Derandomized Evolution Strategies in Computational Neuroscience

Hubert R. Dinse Michael Hüsken Christian Igel Christian Klaes
Marc Nunkesser Stefan Schneider Jan Wiemer
Institut für Neuroinformatik, Lehrstuhl für theoretische Biologie
Ruhr-Universität Bochum
44780 Bochum, Germany
{christian.igel, stefan.schneider}@neuroinformatik.ruhr-uni-bochum.de

1 INTRODUCTION

We propose evolutionary "analysis by synthesis" guided by neurobiological knowledge as a powerful tool in computational neuroscience. The challenge is to force artificial evolution to favor solutions that are reasonable from the biological point of view. Such solutions are only likely to evolve if as much neurobiological knowledge as possible is used in the design process. This can be achieved by providing sufficient experimental data to evaluate the evolved systems and by a considerate choice of the basic structure. Additional knowledge can be incorporated into the fitness function and in constraints that ensure biological plausibility.

The first step towards the design of biological systems is parameter adaptation of predefined models. In the following, we present some recent applications of the CMA-ES, an elaborated, derandomized evolution strategy [4, 5], in the field of parameter optimization of neurobiological models.

2 ADAPTATION OF MACROSCOPIC CORTICAL MODELS

Dynamic neural fields (NFs) have been introduced as macroscopic models of cortical information processing [11, 1] and there is a growing interest in using them in biological modeling and for technical applications. However, the problem arises how to adjust the parameters of NFs for a given task. In [6] we applied gradient-based and evolutionary optimization as well as a simple hybrid algorithm, which switches from evolutionary to gradient-based optimization when the error drops below a threshold, for parameter adaptation of a one-dimensional single-layered NF [1]. Here we extend our investigation to gradient-based and evolutionary optimization of two-layered NFs [3, 8]:

$$\tau_u \dot{u}_x(t) = -u_x(t) + \sum_{x'=1}^{n} w_s(x-x')s_{x'}(t) + h + f_s[u_x(t)]\left[\sum_{x'=1}^{n} w_u(x-x')\, f_u[u_{x'}(t)] - v_x(t)\right] \quad (1)$$

$$\tau_v \dot{v}_x(t) = -v_x(t) + \sum_{x'=1}^{n} w_v(x-x')\, f_u[u_{x'}(t)] \ , \quad (2)$$

where $u_x(t)$ and $v_x(t)$ describe the average membrane potential at time t of a neuron located at site $x \in \{1,\ldots,n\}$ of the inhibitory and excitatory layer, respectively. The external input into the field is denoted by $s_x(t)$. The resting potential is given by h and the membrane time constants by τ_u and τ_v. The strengths of the couplings are determined by $w_u(.)$, $w_v(.)$, and $w_s(.)$, which are bell-shaped functions with two parameters each. The firing rate function $f_u[.]$ and the shunting term $f_s[.]$ are scaled sigmoidal functions with three parameters each.

The task was to reproduce the spatio-temporal pattern shown in Fig. 1. The error function was the mean-squared error. The derivatives of the error with respect to the 15 parameters of the field can be determined by solving the so called variation system forward in time, cf. [6].

The results were similar to the ones reported for the single-layered NF. The adaptation of the dynamical system turned out to be a difficult optimization task, where many trials led to undesired local minima. In our experiments, the evolution strategy was considerably slower than the gradient-based optimization using the BFGS-algorithm, although in NFs the gradient cannot be calculated as efficiently as in the case of standard recurrent neural networks, because methods like computing the gradient backward in time do not

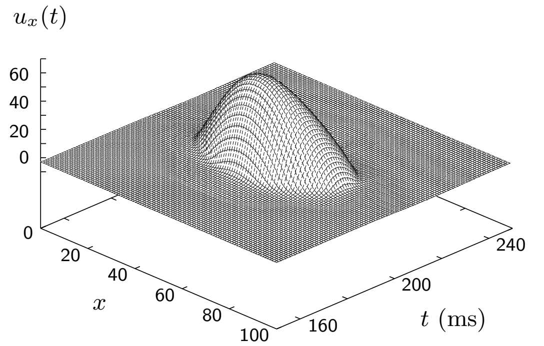

Figure 1: Artificial target data for the adaptation of the two-layered neural field. The spatio-temporal pattern was generated by the field Eqs. (1) and (2).

reduce the complexity [6]. However, evolution found the optimal solution more often in some scenarios.

