A Model for Fast Analog Computations with Noisy Spiking Neurons

Wolfgang Maass

Institute for Theoretical Computer Science
Technische Universitaet Graz
Klosterwiesgasse 32/2
A-8010 Graz, Austria
e-mail: maass@igi.tu-graz.ac.at
Tel.: +43 316 873 5822, Fax: +43 316 873 5805

Abstract

We show that networks of spiking neurons can simulate arbitrary feedforward sigmoidal neural nets in a way which has previously not been considered. This new approach is based on temporal coding by single spikes (respectively by the timing of synchronous firing in pools of neurons), rather than on the traditional interpretation of analog variables in terms of firing rates.

As a consequence we can show that networks of noisy spiking neurons are "universal approximators" in the sense that they can approximate with regard to temporal coding *any* given continuous function of several variables.

1 Results

Traditionally one views the firing rate of a neuron as the representation of an analog variable in analog computations with spiking neurons. However with regard to fast cortical computations this view is inconsistent with experimental data. Thorpe and Imbert (Thorpe, 1989) have demonstrated that visual pattern analysis and pattern classification can be carried out by humans in just 100 msec, in spite of the fact that it involves a minimum of 10 synaptic stages from the retina to the temporal lobe. The same speed of visual processing has been measured by Rolls et al. in macaque monkeys. Furthermore they have shown that a single cortical area involved in visual processing can complete its computation in just 20-30 msec (Rolls, 1994a, 1994b). On the other hand the firing rates of neurons involved in these computations are usually below 100 Hz, and hence at least 20-30 msec would be needed just to sample the current firing rate of a neuron.

We explore in this article the theoretical possibilities of analog computations with noisy spiking neurons in *temporal coding*, where analog variables $x \in [0, \gamma]$ are encoded by the firing time $T - x$ of a spiking neuron (as in Hopfield, 1995), respectively by the median firing time of a pool of neurons in a more noise-robust implementation. It is well-known that EPSP's and IPSP's can *shift* the firing time of a neuron by a small amount. We show that this simple mechanism is in principle sufficient for computing in temporal coding arbitrarily complex bounded continuous multi-variable functions in just 20 msec.

The results reported in this article are rigorous theoretical results. We refer to (Maass, 1995) for all details that have to be omitted in this summary.

Assume that the membrane potential at the trigger zone of a neuron v is determined at time t by n postsynaptic potentials $h_1(t), \ldots, h_n(t)$, which were caused by the firing of presynaptic neurons $a_1, \ldots, a_n$ at times $T_0 - s_1, \ldots, T_0 - s_n$. We will focus on the initial segments of length $d + \triangle$ of these PSP's, during which they can be approximated well by a piecewise linear function

$$h_i(t) = \begin{cases} 0 & , \quad \text{if} \quad t - (T_0 - s_i) < d \\ w_i \cdot (t - (T_0 - s_i) - d) & , \quad \text{if} \quad d \le t - (T_0 - s_i) \le d + \triangle \end{cases} ,$$

where $w_i \ge 0$ in the case of an EPSP and $w_i < 0$ in the case of an IPSP. Let t_v be the time when the sum of these PSP's reaches the firing threshold Θ and neuron v fires. Assume that the n PSP's are at time t_v within their linearly rising (EPSP) respectively linearly decreasing (IPSP) segments of length $\triangle$, i.e.

$$d \;\le\; t_v - (T_0 - s_i) \;\le\; d + \triangle \quad \text{for} \quad i = 1, \ldots, n \quad . \tag{1}$$

Then we have $\Theta = \sum_{i=1}^n h_i(t_v)$, and hence

$$t_v \;=\; \frac{\Theta}{\sum_{i=1}^n w_i} \;+\; T_0 \;+\; d \;-\; \frac{1}{\sum_{i=1}^n w_i} \cdot \underline{w} \cdot \underline{s} \quad . \tag{2}$$

Thus neuron v essentially outputs for inputs $s_1, \ldots, s_n \in [0, \gamma]$ the (normalized) inner product $\underline{w} \cdot \underline{s}$ in temporal coding. The constraint (1) is satisfied if $\sum_{i=1}^n w_i > 0$, $\sum_{i=1}^n w_i \cdot s_i \ge 0$, and Θ is chosen so that

$$\gamma \cdot \sum_{i=1}^n w_i \le \;\Theta\; \le\; (\triangle - \gamma) \;\cdot\; \sum_{i=1}^n w_i \quad . \tag{3}$$

The parameter γ has to be chosen $\le \triangle/2$ in a theoretically rigorous simulation. However we expect that a larger value of γ (around 5 msec) can actually be employed for computations in biological neural systems (see Maass, 1995). If we fix $s_1 \equiv 0$, then we can choose w_1 so that $\sum_{i=1}^{n} w_i = 1$ for *arbitrary* given $w_2, \dots, w_n \in \mathbf{R}$, and hence compute in this way a linear function of $n-1$ inputs $s_2, \dots, s_n$ with *arbitrary* given weights $w_2, \dots, w_n \in \mathbf{R}$.

