhibitory network is controlled by projections from other areas, so that sustained firing can be enabled depending on the computational needs. Such a hypothesis underlies our model for parallel and serial pattern recognition in section 4.2.

Another hypothesis to be made when considering control of STM neurons, is that the lifetime of the sustained activity must be considerably larger than the time scale at which control is performed. By using a pool of STM neurons one can actually garantee a minimum retention time during which there is almost never a spontaneous end of the activity (see section 4.4).

3.3   Stochastic synapses

In real neurons, a spike arriving at a synapse will not systematically initiate the release of neurotransmitter and produce an EPSP in the postsynaptic membrane. For instance, in neurons of the spinal cord, the release probabilities have been estimated to be between $p = 8\%$ and $p = 44\%$ depending on the fraction of synapses which are considered active [Pun et al., 1986].

Often, the spike arriving from a source neuron is distributed to a certain number k of synapses on the target neuron. This number is sometimes surprisingly small (see section 3.1). This duplication of synapses leads to a higher effective excitation probability P_e. Let us assume that an input spike must produce a number n of near simultaneous EPSP's at a pool of k synapses in order to initiate an output spike, then P_e depends on k and n in the following way:

$$P_e = 1 \quad - \quad \sum_{j=0}^{n-1} C_j^k (1-p)^{k-j} p^j \qquad (5)$$

$$C_j^k \quad = \quad \frac{k!}{j!(k-j)!}$$

If these k synapses are autapse, then P_e is equivalent to α_{01} and the decay time constant can be calculated as in section 2.2. If they are part of an excitatory circuit involving relay-neurons, the synaptic release probability p must be multiplied by the firing probability of the presynaptic relay neuron.

An important feature of such circuits is the probabilistic nature of the self-sustained activity and its associated decay-time. In delayed matching or delayed response experiments, it has been reported that failures to perform the task correctly are associated with a premature termination of the sustained activity [Funahashi et al., 1989]. Therefore, with short-term memory based on stochastic synapses, one may expect the performance to decay as the demanded retention time increases. This is actually observed [Fuster et al., 1981] although the decay curve is not a simple exponential and may indicate the uses of a set of STM neurons rather than a single one (see section 4.4).

Theoretically, the lifetime of the sustained activity could also be controlled by altering the neurotransmitter release probability of individual synapses. It is therefore tempting to explore if this release probability is part of the properties of neurons subject to learning mechanisms. An indication that this property is at least variable is the observation that silent synapses can become active [Redman, 1990].

4   Applications of short-term memory neurons

4.1   A neuronal AND-function

A neuron can operate as a coincidence detector, firing only if all its inputs receive a spike during a small time-window. When information is coded as single spikes, this operation is identical to the AND-operation of logical gates. However, when information is coded by the frequency of random spike trains, the AND-function can formally not be realized [Bugmann, 1991a]. For instance, when all inputs are active, say with a frequency f_1 coding for the logical level 1, the output frequency is given by the frequency f_c with which the coincidences occur. As this frequency is always smaller than f_1 [Bugmann, 1991b], the output cannot reach the logical level 1 which would be expected from an AND-function. Actually, the larger the number of inputs, the smaller the output frequency and, for more than 20 inputs, the output frequency does not exceed the background noise level [Bugmann, 1991b]. There are indications that the brain uses neurons as coincidence detectors [Bugmann and Taylor, 1993b] but the associated dramatic frequency drop is generally not observed. How is this possible ?

A solution to that problem is to assume that coincidence detecting neurons are provided with a self-feedback loop. Such neurons have then a gain which can multiply the number of spikes produced by each coincidence in order to compensate for the frequency drop (see section 2.5).
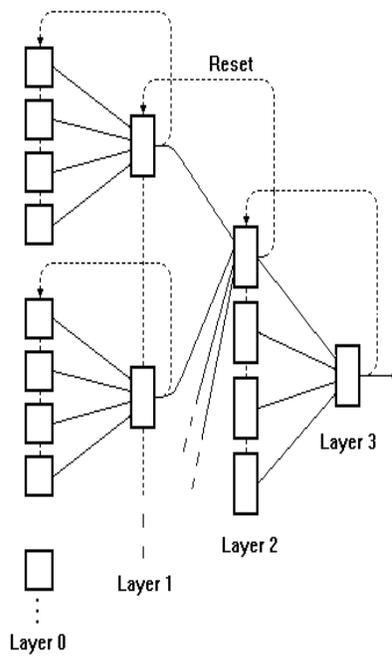
Figure 3: A neural network with STM nodes and inhibitory recurent inhibition. Neurons operate as coincidence detectors, producing an output spike only when all m input neurons are firing and produce m synchronous spikes (in the figure, $m = 4$). As soon as the first output spike is produced, neurons enter a state of sustained firing (the self-feedback circuit is not shown here). They also send an inhibitory reset signal to interrupt the sustained firing of their m input neurons.

