Towards More Biologically Plausible Error-Driven Learning for Artificial Neural Networks

Kristína Malinovská[(✉)], Ľudovít Malinovský, and Igor Farkaš

Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Bratislava, Slovakia
{malinovska,farkas}@fmph.uniba.sk
http://cogsci.fmph.uniba.sk/cnc/

Abstract. Since the standard error backpropagation algorithm for supervised learning was shown biologically implausible, alternative models of training that use only local activation variables have been proposed. In this paper we present a novel algorithm called UBAL, inspired by the GeneRec model. We shortly describe the model and show the performance of the algorithm for XOR and 4-2-4 problems.

Keywords: Error-driven learning · Biological plausibility

In search for an alternative to error backpropagation [5], considered to be biologically implausible [1], O'Reilly proposed the GeneRec model [4]. Instead of propagating error values, neuron activation is propagated in GeneRec bidirectionally. The weight update is based on the difference in the net activation in the minus phase (producing output from input) and the plus phase (desired value is "clamped" on the output layer and the activation spread back to the hidden layer). Building on this principle, we proposed the BAL model [3] for bidirectional heteroassociative mappings, but failed to reach 100% convergence on the canonical 4-2-4 encoder task despite extensive experimental tuning [2]. As an improvement, we propose the Universal Bidirectional Activation-based Learning (UBAL) algorithm with additional learning parameters enabling the model to perform also unidirectional association tasks such as classification. As GeneRec, our model uses activation state differences, but with separate weight matrices M and W for each direction of activation flow. The activation is propagated in four phases (Fig. 1).

As outlined in Table 1, in the forward prediction phase FP, the input is presented to layer p and the activation spreads to layer q and vice versa for the backward prediction BP. Additionally, there are echo activation phases (FE and BE) in which the network's previous outputs q^{FP} and p^{BP} are echoed back to p and q through weights M and W, respectively.

The learning rule in Eqs. 1 and 2 takes as inputs intermediate terms t (target) and e (estimate) from Table 2.

V. Kůrková et al. (Eds.): ICANN 2018, LNCS 11141, pp. 228–231, 2018.
https://doi.org/10.1007/978-3-030-01424-7_23

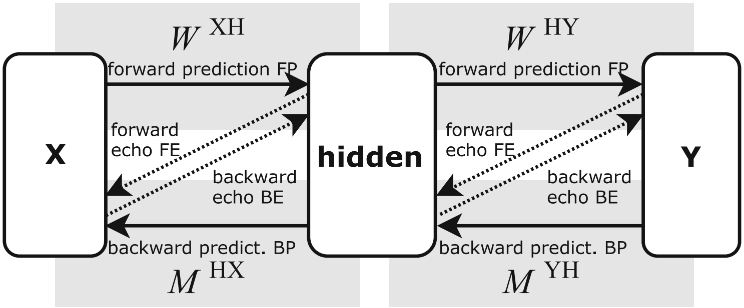

Fig. 1. Activation propagation in a network with input-output layers **x** and **y** and one hidden layer.

Table 1. Activation propagation rules, p and q denote two layers of the network connected by weight matrices W and M. Symbols b and d denote the biases and σ stands for the standard logistic activation function.

Direction and phase	Term	Value
Forward prediction	q_j^{FP}	$\sigma(\sum_i w_{ij} p_i^{\mathrm{FP}} + b_j)$
Forward echo	p_i^{FE}	$\sigma(\sum_j m_{ji} q_j^{\mathrm{FP}} + d_i)$
Backward prediction	p_i^{BP}	$\sigma(\sum_j m_{ji} q_j^{\mathrm{BP}} + d_i)$
Backward echo	q_j^{BE}	$\sigma(\sum_i w_{ij} p_i^{\mathrm{BP}} + b_j)$

Table 2. Definition of terms used in the learning rule.

Term name	Term	Value
Forward target	t_j^{F}	$\beta_q^{\mathrm{F}} q_j^{\mathrm{FP}} + (1 - \beta_q^{\mathrm{F}}) q_j^{\mathrm{BP}}$
Forward estimate	e_j^{F}	$\gamma_q^{\mathrm{F}} q_j^{\mathrm{FP}} + (1 - \gamma_q^{\mathrm{F}}) q_j^{\mathrm{BE}}$
Backward target	t_i^{B}	$\beta_p^{\mathrm{B}} p_i^{\mathrm{BP}} + (1 - \beta_p^{\mathrm{B}}) p_i^{\mathrm{FP}}$
Backward estimate	e_i^{B}	$\gamma_p^{\mathrm{B}} p_i^{\mathrm{BP}} + (1 - \gamma_p^{\mathrm{B}}) p_i^{\mathrm{FE}}$

$$\Delta w_{ij} = \lambda\, t_i^{\mathrm{B}} (t_j^{\mathrm{F}} - e_j^{\mathrm{F}}) \tag{1}$$

$$\Delta m_{ij} = \lambda\, t_j^{\mathrm{F}} (t_i^{\mathrm{B}} - e_i^{\mathrm{B}}) \tag{2}$$

The learning rate λ and parameters β (target prediction strength) and γ (estimate prediction strength) used in the learning rule terms in Table 2 drive the network learning. Depending on their values the network can accomplish different tasks.