3 ADAPTION OF A MICROSCOPIC CORTICAL MODEL

In order to reveal mechanisms of cortical temporal stimulus processing, we carried out a neurobiological experiment, set up a model, and fitted the model to the experimental data by means of the CMA-ES. We employed the CMA-ES for searching the model parameters for the following reasons: First, adjusting the parameters manually, as it is common in computational neuroscience, is very time-consuming in our case. Second, we wanted to carry out a global search for good solutions, i.e., those with a small fitting error.

3.1 Neurobiological experiment

We examined how temporal sequences of sensory stimuli are processed by the early sensory cortex [12]. To this end, we recorded the neural stimulus-response activity in the somatosensory cortex of rats. Trains of three tactile stimuli of variable interstimulus intervals (ISIs) were applied to the hind paw of the rat.

3.2 Model

We hypothesized that the experimentally observed effects were caused by cortical feedback inhibition and depletion of the transmitter of the involved synapses. Our model consists of an excitatory and an inhibitory neuron, each representing a local cortical neuron population [10]. We also included a depletion mechanism for the transmitter of the sensory synapses [2]. The model is described by the following system of coupled ordinary differential equations:

$$\tau_e \frac{\mathrm{d}x_e(t)}{\mathrm{d}t} = -x_e(t) + f[p\, w_{\mathrm{ff}}\, x_s(t) - w_{\mathrm{int}}\, x_i(t)] \quad (3)$$

$$\tau_i \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = -x_i(t) + f[w_{\mathrm{int}}\, x_e(t)] \quad (4)$$

$$\tau_e \frac{\mathrm{d}x_s(t)}{\mathrm{d}t} = -x_s(t) + F_s(t) \quad (5)$$

$$\frac{\mathrm{d}p(t)}{\mathrm{d}t} = -c_t\, p(t)\, F_s(t) + \frac{1-p(t)}{\tau_d} \;. \quad (6)$$

x_e, x_i and x_s are the synaptic drives [9] of the excitatory, the inhibitory, and the sensory neuron. τ_e and τ_i are time constants for excitation and inhibition. w_{int} denotes the coupling strength between excitation and inhibition. F_s is the firing rate of the sensory neuron and models the stimulus. w_{ff} denotes the coupling strength between the sensory neuron and the excitatory cortical neuron. Synaptic depletion is described by the fraction of available transmitter, p, the depletion factor c_t, and the transmitter recovery time constant τ_d. $f[.]$ is a nonlinear firing rate function (linear threshold function). The three maxima of the excitatory firing rate $f[x_e(t)]$ are read out as the stimulus responses corresponding to the experimentally observed responses (a_1, a_2, and a_3 in Fig. 2).

3.3 Adaptation of the model

We applied the CMA-ES for determining the 4 model parameters τ_i, τ_d, w_{int}, and c_t. These parameterize the mechanisms we hypothesized to be responsible for the observed experimental effects. The other parameters were set to plausible values beforehand and kept fixed. The task was to minimize the mean-squared error between experimental responses and model responses over all ISIs.

We carried out a series of CMA-ES trials yielding several local optima of different quality. This indicates the difficulty of the optimization task. One solution fits the experimental data very well (Fig. 2). Furthermore, its parameter values are biologically very plausible. Most of the other solutions fit the experimental data clearly worse. Their parameter values are less plausible than those of the best solution.

Therefore, the CMA-ES results enabled us to conclude that our modeling approach is reasonable.