4.2   Parallel and serial pattern recognition

Some tasks performed by the brain are parallel, others are serial. For, instance, recognizing the character "A" in an image is considered to be a parallel process performed by the visual system. But if an "A" has the size of a building or is observed through a microscope with such a magnification that only parts of it can be seen at any time, then an observer will have to explore visually each part of the character before being able to identify it. This exploration process is typically a serial one. How can such different processes lead to the same final state, the recognition of the character "A" ? Are two convergent information processing systems needed ?

A solution to the serial/parallel problem is to assume that pattern recognition takes place in a multilayer network of feature detecting neurons with STM properties (figure 3). The operation of that network is exemplified with the problem of recognizing a character "A". Let us assume that there is a "grand-mother" cell for the character A (Figure 4). This "A-cell" receives it inputs from feature detecting cells and operates as an AND-gate. Only if all inputs representing the characteristics parts of the A are active, then the A-cell will fire.
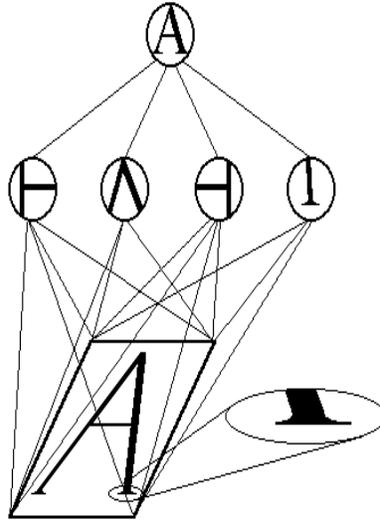
Figure 4: Application of a neural network using short-term memory neurons to the parallel and serial recognition of the character "A". In the parallel case, the entire character can be seen at one time. In the serial case, only a small part of the character can be seen at any time and recognition requires exploration of the image.

In the parallel processing case, the whole letter can be observed simultaneously. All feature detectors are activated at the same time and the recognition takes place immediately, as in a standard MLP network.

In the serial case, the features can also be detected (let us assume that their detection is size invariant) but they cannot be presented simultaneously to the A-cell. However, by using feature detecting cells with short-term memory, these can remain active as long as the exploration lasts. In that way, at the end of the exploration, all feature detecting cells are active at the same time and the A-cell can eventually fire (for simplicity we have omitted to discuss the coding of the spatial relation between individual features which is in one case obtained from retinotopic mapping and, in the other, from exploratory movements).

When the pattern has been recognized, the persistent activity of the feature-detecting cells is no longer necessary and can even be disturbing for future recognition tasks. A good solution seems then to shut down these cells by sending an inhibitory feedback.

Such a system has interesting properties:
- It is self-timed because its computation or recognition speed is entirely determined by the degree of synchronization of the inputs. It has been used as a model of the latencies in the visual system [Bugmann and Taylor, 1993b]. Interestingly, the computation speed is not affected by the time delay in the feedback inhibition from the next layer. The image information just flows through the network as fast as loops can be turned on. The feedback delay however, reduces the frequency at which new images can be processed because of the delayed resetting. As all inputs occurring before the feedback

arrives are treated as simultaneous, the feedback delay also determines the temporal fusion threshold, the minimum time which must separate two images perceived as distinct ($\approx$ 30ms in the visual system [Pöppel, 1987]).

- What happens when the character is unknown and no recognition can take place ? In that case, there is no resetting signal and the sustained activity of the feature detecting neurons persists. Such a persistence could be used as a signal that there is a problem in recognizing the character and be used to take some decision. For instance, focusing more attention, searching for other cues or giving up. We may note that it takes approximately 0.5sec before a neuronal activity becomes conscious [Libet, 1989]. Persistence could also be used by a "novelty detector" à la Grossberg [Grossberg and Schmajuk, 1989].

