Temporal Hebbian learning in Rate-Coded Neural Networks: A theoretical approach towards classical conditioning

Bernd Porr and Florentin Wörgötter

University of Stirling, Department of Psychology, Stirling FK9 4LA, Scotland
http://www.cn.stir.ac.uk

Abstract. A novel approach for learning of temporally extended, continuous signals is developed within the framework of rate coded neurons. A new temporal Hebb like learning rule is devised which utilizes the predictive capabilities of bandpass filtered signals by using the derivative of the output to modify the weights. The initial development of the weights is calculated analytically applying signal theory and simulation results are shown to demonstrate the performance of this approach. In addition we show that only few units suffice to process multiple inputs with long temporal delays.

1 Introduction

The learning rule by Donald Hebb has always been one of the basic concepts of computational neuroscience [1]. Learning is achieved through changes of the synaptic weights between neurons. The synaptic weight is increased when the pre- and postsynaptic spiking activity is correlated and decreased vice versa. This learning rule is symmetric in time within a certain time window. In contrast to this traditional approach the temporal Hebb rule is asymmetric in time: A weight will be strengthened only if the input precedes the output by a short interval. If the order of input and output is reversed the weight will be decreased [2–4].

We emphasize here the most ubiquitous temporal learning known as ”classical conditioning”. It takes place on rather long time scales and normally requires two input stimuli, the unconditioned and the conditioned stimulus, which often arise from different sensorial modalities. The unconditioned stimulus (food) is followed by an output event (salivation). After learning the conditioned stimulus (bell), which always precedes the conditioned stimulus will elicit the same event and, thus, the unconditioned stimulus can be interpreted as a predictor of the conditioned stimulus (the sound of a bell predicts the feeding).

So far temporal sequence learning has in general been achieved by the “temporal difference (TD)-algorithm” [7, 8]. This algorithm is discrete in time and amplitude. Instead, in our approach we have chosen a rate code description as the level of abstraction, which allows to treat continuous signals and this has the advantage that we can develop the theory in an analytically solvable way.

In order to implement our algorithm we will introduce neuronal filter circuits which will act as damped oscillators and devise a new learning rule. This rule directly uses the temporal change (the derivative) of the neuronal output activity to modify the weights in the network. In addition, our system has the feature, that very few components are required to cover large time intervals (in contrast to the TD-algorithm) and it allows to use multiple inputs, which are the two critical prerequisites for “classical conditioning”.

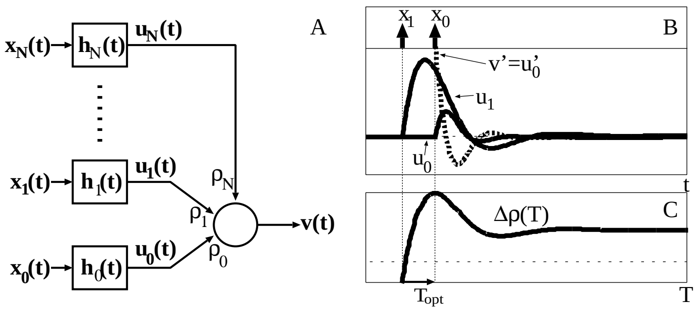

Fig. 1. A) The basic circuit in the time domain. B) shows the inputs x, the impulse responses u for a choice of two different resonators h and the derivative of the output v'. C) shows the initial weight change $\rho_1(T)_{t=0}$. T_{opt} is the optimal time difference between x_1 and x_0 for a maximum change of the weight ρ_1. The time difference of $x_0, x1$ in B is adjusted to T_{opt}, as well. Parameters: $f_0 = 0.1$, $f_1 = 0.05$, $Q_0 = Q_1 = Q = 1$.