4 DISCUSSION

Evolutionary algorithms have proven to be beneficial for the optimization of models in computational neuroscience. They can be applied even if the quality

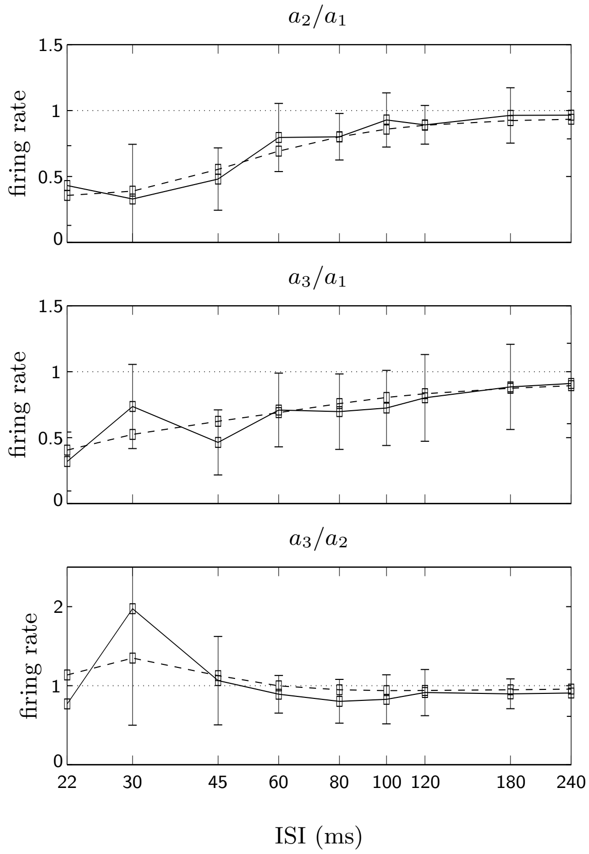

Figure 2: Experimental data (solid line) and model data of the best solution (dashed line). a_1, a_2 and a_3 are the neural response strengths to the 1st, 2nd, and 3rd stimulus.

function is nondifferentiable, discontinuous, and noisy. Further, they allow for the incorporation of expert knowledge in several ways. Even if the fitness function is differentiable, there are reasons to prefer evolutionary optimization, as the gradient may be very difficult to compute. However, if the gradient can be derived with reasonable effort, then it should be exploited, e.g., in a hybrid algorithm as in [6].

Parameter adaptation of nonlinear dynamical systems has turned out to be a difficult optimization task, where—at least in our examples—the optimization algorithms tend to get trapped in local minima; bifurcations of the dynamics exacerbate the search process. These difficulties can partly be overcome by incorporating additional expert knowledge, as in [7], where a macroscopic model is fit to neurobiological data.

References

[1] S. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27:127–132, 1977.

[2] P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press, Cambridge, 2001.

[3] W. Erlhagen, A. Bastian, D. Jancke, A. Riehle, and G. Schöner. The distribution of neural population activation as a tool to study interaction and integration in cortical representations. *Journal of Neuroscience Methods*, 94(1):53–66, 1999.

[4] N. Hansen and A. Ostermeier. Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The $(\mu/\mu, \lambda)$-CMA-ES. In *5th European Congress on Intelligent Techniques and Soft Computing (EUFIT'97)*, pages 650–654, Aachen, Germany, 1997. Verlag Mainz, Wissenschaftsverlag.

[5] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.

[6] C. Igel, W. Erlhagen, and D. Jancke. Optimization of dynamic neural fields. *Neurocomputing*, 36(1-4):225–233, 2001.

[7] C. Igel, W. von Seelen, W. Erlhagen, and D. Jancke. Evolving field models for inhibition effects in early vision. In J. M. Bower, editor, *Computational Neuroscience: Trends in Research 2001*. Elsevier Science, 2002. In press.

[8] D. Jancke, W. Erlhagen, H. R. Dinse, A. C. Akhavan, M. Giese, A. Steinhage, and G. Schöner. Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *Journal of Neuroscience*, 19(20):9016–9028, 1999.

[9] D. J. Pinto, J. C. Brumberg, D. J. Simons, and G. B. Ermentrout. A quantitative population model of whisker barrels: Re-examining the Wilson-Cowan equations. *Journal of Computational Neuroscience*, 3:247–264, 1996.

[10] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Kybernetik*, 12:1–24, 1972.

[11] H. R. Wilson and J. D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13:55–80, 1973.

[12] R. Zepka, J. Wiemer, S. Schneider, S. Cords, and H. R. Dinse. Plasticity of short-term plasticity: Cortical neuron responses to tactile stimuli of variable isi recorded in somatosensory cortex of adult rats before and after intracortical microstimulation (ICMS). In *Society for Neuroscience Abstracts 27*, 2001.