- In the case of shared feature detectors, the postulated feedback signal is probably too simple. For instance, after exploring the vertical bar of a "T" the network may recognize an "I" and shut down the related feature detectors, preventing a subsequent recognition of the "T". A more sophisticated feedback structure is certainly necessary. Actually, nearly half of the cortical neurons (layer 5 and 6) are involved in computing the feedback signals. This shows the importance and complexity of the task.

- If for some reason an output neuron in layer n is prevented from firing, input neurons in the previous layer $n - 1$ will exhibit sustained firing, inhibit neurons in the layer $n - 2$ and cause sustained firing in layer $n - 3$. Let us transpose this to observations in the delayed response experiments. We will assume that the output neuron codes for the delayed response and is inhibited during the delay. Then, a sustained STM behaviour would be observed among those neurons feeding into the response neuron. By designing experiments involving the retention of information at various levels n' of information processing, one may be able to generate STM neurons at the levels $n' - 1$ and $n' - 3$. Our model would predict the observation of STM neurons in every second layer.

How can we train a target cell to become an A-cell ? We may note that the existence of self-feedback connections and interlayer inhibition does not affect the selectivity of a given neuron to the input conditions. Therefore, for the feedforward path, any of the existing methods to train feedforward multilayer networks can be used.

The new aspect of the problem is to train the interlayer feedback projections and to determine which input should be inhibited by the activity of a target cell. For such a task, principles derived from classical conditioning could possibly be used. In classical conditioning, a node previously active (CS) is linked to a node subsequently activated by US (see section 4.4). For inhibitory connections, one could invert such a rule. A node becoming active will be connected in an inhibitory way to all nodes which are already active. For biological realism, since inhibitory synapses are not known to be adaptive, inhibitory interneurons should be assumed, with excitatory synapses on them being trained. It would also comply with the more complex feedback function mentionned earlier. These are general ideas and future work is clearly needed.
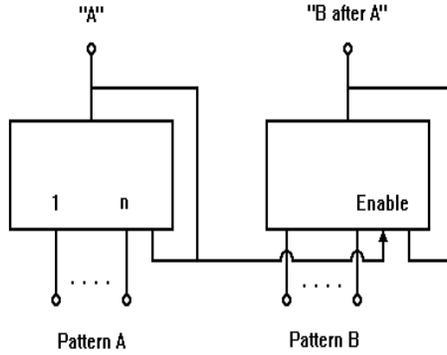
Figure 5: Sequence detector using STM neurons. The neuron on the right requires the presentation of the pattern "B" *and* an input from the "A"-neuron to start firing.

4.3 Sequence recognition

With the previously described network (figure 3), any sequence containing the characters "D", "O" and "G" would be recognized by a neuron tuned to "G O D". Therefore, for certain recognition tasks, order-selective STM neurons are needed. We present here a simple circuit with such a property.

The STM circuit in figure 5 stores the occurence of a sequence "AB". It uses a first STM neuron recognizing "A". Its output is used as an enabling signal for a second STM neuron recognizing "B". Therefore, the neuron "B" can be activated only if the neuron "A" is already firing. Thus, the simple addition of adequate lateral connections in the network of the previous section makes it selective to the order of presentation of the elements of a composite pattern.

4.4 A neuronal timer

In classical conditioning experiments the animal not only learns that a conditioning stimulus (CS), for instance the sound of a bell, will be followed by an unconditional stimulus (US), for instance food. It learns also *when* the US will occur after the CS. Therefore, there is a neuronal mechanism which allows encoding and retrieving time delays between events. A model has been proposed, using a set of spectral timing functions based on biochemical mechanisms [Grossberg and Schmajuk, 1989]. Here we propose a circuit using STM neurons which produces a burst of spikes after a given time. It has the same timing properties as the gated signal used in Grossberg's model and has stochastic properties providing a realistic basis for the Weber law relating accuracy of timing and length of the time. Further, a large range of timing properties can be achieved without requiring a large range of membrane time constants.

The circuit (figure 6) is composed of a pool of m STM neurons triggered by the CS. These have all the same lifetime and, while firing, produce a spike at each time-step. All m STM neurons feed into an m-pRAM with parameters set to cause firing only if the number of simultaneous input spikes is, say, between $n + 1$ and $n - 1$. Therefore,

as the decaying number of active STM neurons approaches n, the pRAM will fire a few spikes.