2 The neuronal circuit

We consider a system of $N + 1$ units h receiving a continuous input signal x and producing a continuous output u. The input units connect with weights ρ to one output unit v (Fig. 1A). All input units are in principle equivalent but we will use h_0 to denote the one unit which transmits the unconditioned stimulus. The output v is then given as:

$$v = \rho_0 u_0 + \sum_{i=1}^{N} \rho_i u_i \qquad \text{with} \qquad u_i(t) = x_i(t) * h_i(t) \tag{1}$$

The transfer function h shall be that of a *resonator* with Frequency f and Quality Q which transforms a δ-pulse input into a damped oscillation (see Fig. 1B). For later use we will define the transfer function of the resonator in the *Laplace* domain as: $H(s) = [(s + p)(s + p^*)]^{-1}$ with $p = a + ib$, $a = \pi f/Q$ and $b = \sqrt{(2\pi f)^2 - a^2}$.

Learning (viz. weight change) takes place according to a new learning rule changing the weights in each time-step by $\Delta\rho_i$, with:

$$\Delta\rho_i = \mu u_i v' \qquad \mu \ll 1 \tag{2}$$

where the weight change depends on the correlation between u_i and the *derivative* of the output v (see Fig. 1B).

3 Performance

3.1 Goal

We state our goal as: After learning, the output unit shall produce a well discernible signal v (e.g., of high amplitude and steeply rising) in response to the earliest occurring conditioned stimulus $x_j, j \geq 1$.

3.2 Analytical solution of the initial weight change

Similar to other approaches [9] we compute the initial development of the weights as soon as learning starts, because this is indicative of the continuation of the learning process. As usual we require that all weight changes occur on a much longer time scale (i.e., very slowly) as compared to the decay of the oscillatory responses u. This allows us to treat the system in a steady state condition, thus, $\Delta\rho \to 0$ for $\Delta t \to 0$.

In order to calculate this weight change we will now introduce several restrictions: A) The weight of the unconditioned stimulus is set to $\rho_0 = 1$ and kept constant throughout learning. B) We will consider only one unit that can learn, thus, $N = 1$. C) Accordingly we have to deal with only two input functions x_0, x_1 and we define them as (delayed) δ-pulses: $x_0(t) = \delta(t+T), T \geq 0$ and $x_1(t) = \delta(t)$. These restrictions will allow us to develop the theory but can be waived in the end without affecting our basic findings.

Because we assume steady-state, we can rewrite the product in the learning rule (Eq. 2) as a correlation integral between input and output:

$$\Delta\rho_1(T) = \mu \int_0^{\infty} u_1(T+\tau)v'(\tau)d\tau \tag{3}$$

The value of T represents the delay between conditional and unconditional stimulus. In order to assure optimal learning progress we must, therefore, determine that particular value T_{opt} for which a maximal weight change is observed. Because we are interested in the initial weight change, we can assume $\rho_1(t) = 0$ for $t = 0$. and Eq. 3 turns into:

$$\Delta\rho_1(T)_{t=0} = \mu \int_0^{\infty} u_1(T+\tau)u_0'(\tau)d\tau \tag{4}$$

In order to solve Eq. 4 we have to switch to the LAPLACE domain where $H(s)$ is the transfer function of the resonatore as defined earlier. We get:

$$\Delta\rho_1(T)_{t=0} = \mu\frac{1}{2\pi}\int_{-\infty}^{+\infty} H_1(i\omega)\left[-i\omega e^{Ti\omega}H_0(-i\omega)\right]d\omega \tag{5}$$

we set $P^+ = |p_1|^2 + |p_0|^2$ and $P^- = |p_1|^2 - |p_0|^2$ and get after integration with the method of residues:

$$\Delta\rho_1(T)_{t=0} = \mu\frac{b_1P^-\cos(b_1T) + (a_1P^+ + 2a_0|p_1|^2)\sin(b_1T)}{b_1(P^+ + 2a_1a_0 + 2b_1b_0)(P^+ + 2a_1a_0 - 2b_1b_0)}e^{-Ta_1} \tag{6}$$