In Fig. 2 we present results from experiments with the 4-2-4 encoder indicating that using a reasonable learning rate the network always converges to a solution. Unlike its predecessor BAL, given a certain parameter setup (Table 3), UBAL converges in the XOR task as shown in Fig. 3. Preliminary results from further experiments suggest that UBAL could get us closer towards a biologically plausible alternative to error backpropagation.

Table 3. Parameters β a γ in our experiments, $\beta^B = 1 - \beta^F$.

	4-2-4 Encoder	XOR
	X — H — Y	X — H — Y
β^{F}	1.0 – 0.5 – 0.0	0.01 – 1.0 – 0.0
γ^{F}	0.5 – 0.5	0.0 – 0.0
γ^{B}	0.5 – 0.5	0.0 – 0.0

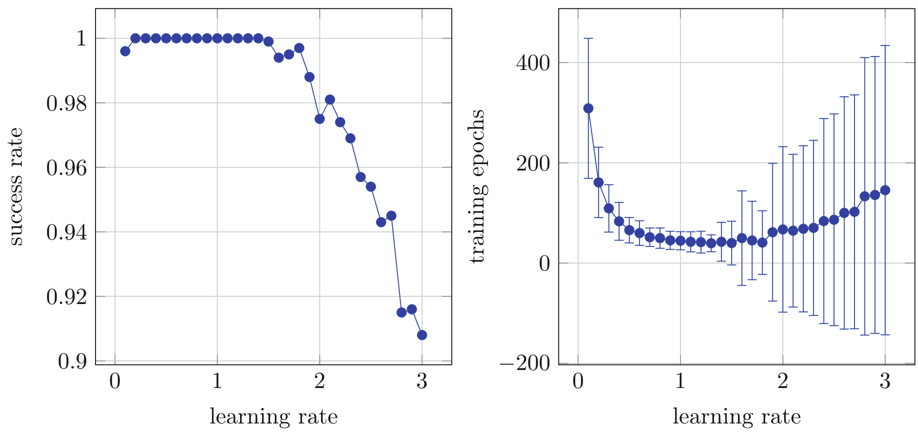

Fig. 2. Results from 4-2-4 encoder experiments with varying λ (1000 nets). Success rate indicates how many networks were able to learn the task with 100% accuracy.

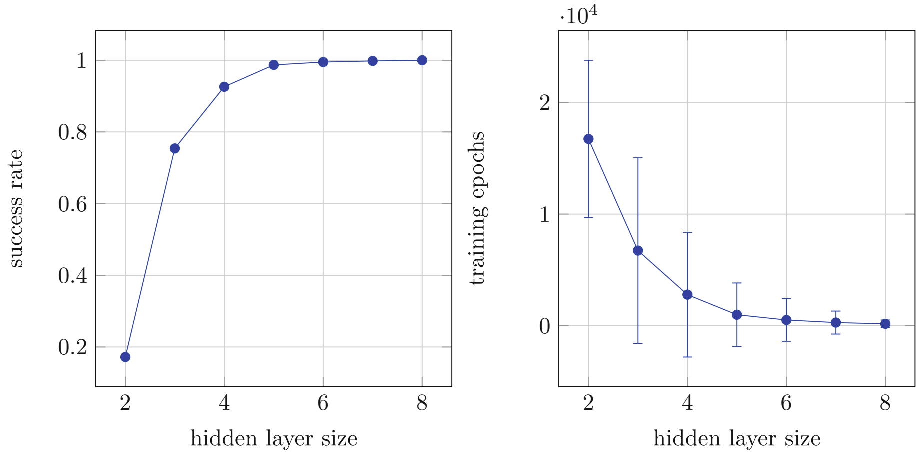

Fig. 3. Results from XOR experiments with varying hidden layer size (1000 nets) and $\lambda = 0.2$. Maximum training epochs: 20000.

Acknowledgment. This work was supported by grants VEGA 1/0796/18 and KEGA 017UK-4/2016.

References

1. Crick, F.: The recent excitement about neural networks. Nature **337**(6203), 129–132 (1989)
2. Csiba, P., Farkaš, I.: Computational analysis of the bidirectional activation-based learning in autoencoder task. In: International Joint Conference on Neural Networks (IJCNN), pp. 1–6. IEEE (2015)
3. Farkaš, I., Rebrová, K.: Bidirectional activation-based neural network learning algorithm. In: Mladenov, V., Koprinkova-Hristova, P., Palm, G., Villa, A.E.P., Appollini, B., Kasabov, N. (eds.) ICANN 2013. LNCS, vol. 8131, pp. 154–161. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40728-4_20
4. O'Reilly, R.: Biologically plausible error-driven learning using local activation differences: the generalized recirculation algorithm. Neural Comput. **8**(5), 895–938 (1996)
5. Rumelhart, D., Hinton, G., Williams, R.: Learning Internal Representations by Error Propagation, pp. 318–362. no. 1. The MIT Press, Cambridge (1986)