Biologically, the role of the m-pRAM would be played by an interneuron inhibiting a companion neuron. The threshold of the interneuron is set to cause firing only when more than n inputs are active. The companion neuron is provided with an after-hyperpolarization (AHP) mechanism [Steriade, 1984] such that, at the end of the inhibition, the AHP rebound causes it to produce a burst of spikes. Another possible circuit consists of a coupled neuron-interneuron receiving the same inputs from the pool of STM neurons. The interneuron inhibits the neuron and has a slightly higher threshold than the neuron. Thus, as the number of active STM neurons decreases, the interneuron stops firing, allowing the neuron to produce a few spikes before the number of STM neurons decreases further.
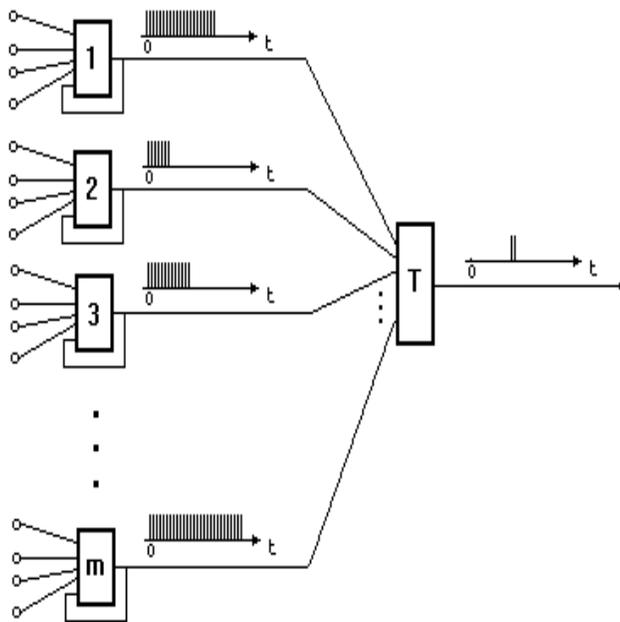

Figure 6: A neural timer. It consists of a set of m STM-neurons trigered by a stimulus. The pRAM "T" fires a few spikes when the number of active STM-neurons has decayed to n $(n < m)$.

The probability $P(n,t)$ that the number of active STM neurons has decayed to n at the time $t = k\Delta t$ is given by:

$$P(n,t) = C_n^m e^{-\frac{nk\Delta t}{\tau}} \left(1 - e^{-\frac{k\Delta t}{\tau}}\right)^{m-n} \tag{6}$$

Examples of the function $P(n,t)$ are shown in figure 7 in the case of the detection of $n = 5$ or more active STM neurons in a pool of $m = 5, 6, 7, 10, 15$ or 20 STM neurons having all the same lifetime τ. The probability $P(n,t)$ has its maximum at the time

14

$t_{max} = \tau ln(m/n)$. Interestingly, the most likely time of firing t_{max} can be controlled by only varying the fan-in of the m-pRAM while keeping constant the threshold n and the lifetime of the STM neurons.
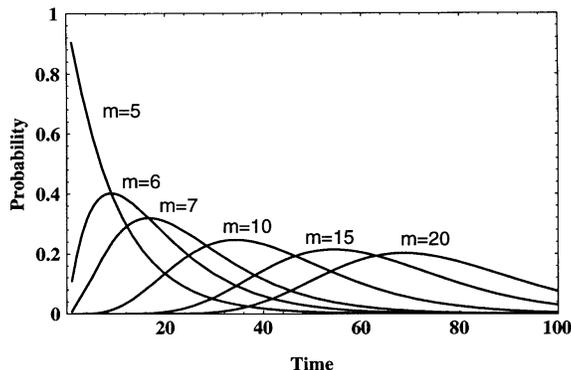

Figure 7: Time-dependent probability that the number of active STM-neurons be exactly n (in this case $n = 5$). As the pool m of STM-neurons becomes larger, it takes a longer time before the number n is reached. In the calculations of these curves, Δt was taken to be 1, and the time indicated on the horizontal axis is measured in timesteps k. We used $\alpha_{01} = 0.98$ corresponding to $\tau = 49.5$.