This result is plotted in Fig. 1. If we assume identical resonators h_0, h_1, we have $a_0 = a_1 = a$ and $b_0 = b_1 = b$ and Eq. 6 simplifies to:

$$\Delta\rho_1(T)_{t=0} = \mu \frac{1}{4ab} \sin(bT) e^{-aT} \tag{7}$$

This is simply the impulse response of the resonator itself. Thus, in this special case the learning curve can be tuned by the choice of a certain impulse response of the resonator. In other words it is possible to implement an *analog temporal Hebb* learning behaviour and treat this analytically. Before showing simulation results we note that the above obtained analytical results can be extended to cover the most general system structure as represented in Fig. 1A. With Eq. 1 transformed in the LAPLACE domain the learning rule turns to:

$$V(s) = \sum_{j=0}^{N} \rho_j X_j(s) H_j(s) \tag{8}$$

$$\Delta\rho_i = -\mu \frac{i\omega}{2\pi} \int_{-\infty}^{+\infty} V(-i\omega) U_i(i\omega) d\omega, \tag{9}$$

where Eq. 9 is the general form of Eq. 5. It should be noted that for all $\Delta\rho_i$ this integral can still be evaluated analytically in the same way as in the special case with two resonators discussed above.

3.3 Simulations

In order to validate the approach several simulations of increasing complexity were performed. As the analytical solution treats only the initial learning step we determined the learning behavior for $t >> 0$ with a sequence of two δ pulses x_1, x_0 as inputs. Fig. 2A shows that at the initial learning step the output function v still coincides with the unconditioned stimulus response u_0. After 10 repetitions of the pulse-sequence (Fig. 2 B), the output function has shifted forward to u_1. This forward shift can be interpreted in the sense of classical conditioning: after learning the output signal v predicts the occurrence of u_0 having been conditioned by u_1.

A sufficiently strong correlation occurs in this setup only for a small temporal difference between the input pulses. To improve on this one can use several resonators $h_1, ..., h_n$ with different frequencies which will all receive input from the conditioning stimulus x_1, defining: $T_0 = T; x_j = x_1, T_j = 0$ for $j \geq 1$ (Fig. 1A and Eq. 8). In Fig. 2 C-E we use five resonators to represent x_1 and show the results after 10 learning steps for increasingly larger intervals T which separate the input pulses. We note that in all cases the output is a superposition of the different resonator outputs and that a well discernible early signal component exists in response to the conditioning stimulus x_1. This is due to the fact that in this setup the learning process can be seen as a sequential improvement of the rising edge at the onset of the conditioned stimulus.

In the next simulation we copied all conditions from the last simulation but replaced the δ-pulses at x_0 and x_1 by temporally extended signals (Fig 2 F/G). This was done to show that the system is able to deal with complex input signals. In an application the

