Gradient Learning in Networks of Smoothly Spiking Neurons

Jiří Šíma*

Institute of Computer Science, Academy of Sciences of the Czech Republic,
P. O. Box 5, 18207 Prague 8, Czech Republic, sima@cs.cas.cz

A prominent position among neural network models has recently been occupied by biologically more plausible networks of spiking (pulse) neurons [3]. One of the main open issues is the development of a practical learning algorithm for these networks. The first published learning algorithm *SpikeProp* [2] is a backpropagation-like procedure based on the gradient descent method which learns the desired firing times of output neurons by adapting the weights in the Spike Response Model SRM_0 of spiking neurons. Plenty of experiments with *SpikeProp* was carried out which clarified, e.g., the role of the parameter initialization and negative weights [5]. The performance of the original algorithm was improved by adding the momentum term [8] and the learning rules for synaptic delays, thresholds, and time constants [6] which resulted in faster convergence and smaller network sizes. An essential speedup was also achieved by approximating the firing time function using the logistic sigmoid [1]. The *SpikeProp* algorithm was also extended to recurrent architectures [7]. See paper [4] for a recent review of other supervised learning methods in spiking neural networks.

Nevertheless, the *SpikeProp* method and its modifications do not usually allow more than one spike per neuron. At the core of these difficulties the fact is that the spike creation or removal due to the weight updates is by nature a very discontinuous process while in contrast the gradient learning algorithms do essentially require the error function to be perfectly smooth for their implementations. Our **main results** concerning this problem are summarized in the following three paragraphs.

1. We have proposed a **new model of a smoothly spiking neuron** which approximates the simplified Spike Response Model SRM_0 [3] and, at the same time, is perfectly smooth. The development of such a model is not straightforward since the creation or deletion of pulses, on one hand, and the differentiability of the function computed by the spiking neuron, on the other hand, are in a direct contradiction. Nevertheless, we have succeeded to make the SRM_0 model smooth not only with respect to the time but also along the weight and delay parameter trajectories which makes it suitable for gradient learning. For this purpose, auxiliary differentiable functions were employed for smoothing out the discontinuities of the original model. In particular, a new pulse arises through a continuous "division" of the preceding spike into two spikes while the spike disappearance is implemented by a continuous "fusion" with its predecessor, which

* Research partially supported by projects 1ET100300517 (AS CR), 1M0545 (MŠMT CR), and AV0Z10300504 (ICS AS CR). The author also thanks Petra Vídnerová for her help with computer experiments.

is controlled by the first derivative of the neuron potential and introduces the derivative values already into the computational phase. In addition, the previous spike may simultaneously find itself in the process of its creation or deletion, which leads to a recursive definition and makes the model relatively complex.
2. For our model of smoothly spiking neurons we have derived a **nontrivial back-propagation rule** for computing the gradient of the error function with respect to both the weight and delay parameters which can be used for supervised learning of desired spike trains in corresponding feedforward networks. We have solved the following related technical issues. In order to take also the temporal dimension into account, we have implemented the chain rule for computing the partial derivatives of the composite error function so that each neuron propagates backwards a variable number of partial derivative terms corresponding to different time instants including the second-order derivative terms. Moreover, the recursive definition of the computational dynamics of a single smoothly spiking neuron leads to a complicated recursive formula also for its derivative which was explicitly resolved and transformed into a relatively simple local adaptation rule for a spiking neuron. In addition, the firing times are still defined by an implicit function whose partial derivatives with respect to various variables were calculated using a total differential identity and the implicit function theorem.
3. We have implemented the proposed learning algorithm and performed several **preliminary experiments** which confirm the method functionality and its reasonable time complexity. In particular, the smoothness of spike creation/deletion process was exhibited. Thus the new gradient learning method can naturally cope with multiple spikes whose number changes in time. Experiments with more complicated temporal training patterns remain for further research which this paper invites to participate in.

References

1. Berkovec, K.: Learning in networks of spiking neurons. (in Czech) M.Sc. Thesis, Faculty of Mathematics and Physics, Charles University, Czech Republic (2005)
2. Bohte, S.M., Kok, J.N., La Poutré, H.: Error-backpropagation in temporally encoded networks of spiking neurons. Neurocomputing **48** (1-4) (2002) 17–37
3. Gerstner, W., Kistler, W.M.: Spiking Neuron Models: Single Neurons, Populations, Plasticity. Cambridge University Press, Cambridge, UK (2002)
4. Kasiński, A., Ponulak, F.: Comparison of supervised learning methods for spike time coding in spiking neural networks. International Journal of Applied Mathematics and Computer Science **16** (1) (2006) 101–113
5. Moore, S.C.: Back-propagation in spiking neural networks. M.Sc. Thesis, Department of Computer Science, University of Bath, UK (2002)
6. Schrauwen, B., Van Campenhout, J.: Extending SpikeProp. Proceedings of the IJCNN 2004 IEEE International Joint Conference on Neural Networks, Budapest, Hungary, Vol. 1, IEEE, Piscataway, NJ (2004) 471–475
7. Tiňo, P., Mills, A.J.S.: Learning beyond finite memory in recurrent networks of spiking neurons. Neural Computation **18** (3) (2006) 591–613
8. Xin, J., Embrechts, M.J.: Supervised learning with spiking neuron networks. Proceedings of the IJCNN 2001 IEEE International Joint Conference on Neural Networks, Washington D.C., Vol. 3, IEEE, Piscataway, NJ (2001) 1772–1777