We may also look at the retention time of a short-term memory using a neuron receiving inputs from m STM neurons and having a threshold set to cause firing as long as n or more STM neurons among m are active. Here we call that a "t-neuron". The probability $M(n, t)$ that at least n STM neurons are still firing at time $t = k\Delta t$ is given by:

$$M(n, t) = \sum_{i=n}^{m} P(i, t) \tag{7}$$

Examples of the function $M(n, t)$ are given in figure 8 for various combinations of n and m. We have chosen $\alpha_{01} = 0.98$ and $n = m/2$ and it appears clearly that, the larger the value of m, the more precise the timer will be. For a very large number m of inputs, the function $M(n, t)$ tends toward a step function with the activity ending exactly at $t = \tau ln(m/n)$.

If the time of occurence of an input had to be estimated from the activity of a single pRAM with self-feedback, then one could only say: "It has just happened" if the unit is active or "It was a long time ago" if the unit is inactive. We see here that a more precise time estimation can achieved by averaging the activity of a large number of independent units. Using single neurons with self-feedback would be the most economic way to realize a large number of such independent memories. It would therefore be interesting to find out if the number of true autapses is larger in regions of the brain involved in time estimation. We may note that the use of electrotonic junctions rather than chemical synapses may be a way to reduce the stochasticity in a system and increase the lifetime of the activity in a loop.
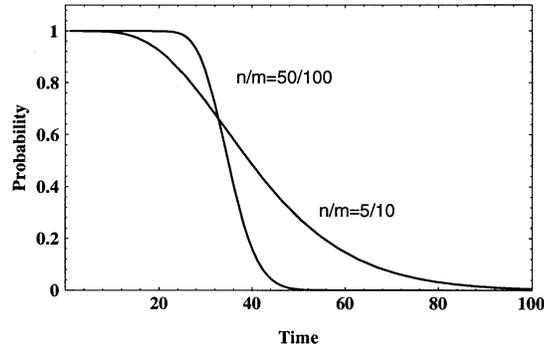
Figure 8: Time-dependent probability that n STM-neurons among m are in an active firing state. As the number of neurons involved increases, the decay of the number of active STM-neurons below n occurs in a smaller time-window and the timer becomes more precise. The parameters are the same as in the previous figure.

There are various data on timing accuracy in humans and animals [Wearden, 1990,1992] and it might be interesting to find out what combinations of n, m and time constants τ would best fit these data.

In experiments by Fuster and coworkers [1981] monkeys which had to perform a delayed response task showed a gradual decrease in performance as the retention time increased. The performance curves remained at 100% for delays smaller than 8sec and showed a progressive downward inflection for increasing delays. A level of approximately 85% was reached at a delay of 32sec. This is typical of a retention strategy based on a pool of independent STM neurons. If a single STM neuron is used, then the performance curve would show an exponential decay as in figure 2 or the first curve in figure 7. However, one may not exclude the possibility that such non-exponential decays emerge as a result of the dynamics of networks of excitatory neurons with threshold. It may also indicate a more complex retention strategy (See for instance the eye movements during the delay in [Fuster et al., 1981]).

The use of stochastic short-term memory units for the identification of temporal patterns has been suggested [Manevitz, 1992]. A model of temporal sequence learning, retrieval and recognition using a pool of units with memory properties similar to those of our t-neurons has been sucessfully applied to various problems [Reiss and Taylor, 1992].

5   Conclusion

Whether biological short-term memories are based on self-feedback connections or use lateral excitatory interconections in small clusters of neurons, we suggest that they have a stochastic behaviour characterized by a lifetime. The mathematical analysis of a pRAM neuron with self-feedback allows a simple description of the parameters determining this lifetime. The pRAM neuron allows also a simple implementation of a well controlled stochastic short-term memory.

We have outlined in this paper a number of interesting applications of neurons with self-sustained activity. Using the frequency gain due to the self-feedback loop one can conceive of a neuronal AND-function. Using the information storage property of such neurons, one can conceive of a simple multilayer neural network with synchronous and asynchronous pattern recognition capabilities. One can also construct a sequence detector. Using the limited life-time of the sustained activity one can create powerful timing circuits.

Aknowledgements

6   References

Alavi F.N. and Taylor J.G. (1992) *A simulation of the gated thalamo-cortical model,* in Alexander, I. and Taylor, J.G. (eds) "Artificial neural networks 2", Elsevier (Proc ICANN'92), 929-931.