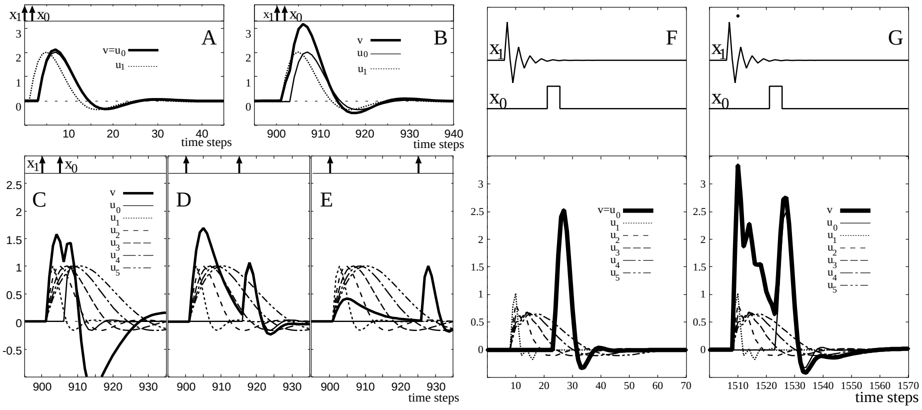

Fig. 2. Simulation results. Arbitrary time steps were used in the simulations. Pulse sequences were repeated every 100 time-steps, the first starting at zero. Resonators were implemented digitally as IIR-filters, which leads to a small onset-delay of 2 time-steps. They had always a value of $a = 0.1\pi$. A,B) are obtained with two identical resonators h_0, h_1 with: $b_{0,1} = 1.09$. The other parameters were $\mu = 0.01$ and $T = 2$. A) Result for $t = 0$, B) for $t = 900$. C-E) are calculated with six resonators, five (curves of $u_1, ..., u_5$) receiving input from x_1 and one (curve of u_0) from x_0. Resonator parameters are: $f_0 = 1.09$, $f_j = \frac{f_0}{j}, j \geq 1$, μ was set to 0.11. Results are shown for $t = 900$. C) $T = 5$. D) $T = 15$, E) $T = 25$. F,G) have the same conditions as D) x_0 was also a square wave but x_1 was sine burst.

signal x_0 (the unconditioned stimulus) could be a light stimulus and the signal x_1 (the conditioned stimulus) a sound event. After learning the system is again able to detect the sine burst as the predictor of the square pulse x_0.

4 Discussion

From the above results it can be deduced that the system generalizes without problems to more generic input combinations (e.g. more than two inputs). An important feature of other algorithms for sequence learning is that after convergence the output function will approximate an “expectation potential”, which is essentially a rectangular function, starting at x_1 and ending at x_0 (see [10], chapter 10). This is the reason, why our theory was developed using resonators (band-passes), because they allow to compose arbitrary shapes of v, such as expectation potentials. We realize, however, that a pure low-pass characteristic would also suffice to induce temporal learning, which relates our approach to the theory of adaptive filters [11]. This helps to understand the predictive properties of our system, because from this context it is known that the derivative of a low-pass filtered signal acts as a predictor of the signal [12]. However, adaptive filters rely on the (gradually shifting) self-similarity of a single signal - hence on its auto-correlation properties. Whereas our approach gains its predictive power from the cross-correlation properties between two (or more) signals such as in biological classical conditioning. Therefore, we expect that this approach will also prove useful in a more technologically oriented context.

References

1. Donald O. Hebb. *The organization of behavior*. Science Ed., New York, 1967.
2. A. H. Klopf. A neuronal model of classical conditioning. *Psychobiology*, 16(2):85–123, 1988.
3. Richard Kempter, Wulfram Gerstner, and J. Leo van Hemmen. Hebbian learning and spiking neurons. *Physical Review E*, 59:4498–4514, 1999.
4. S. Song, K. D. Miller, and L. F. Abbott. Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, 3:919–926, 2000.
5. Gerstner, Kempter, van Hemmen, and Wagner. A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383:76–78, 1996.
6. Ken D. Miller. Receptive fields and maps in the visual cortex: Models of ocular dominance and orientation columns. In E Donnay, J.L. van Hemmen, and K. Schulten, editors, *Models of Neural Networks III*, pages 55–78. Springer-Verlag, 1996.
7. Wolfram Schultz, Peter Dayan, and P. Read Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.
8. R.S. Sutton and A.G. Barto. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev*, 88(2):135–170, 1981.
9. E. Oja. A simplified neuron model as a principal component analyzer. *J Math Biol*, 15(3):267–273, 1982.
10. P Dayan and L. F. Abbott. *Theoretical Neuroscience*. MIT Press, Cambridge MA, 2001.
11. S. S. Haykin. *Adaptive filter theory*. Prentice Hall, Englewood Cliffs, 1991.
12. S. M. Bozic. *Digital and Kalman filtering: an introduction to discrete-time filtering and optimum linear estimation*. E. Arnold, London, 1994.