Computational Neurogenetic Modeling: Integration of Spiking Neural Networks, Gene Networks, and Signal Processing Techniques

Nikola Kasabov, Lubica Benuskova, Simei Gomes Wysoski©

Knowledge Engineering and Discovery Research Institute,
Auckland University of Technology, 581-585 Great South Rd,
Auckland, New Zealand
{nkasabov,lbenusko,swysoski}@aut.ac.nz
http:www.kedri.info

Abstract. The paper presents a theory and a new generic computational model of a biologically plausible artificial neural network (ANN), the dynamics of which is influenced by the dynamics of internal gene regulatory network (GRN). We call this model a "computational neurogenetic model" (CNGM) and this new area of research Computational Neurogenetics. We aim at developing a novel computational modeling paradigm that can potentially bring original insights into how genes and their interactions influence the function of brain neural networks in normal and diseased states. In the proposed model, FFT and spectral characteristics of the ANN output are analyzed and compared with the brain EEG signal. The model includes a large set of biologically plausible parameters and interactions related to genes/proteins and spiking neuronal activities. These parameters are optimized, based on targeted EEG data, using genetic algorithm (GA). Open questions and future directions are outlined.

1 Introduction

We introduce a novel computational approach to brain neural network modeling that integrates ANN with an internal dynamic GRN. Interaction of genes in model neurons affects the dynamics of the whole ANN through neuronal parameters, which are no longer constant, but change as a function of gene expression. Through optimization of the GRN, initial gene/protein expression values and ANN parameters, particular target states of the neural network operation can be achieved. It is illustrated by means of a simple neurogenetic model of a spiking neural network (SNN). The behavior of SNN is evaluated by means of the local field potential (LFP), thus making it possible to attempt modeling the role of genes in different brain states, where EEG data is available to test the model. We use the standard FFT signal processing technique to evaluate the SNN output and compare with real human EEG data. For the objective of this work, we consider the time-frequency resolution reached with the FFT to be sufficient. However, should higher accuracy be critical, Wavelet Transform, which considers both time and frequency resolution, could be used instead. Broader theoretical

W. Duch et al. (Eds.): ICANN 2005, LNCS 3697, pp. 509-514, 2005.

and biological background of CNGM construction is given in [1]. A simpler linear version of an internal GRN with preliminary results on epilepsy modeling can be found in [2]. In this paper we (1) introduce and simulate a more realistic nonlinear model of GRN, (2) present a list of real proteins/genes that are involved in CNGM, (3) compare the CNGM performance to real human EEG data using the same signal processing technique, (4) suggest an optimization procedure to obtain a CNGM with parameters leading to modeling of the real EEG signal.

2 A general CNGM and its optimization via evolution

In general, we consider two sets of genes – a set G_{gen} that relates to general cell functions and a set G_{spec} that defines specific neuronal information-processing functions (receptors, ion channels, etc.). The two sets form together a set $\mathbf{G}=\{G_1, G_2, \ldots, G_n\}$. We assume that the expression level of each gene $g_j(t+\Delta t')$ is a nonlinear function of expression levels of all the genes in **G**, inspired by discrete models from [3], [4]:

$$g_j(t+\Delta t') = \sigma\left(\sum_{k=1}^{n} w_{jk}\, g_k(t)\right) \quad (\mathbf{1})$$

We work with normalized gene expression values in the interval (0, 1). The coefficients $w_{ij} \in (-5,5)$ are the elements of the square matrix **W** of gene interaction weights. Initial values of gene expressions are small random values, i.e. $g_j(0) \in (0, 0.1)$.

In the current model we assume that: (1) one protein is coded by one gene; (2) relationship between the protein level and the gene expression level is nonlinear; (3) protein levels lie between the minimal and maximal values. Thus, the protein level $p_j(t+\Delta t)$ is expressed by

$$p_j(t+\Delta t) = (p_j^{\max} - p_j^{\min})\,\sigma\left(\sum_{k=1}^{n} w_{jk}\, g_k(t)\right) + p_j^{\min} \quad (\mathbf{2})$$

The delay $\Delta t < \Delta t'$ corresponds to the delay caused by the gene transcription, mRNA translation into proteins and posttranslational protein modifications [9]. Delay $\Delta t'$ includes also the delay caused by gene transcription regulation by transcription factors.