Alavi F.N. and Taylor J.G. (1993) *A global competitive network for attention,* Neural Networks World (to appear).

Amit D.J. and Tsodyks M.V. (1991) *Quantitative study of attractor neural network retrieving at low spike rate I: Substrate - spikes, rates and neuronal gain,* Network, **2**, 259-273 and *Low-rate retrieval in symmetric networks,* Network, **2**, 275-294.

Bernander O., Douglas R., Martin K. and Koch C. (1991) *Synaptic background activity influences spatiotemporal integration in single pyramidal cells,* Proc. Natl. Acad. Sci. USA, **88**, 11569-11573.

Bressloff P.C. and Taylor J.G. (1993) *Low firing-rates in a compartmental model neuron,* J. Phys. A: Math. Gen, **26**, L165-L170.

Bugmann G. (1991a) *Can neurons realize multiple AND-functions ?,* Proc. 4th Intl. Conf. on Neural Networks and their Appl. (Neuronimes'91), Nimes, France, 757-760.

Bugmann G. (1991b) *Summation and multiplication: Two distinct operation domains of leaky integrate-and-fire neurons,* Network, **2**, 489-509.

Bugmann G. and Taylor J.G. (1993a) *Realizing a formal neuron by using pRAMs,* in proc. of the Weightless Neural Network Workshop (WNNW'93), Allison N.M. (ed), The Department of Electronics, Univ. of York, UK, 105-110.

Bugmann G. and Taylor J.G. (1993b) *A model for latencies in the visual system.* in Proc. of the International Conference on Artificial Neural Networks (ICANN'93), Amsterdam, Springer.

Burgess N. and Hitch G.J. (1992) *Toward a network model of the articulatory loop*, J. of Memory and Language, **31**, 429-460.

Caianiello, E.R. (1961) *Outline of a theory of thought-processes and thinking machines*, J. Theoret. Biol., **2**, 204-235.

Chevalier G. and Deniau J.M. (1990) *Dishinhibition as a basic process in the expression of striatal functions*, Trends in Neurosci., **13**, 277-280.

Douglas, R.J. and Martin, K.A.C. (1990) *Control of neuronal output by inhibition at the axon initial segment*, Neural Comp., **2**, 283-292.

Douglas, R.J. and Martin, K.A.C. (1991) *A functional microcircuit for cat visual cortex*, J. Physiol., **440**, 735-769.

Ferster D. (1986) *Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex*, J. Neurosci., **6**, 1284-1301.

Freund T.F., Martin K.A.C. and Whitteridge D. (1985) *Innervation of cat visual area 17 and 18 by physiologically identified X- and Y-type thalamoc afferents. I. Arborization patterns and quantitative distribution of postsynaptic elements*, J. Comparative Neurology, **242**, 263-274.

Funahashi S., Bruce C.J. and Goldman-Rakic P.(1989) *Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex*, J. Neurophysiology, **61**, 331-349.

Fuster J.M., Bauer R.H. and Jervey J.P. (1981) *Effects of cooling inferotemporal cortex on performances of visual memory tasks*, Exp. Neurol., **71**, 398-409.

Fuster J.M. (1990) *Inferotemporal units in selective visual attention and short-term memory*, J. Neurophys., **64**, 681-697.

Gochin P.M., Miller E.K., Gross C.G. and Gerstein G.L. (1990) *Functional interaction among neurons in inferior temporal cortex of the awake monkey*, Exp. Brain Res., **84**, 505-516.

Gorse D. and Taylor J.G. (1990) *A general model of stochastic neural processing*, Biol. Cybern., **63**, 299-306.

Grossberg S. (1982) *Studies of Mind and Brain*, Reidel, Boston

Grossberg S. and Schmajuk N.A. (1989) *Neural dynamics of adaptative timing and temporal discrimination during associative learning*, Neural Networks, **2**, 70-102.

Kirillov A.B., Myre C.D. and Woodward D.J. (1993) *Bistability, switches and working memory in a two-neuron inhibitory-feedback model*, Biol. Cybern., **68**, 441-449.

Libet B. (1989) *Conscious subjective experience vs. unconscious mental functions: A theory of the cerebral processes involved*, in Cotteril R.M.J. (ed) "Models of brain function", Cambridge Univ. Press, 35-49.

Maex, R. and Orban, G.A. (1992) *A model circuit for cortical temporal low-pass filtering*, Neural Computation, 4, 932-945.