If all weights w_i are multiplied with a factor $\lambda > 1$, this will increase the slope of the membrane potential $\sum_{i=1}^{n} h_i(t)$ at the time when it crosses the threshold. As a consequence the output of this neuron in temporal coding will become more noise-robust, both from the point of view of the common mathematical models for noise in spiking neurons (Gerstner, 1994, Maass, 1995, 1996), and from the point of view of experimental results (Mainen, 1995).

In an even more noisy setting when synapses and/or neurons fail with significant probability, one may replace each single neuron v in our construction by a pool P_v of neurons with approximately identical connection. In this case the mean firing time of those neurons in P_v that do fire plays the role of the firing time t_v of the single neuron v . One should note that in this generalized interpretation a network of spiking neurons can in principle perform analog computations in temporal coding with high reliability, but single-neuron recordings from a small number of neurons in this network would not necessarily yield repetitions of the same firing pattern for repetitions of the same computation.

So far we have shown that a spiking neuron can compute (for a certain range of its parameters) in temporal coding any *linear* function $\underline{s} \mapsto \underline{w} \cdot \underline{s}$. In particular it can detect to what degree a learned pattern $\underline{w}$ (which is stored in the efficacies of its synapses) is present in a stimulus $\underline{s}$.

We will now show that by using *two* layers of spiking neurons, one can in fact approximate in temporal coding *any* given bounded continuous function $F : [0, \gamma]^n \to [0, \gamma]^k$, with any desired degree of precision. It is well-known that any such function F can be approximated by a feedforward sigmoidal neural net with two layers. Furthermore it is known that the exact form of the activation function of the gates is irrelevant for this result (Leshno, 1993). Hence one can employ in particular the piecewise linear activation function π_γ with

$$\pi_\gamma(x) = \begin{cases} \gamma & , \text{if } \ x > \gamma \\ x & , \text{if } \ 0 \le x \le \gamma \\ 0 & , \text{if } \ x < 0 \end{cases} .$$

We show that one can simulate with spiking neurons in temporal coding a given sigmoidal gate G with activation function π_γ , i.e. a gate that outputs

$$G(s_1, \cdots, s_n) = \begin{cases} \gamma & , \text{if } \ \underline{w} \cdot \underline{s} > \gamma \\ \underline{w} \cdot \underline{s} & , \text{if } \ 0 \le \underline{w} \cdot \underline{s} \le \gamma \\ 0 & , \text{if } \ \underline{w} \cdot \underline{s} < 0 \end{cases}$$

for any $s_1, \ldots, s_n \in [0, \gamma]$. One adds auxiliary neurons to the preceding construction whose firing is time-locked with the onset of the stimulus. These can prevent a firing of neuron v before time $T - \gamma$ for $T := \Theta + T_0 + d$ (assume for simplicity that $\sum_{i=1}^{n} w_i = 1$), and they can force v to fire at the latest by time T (see Maass, 1995, for details).

With k layers of spiking neurons one can thus simulate in temporal coding with any desired precision any given k-layer net of sigmoidal neurons that employ the activation function π_γ . Hence in combination with the abovementioned results we have shown that with 2 layers of spiking neurons one can in principle approximate in temporal coding any given bounded continuous function with a computation time of ≤ 20 msec.

Finally we would like to point out that with biological neurons one can carry out very similar multi-layer computations *without* simulating explicity the activation function π_γ (i.e. without forcing v to fire within a specific time window $[T - \gamma, T]$) . The natural biological form of EPSP's and IPSP's has in general the effect that if v fires before $T - \gamma$, the resulting PSP in a neuron v' on the next layer is at the time when v' fires no longer within its initial *linearly* increasing (respectively decreasing) phase. As a result the "too large" value of the output of v in temporal coding is effectively *reduced* in size through the form of the PSP in v' . Dually, if v fires *after* the time window $[T - \gamma, T]$, the resulting PSP will in general still have its initial value 0 at the time when v' fires, which amounts to replacing the (theoretically) negative value of this output of v in temporal coding by the value 0 . In this way the *bounded* functional form of PSP's in v' achieves a similar effect as the application of a bounded activation function to the value $\underline{w} \cdot \underline{s}$ that is output by v in temporal coding.

Our new model for linear and nonlinear computations with spiking neurons also provides a new perspective on learning. For example in this model the Hebb rule would want to increase the efficacy w_i of the synapse between a presynaptic neuron a_i and v if the outputs of a_i and v in *temporal* coding are correlated, i.e. if v fires within a certain time window after a_i (and decrease the efficacy otherwise). Exactly this type of synaptic modulation has recently been found in pyramidal neurons (Markram, 1995).