The GRN model from equations (1) and (2) is a general one and can be integrated with any ANN model into a CNGM. Unfortunately the model requires many parameters to be either known in advance or optimized during a model simulation. In the presented experiments we have made several simplifying assumptions:

1. Each neuron has the same GRN, i.e. the same genes and the same interaction gene matrix **W**.
2. Each GRN starts from the same initial values of gene expressions.
3. There is no feedback from neuronal activity or any other external factors to gene expression levels or protein levels.
4. Delays Δt are the same for all proteins and reflect equal time points of gathering protein expression data.

We have integrated the above GRN model with the SNN illustrated in Fig. 1. Our spiking neuron model is based on the Spike Response Model [5], with excitation and

inhibition having both fast and slow components [6], [7] both expressed as double exponentials with amplitudes and the rise and decay time constants.

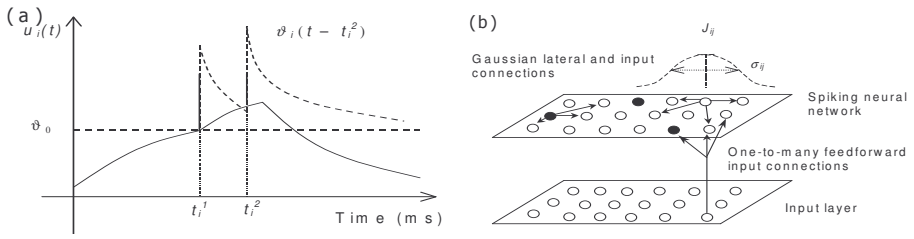

Fig. 1. (a) Spiking neuron model. When the membrane potential $u_i(t)$ of the i^{th} spiking neuron reaches the firing threshold $\vartheta_i(t)$ at time t^k_i, the neuron fires an output spike. $\vartheta_i(t)$ rises after each output spike and decays back to the resting value ϑ_0. (b) The SNN architecture. About 10–20% of N = 120 neurons are inhibitory neurons that are randomly positioned on the grid (filled circles). External input is random with average frequency between 10–20 Hz

Table 1. Neuronal parameters and their corresponding proteins (receptors/ion channels)

Neuron's parameter P_j	Relevant protein p_j
Amplitude and time constants of:	
Fast excitation	AMPAR
Slow excitation	NMDAR
Fast inhibition	GABRA
Slow inhibition	GABRB
Firing threshold and its decay time constant	SCN and/or KCN and/or CLC

Neuronal parameters and their correspondence to particular proteins are summarized in Table 1[1]. Several parameters (amplitude, time constants) are linked to one protein. However their initial values in equation (3) will be different. Relevant protein levels are directly related to neuronal parameter values P_J such that

$$P_j(t) = P_j(0)\, p_j(t) \tag{3}$$

where $P_j(0)$ is the initial value of the neuronal parameter at time $t = 0$. Moreover, besides the genes coding for the proteins mentioned above, we include in our GRN nine more genes that are not directly linked to neuronal information-processing parameters. These genes are: c-jun, mGLuR3, Jerky, BDNF, FGF-2, IGF-I, GALR1, NOS, S100beta. We have included them for later modeling of some diseases.

We want to achieve a desired SNN output through optimization of the model 294 parameters (we are optimizing also the connectivity and input frequency to the SNN). We evaluate the LFP of the SNN, defined as LFP = $(1/N)\Sigma\, u_i(t)$, by means of FFT in order to compare the SNN output with the EEG signal analyzed in the same way. It has been shown that brain LFPs in principle have the same spectral characteristics as

[1] Abbreviations: AMPAR = (amino- methylisoxazole- propionic acid) AMPA receptor, NMDAR = (*N*-methyl-D-aspartate acid) NMDA receptor, GABRA = (gamma-aminobutyric acid) GABA receptor A, GABRB = GABA receptor B, SCN = Sodium voltage-gated channel, KCN = kalium (potassium) voltage-gated channel, CLC = chloride channel.