Manevitz, L.M. (1992) *Temporal sparse memory: Identifying temporal patterns via homeomorphic contraction of memory*, in Alexander, I. and Taylor, J.G. (eds) "Artificial neural networks 2", Elsevier (Proc ICANN'92), 1651-1654.

Meyer G.(1987) *Form and spatial arrangement of neurons in the primary motor cortex of man*, J. Comp. Neurol., **262**, 402-428.

Mushiake H., Inase M. and Tanji J. (1990) *Selective coding of motor sequences in the supplementary motor area of the monkey cerebral cortex*, Exp. Brain Res., **82**, 208-210.

Nicoll, R.A. and Jahr, C.E. (1982) *Self-excitation of olfactory bulb neurons*, Nature, **296**, 441-444.

Pöppel E. (1987) *Time perception*, in "Encyclopedia of Neuroscience", Birkhäuser, Boston, 1215-1216.

Pun, R.Y.K., Neale, E.A., Guthrie, P.B. and Nelson, P.G. (1986) *Active and inactive synapses in cell culture*, J. Neurophysiology, **56**, 1242-1256.

Rashevsky N. (1960) *Mathematical Biophysics*, Dover, New York

Redman, S. (1990) *Quantal analysis of synaptic potentials in neurons of the central nervous system*, Physiological Reviews, **70**, 165-198.

Reiss M. and Taylor J.G. (1991) *Storing temporal sequences*, Neural Networks, 4, 773-787.

Rolls, E. (1990) *Principles underlying the representation and storage of information in neural networks in the primate hippocampus and cerebral cortex*, in Zornetzer, S.F., Davis, J.L. and Lau, C. (eds) "An introduction to neural and electronic networks", Academic Press, London.

Smyrnis N., Taira M., Ashe J. and Georgopoulos A.P. (1992) *Motor cortical activity in a memorized delay task*, Exp. Brain Res., 92, 139-151.

Steriade M. (1984) *The excitatory-inhibitory response sequence in thalamic and neocortical cells: State-related changes and regulatory systems*, in Edelmann G.M. (ed) "Dynamic aspects of neocortical function", Wiley, New York, 107-157.

Tanaka K., Fukada Y. and H.A. Saito (1989) *Underlying mechanism of the response specificty of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey*, J. Neurophysiology, **62**, 642-656.

Taylor, J.G. (1972) *Spontaneous behaviour in neural networks*, J. Theor. Biol., **36**, 513-528.

Taylor J.G. (1992a) *From single neuron to cognition*, in Alexander, I. and Taylor, J.G. (eds) "Artificial neural networks 2", Elsevier (Proc ICANN'92), 11-16.

Taylor, J.G. (1992b) *Towards a neural network model of the mind*, Neural Networks World, **2**, 797-812.

Taylor, J.G. (1993) *Modelling the mind*, in: Orchard G. (ed) Proc. of Neurocomputing, Research and Applications, Queen's University, Belfast, June 1992, IOP Publ. Ltd, Bristol, 1-20.

Treves, A. and Rolls, E.T. (1992) *Computational constraints suggest the need for 2 distinct input systems to the hippocampal CA3 network*, Hippocampus, **2**, 189-200.

Van der Loos, H. and Glaser, E.M. (1972) *Autapses in neocortex cerebri: Synapses between a pyramidal cell's axon and its own dendrites*, Brain Res., **48**, 355-360.

Verri A., Straforini M. and Torre V. (1992) *Computational aspects of motion perception in natural and artificial neural networks*, in Orban G.A. and Nagel H.H. (eds.) "Artficial and biological vision systems", Springer, 71-92.

Wearden J.H. (1990) *The erching resonse and the law of animal timing*, J. Exp. Psychology: Animal Behaviour Processes, **16**, 150-161.

Wearden J.H. (1992) *Temporal generalization in humans*, J. Exp. Psychology: Animal Behaviour Processes, **18**, 134-144.

Weber, A.J., Kalil, R.E. and Behan, M. (1989) *Synaptic connections between corticulogeniculate axons and interneurons in the dorsal lateral geniculate nucleus of the cat*, J. Comp. Neurol., **289**, 156-164.

Zipser, D. (1991) *Recurrent network model of the neural mechanism of short term active memory*, Neural Computation, **3**, 179-193.