2 Remarks

1. In an alternative biological interpretation of our construction one can exploit finer details of dendritic integration in order to simulate *several* layers of a sigmoidal neural net with gates of fan-out 1 by a *single* spiking neuron. The idea is here to exploit boosting phenomena via voltage-dependent channels at branching points ("hot spots", see e.g. Mel, 1993, Sheperd, 1995) of the dendritic

tree in order to simulate by a hot spot a sigmoidal gate in temporal coding (where the output corresponds to the "firing time" of such hot spot). In this interpretation the weight on the edge from this gate to a gate on the next level of the simulated sigmoidal neural net would correspond to the amplitude of the "action potentials" resulting from this hot spot, combined with the distance of this hot spot to the next "higher" hot spot in the dendritic tree (respectively to the soma). In this biological interpretation one can avoid the use of (unreliable) synapses for the simulation of internal connections in the sigmoidal neural net.

2. Our construction shows as a special case that a *linear* function of the form $\underline{s} \to \underline{w} \cdot \underline{s}$ for real-valued vectors $\underline{s} = \langle s_1, \ldots, s_n \rangle$ and $\underline{w} = \langle w_1, \ldots, w_n \rangle$ can be computed very efficiently (and very fast) by a spiking neuron v in temporal coding. In this case no auxiliary neurons are needed.

   A fast computation of linear functions is obviously relevant in many biological contexts, such as coordinate transformations between different frames of reference, or the analysis of a complex stimulus $\underline{s}$ in terms of many stored patterns $\underline{w}, \underline{w}', \ldots$.

   For example in an olfactory neural system (see e.g. Hopfield, 1991, 1995) the stimulus $\underline{s}$ may be thought of as a superposition of various stored basic odors $\underline{w}, \underline{w}', \ldots$. In this case the output $y = \underline{w} \cdot \underline{s}$ of neuron v in temporal coding may be interpreted as the amount by which the basic odor $\underline{w}$ (which is stored in the efficacies of the synapses of v) is present in the stimulus $\underline{s}$. Furthermore another neuron $\tilde{v}$ on the next layer might receive as its input $\underline{y} = \rangle y, y', \ldots, \langle$ from several such neurons $v, v', \ldots$, i.e. $\tilde{v}$ receives the "mixing proportions" $y = \underline{w} \cdot \underline{s}$, $y' = \underline{w}' \cdot \underline{s}$, for various stored basic odors $\underline{w}, \underline{w}', \ldots$ in temporal coding. This neuron $\tilde{v}$ on the second layer can then continue the pattern analysis by computing for this input $\underline{y}$ the inner product $\underline{W} \cdot \underline{y}$ with some stored "higher order pattern" $\underline{W}$ (e.g. the composition of basic odors that is characteristic for an individual animal) that is encoded in the efficacies of the synapses of neuron $\tilde{v}$. Such multi-layer pattern analysis is faciliated by the fact that the here considered neurons encode their output in the same way in which their input is encoded (in contrast to the approach in Hopfield, 1995).

   One also gets in this way a very fast implementation of Linskers network (Linsker, 1988) with spiking neurons in temporal coding.

3 References

Gerstner, W. and van Hemmen, J. L. (1994) *How to describe neuronal activity: spikes, rates, or assemblies?* Advances in Neural Information Processing Systems, vol. 6, Morgan Kaufmann (San Mateo) 463-470.

Hopfield, J. J. (1991) *Olfactory computation and object perception.* Proc. Nat. Ac. Sci. USA, vol. 88, 6462-6466.

Hopfield, J. J. (1995) *Pattern recognition computation using action potential timing for stimulus representations.* Nature, vol. 376, 33-36.

Leshno, M., Lin, V. Y., Pinkus, A., and Schocken, S. (1993) *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function.* Neural Networks, vol. 6, 861-867.

Linsker, R. (1988) *Self-organization in a perceptual network.* Computer Magazine, vol. 21, 105-117.

Maass, W. (1995) *Fast sigmoidal networks via noisy spiking neurons*, to appear in Neural Computation.

Maass, W. (1996) *On the computational power of noisy spiking neurons.* Advances in Neural Information Processing Systems, vol. 8, MIT-Press (Cambridge), 211-217.

Mainen, Z.F. and Sejnowski, T. J. (1995) *Reliability of spike timing in neocortical neurons.* Science, vol. 268, 1503-1506.

Markram, H. and Sakmann, B. (1995) *Action potentials propagating back into dendrites trigger changes in efficacy of single-axon synapses between layer V pyramidal neurons.* Abstract 788.11 in Vol. 21 of the Proc. of the Conference of the Society for Neuroscience.

Mel, B. W. (1993) *Synaptic integration in an excitable dendritic tree.* J. of Neurophysiology, vol. 70, 1086-1101.

Rolls, E. T. (1994a) *Brain mechanisms for invariant visual recognition and learning.* Behavioural Processes, vol. 33, 113-138.

Rolls, E. T. and Tovee, M. J. (1994b) *Processing speed in the cerebral cortex, and the neurophysiology of visual backward masking.* Proc. Roy. Soc. B., vol. 257, 9-15.

Shepherd, G. M. (1994) *Neurobiology*, 3rd ed. Oxford University Press (New York).

Thorpe, S. J. and Imbert, M. (1989) *Biological constraints on connectionist modelling.* In: Connectionism in Perspective, Pfeifer, R., Schreter, Z., Fogelman-Soulié, F., and Steels, L., eds., Elsevier (North-Holland).