EEG [8]. Because the updating time for SNN dynamics is inherently 1ms, just for computational reasons, we will employ the delays Δt in equation (2) being equal to just 1s instead of minutes or tens of minutes [9]. In order to find an optimal GRN within the SNN model so that the frequency characteristics of the LFP of the SNN model are similar to the brain EEG characteristics, we use the following procedure:

1. Generate a population of CNGMs, each with randomly generated values of coefficients for the GRN matrix W, initial gene expression values $g(0)$, initial values of SNN parameters $P(0)$, and different connectivity;
2. Run each SNN over a period of time T and record the LFP;
3. Calculate the spectral characteristics of the LFP using FFT;
4. Compare the spectral characteristics of SNN LFP to the characteristics of the target EEG signal. Evaluate the closeness of the LFP signal for each SNN to the target EEG signal characteristics. Proceed further according to the standard GA algorithm to possibly find a SNN model that matches the EEG spectral characteristics better than previous solutions;
5. Repeat steps 1 to 4 until the desired GRN and SNN model behavior is obtained;
6. Analyze the GRN and the SNN parameters for significant gene patterns that cause the SNN model behavior.

3 Simulation and results

First, we present the results of analysis performed on real human interictal EEG data obtained with permission from [10]. Fig. 2 shows the brain EEG signal, its FFT power spectrum and evolution of the relative intensity ratios (RIRs) for different clinically relevant sub-bands over time. These sub-bands are: delta (0.5-3.5 Hz), theta (3.5-7.5 Hz), alpha (7.5-12.5 Hz), beta 1 (12.5-18 Hz), beta 2 (18-30 Hz), gamma (above 30 Hz). Each point depicts the RIR over the previous 1s.

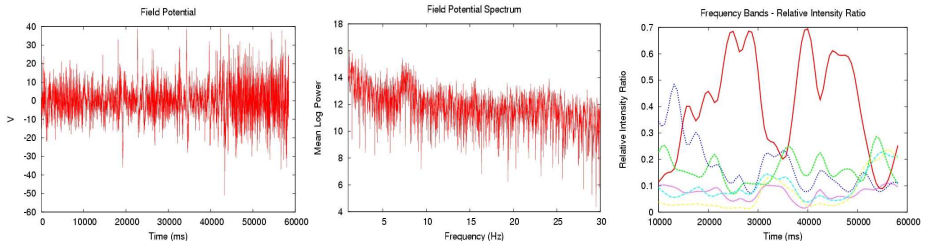

Fig. 2. a) Human interictal EEG Signal; b) classical FFT analysis of the EEG signal, sampling rate is 256 Hz; c) temporal evolution of RIRs for the clinically relevant frequency sub-bands for the EEG signal. The dominant sub-band is delta (0.5-3.5 Hz)

We calculated the average RIRs over the whole time of simulation (i.e., T = 1 min) and used this vector of values as a fitness function for our GA. After 50 generations with 6 solutions in each population we obtained the following result for the best solution, illustrated in Fig. 3. Solutions for reproduction were being chosen according to the roulette rule and the crossover between parameter values was performed as an arithmetic average of the parent values. We performed the same FFT analysis as for

the real EEG data with the Min/Max frequency = 0.1 / 50 Hz. This particular SNN had an evolved GRN with only 5 genes out of 16 changing periodically their expression values (s100beta, GABRB, GABRA, mGLuR3, c-jun) and all other genes having constant expression values (see e.g. Fig. 4), either minimal or maximal.

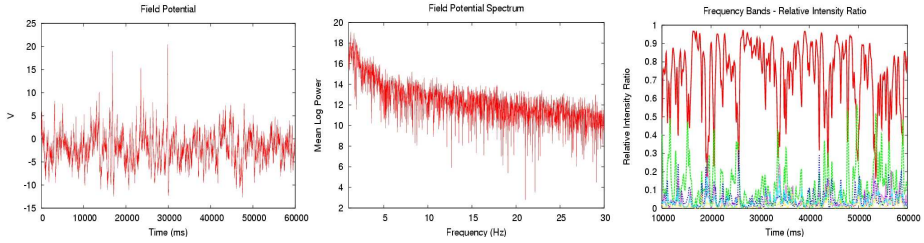

Fig. 3. a) Local field potential of the SNN with GRN; b) classical FFT analysis of the SNN LFP, sampling rate is 1000 Hz; c) temporal evolution of RIRs for clinically relevant frequency sub-bands for the LFP. The dominant sub-band is again delta (0.5-3.5 Hz)

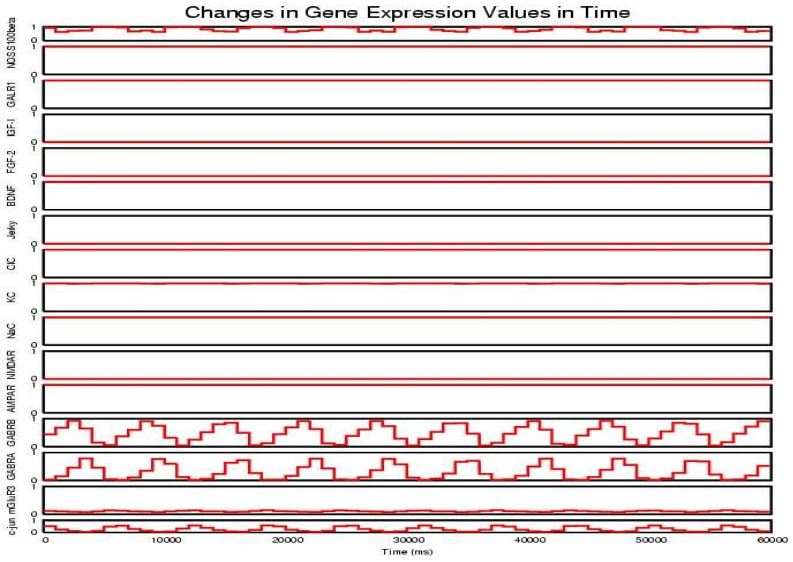

Fig. 4. Changes in gene expression values over time in the model GRN

4 Discussion

Our preliminary results show that the same signal processing techniques can be used for the analysis of both the simulated LFP of the SNN CNGM and the real EEG data to yield conclusions about the SNN behavior and to evaluate the CNGM at a gross level. With respect to our neurogenetic approach we must emphasize that it is still in an early developmental stage and the experiments assume many simplifications. In particular, we would have to deal with the delays in equation (2) more realistically to

be able to draw any conclusions about real data and real GRNs. The LFP obtained from our simplified model SNN is of course not exactly the same as the real EEG, which is a sum of many LFPs. However LFP's spectral characteristics are very similar to the real EEG data, even in this preliminary example. Based on our preliminary experimentation, we have come to the conclusion that many gene dynamics, i.e. many interaction matrices **W**s that produce various gene dynamics (e.g., constant, periodic, quasiperiodic, chaotic) can lead to very similar SNN LFPs. In our future work, we want to explore statistics of plausible **W**s more thoroughly and compare it with biological data to draw any conclusions about underlying GRNs. Further research questions are: How many GRNs would lead to similar LFPs and what do they have in common? How to use CNGM to model gene mutation effects? How to use CNGM to predict drug effects? And finally, how to use CNGM for the improvement of individual brain functions, such as memory and learning?

Acknowledgments
Supported by the NERF grant AUTX02001 funded by FRST, and by KEDRI funds.

References

1. Kasabov, N., Benuskova, L.: Computational Neurogenetics. Journal of Computational and Theoretical Nanoscience, Vol. 1. (2004) 47-61
2. Kasabov, N., Benuskova, L., Wysoski, S.G.: Computational neurogenetic modelling: gene networks within neural networks. Proc. IEEE Intl. Joint Conf. Neural Networks, Vol. 2. IEEE Press, Budapest (2004) 1203-1208
3. Weaver, D.C., Workman, C.T., Stormo, G.D.: Modeling regulatory networks with weight matrices. Proc. Pacific Symp. Biocomputing - Hawai, Vol. 4. World Scientific Publ. Co, Singapore (1999) 112-123
4. Wessels, L.F.A., vanSomeren, E.P., Reinders, M.J.T.: A comparison of genetic network models. Proc. Pacific Symp. Biocomputing, Vol. 6. (2001) 508-519
5. Gerstner, W., Kistler, W.M.: Spiking Neuron Models. Cambridge Univ. Press, Cambridge-MA (2002)
6. Destexhe, A.: Spike-and-wave oscillations based on the properties of $GABA_B$ receptors. J. Neuroscience, Vol. 18. (1998) 9099-9111
7. Kleppe, I.C., Robinson, H.P.C.: Determining the activation time course of synaptic AMPA receptors from openings of colocalized NMDA receptors. Biophys. J., Vol. 77. (1999) 1418-1427
8. Destexhe, A. Contreras, D., Steriade, M.: Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. J. Neuroscience, Vol. 19. (1999) 4595-4608
9. Lodish, H. Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J.: Molecular Cell Biology. 4th edn. W.H. Freeman & Co, New York (2000)
10. Quiroga, R.Q.: Dataset #3: Tonic-clonic (Grand Mal) seizures. www.vis.caltech.edu/~rodri/data.htm (1998)