TECHNICAL NOTE

Music Neurotechnology for Sound Synthesis: Sound Synthesis with Spiking Neuronal Networks

Eduardo R. Miranda and John Matthias

ABSTRACT

Music neurotechnology is a new research area emerging at the crossroads of neurobiology, engineering sciences and music. Examples of ongoing research into this new area include the development of brain-computer interfaces to control music systems and systems for automatic classification of sounds informed by the neurobiology of the human auditory apparatus. The authors introduce *neurogranular sampling*, a new sound synthesis technique based on spiking neuronal networks (SNN). They have implemented a neurogranular sampler using the SNN model developed by Izhikevich, which reproduces the spiking and bursting behavior of known types of cortical neurons. The neurogranular sampler works by taking short segments (or sound grains) from sound files and triggering them when any of the neurons fire.

Neurobiology has given inspiration to computer music technology since the invention of the perceptron in the late 1950s by Frank Rosenblatt and the subsequent development of a variety of artificial neural network architectures, most of which have been used in computer-music applications, with various degrees of success. Essentially a perceptron is an artificial neuron and is the basic building block of artificial neural networks. It receives one or more inputs and produces one or more outputs, each of which is a simple non-linear function of the sum of its inputs. A neural network creates connections between artificial neurons, and the organization and weights of these connections determine the overall output of the network. Neural networks are normally used in engineering to predict events and recognize patterns. However, the types of neuronal models discussed in this paper are sophisticated models designed to help us understand how the brain works rather than to be used in engineering applications.

We are interested in exploring the behavior of computational models of brain functioning to make music. We find their ability to generate very complex biological-like behavior from the specification of relatively simple parametric variables compelling and inspiring. They allow for the design of complex sound generators and sequencers controlled by a handful of parameters. We envisage the design of new musical instruments based on such models.

Many recent advances in the neurosciences, especially in computational neuroscience, have led to a deeper understanding of the behavior of individual and neuronal networks (large groups of biological neurons) [1,2], and we can now begin to apply biologically informed neuronal functional paradigms to problems of design and control [3,4]. For instance, such technology may lead to the development of autonomous machines capable of adapting themselves to different conditions, mimicking the ability of the brain to modify its own internal representations of the world. We have coined the term *music neurotechnology* to refer to a new research area that is emerging at the crossroads of neurobiology, engineering sciences and music. Examples of ongoing research into music neurotechnology include the development of brain-computer interfaces to control music systems [5–7] and systems for automatic classification of sounds informed by the neurobiology of the human auditory apparatus [8], to cite but two. This paper reports on the latest results of our ongoing research into music neurotechnology for sound synthesis, with a view to developing new musical instruments based on neuronal functional paradigms. We introduce the *neurogranular sampler*, which uses spiking neuronal network (SNN) models to control the triggering of sound grains taken from a given sampled sound.

Eduardo R. Miranda (composer, computer scientist), Interdisciplinary Centre for Computer Music Research (ICCMR), University of Plymouth, Plymouth, PL4 8AA, U.K. E-mail: <eduardo.miranda@plymouth.ac.uk>.

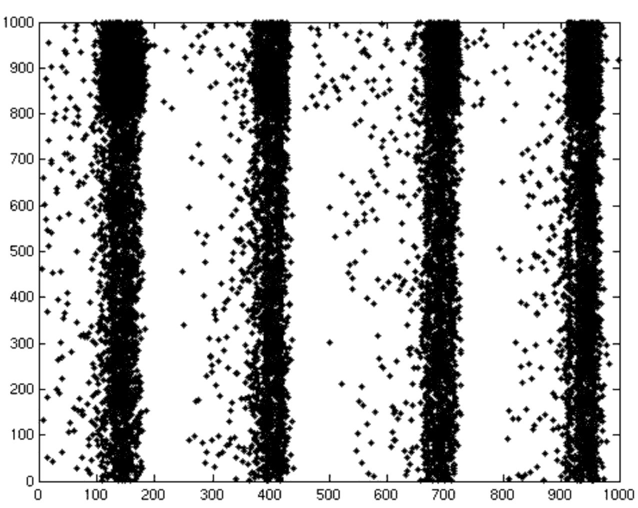
John Matthias (composer, musician, physicist), Interdisciplinary Centre for Computer Music Research (ICCMR), University of Plymouth, Plymouth, PL4 8AA, U.K. E-mail: <john.matthias@plymouth.ac.uk>.

SPIKING NEURONAL NETWORKS

In order to understand the functioning of the neurogranular sampler, one needs to understand the basics of spiking neuronal networks.

A neuron can be thought of as a cell that fires a traveling, spiking signal to connected neurons when the voltage on its membrane exceeds a certain threshold voltage. A neuron receiving several spikes simultaneously (or within a very small

Fig. 1. An example of collective firing behavior. (© Eduardo Miranda) Neuron numbers are plotted (y-axis) against time (x-axis) for a simulation of 1,000 neurons over a period of 1 second. Each dot represents a firing event.



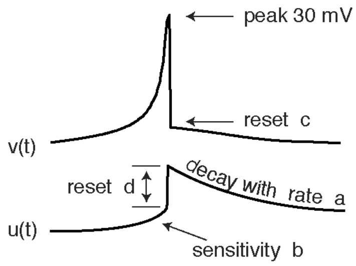


Fig. 2. Known types of neurons correspond to different values of the parameters *a*, *b*, *c* and *d* in Izhikevich's model. (Screen grab © Eugene Izhikevich. Electronic version of this figure and reproduction permissions are freely available at <www.izhikevich.com>.)

time-window) is likely to have its voltage pushed beyond the threshold level and will in turn send spike signals to its connected neighbors. Furthermore, synapses that cause spiking signals tend to become potentiated and those that do not cause spiking signals become depleted, a phenomenon known as *synaptic plasticity*. There are also many varieties of spiking behavior, such as regular spiking and bursting (short bursts containing many spikes very close together in time), to cite but two. Millions of adaptive, interconnected neurons produce very rich behavior, especially on a collective level of description, and patterns of firing regularly occur in large groups of neurons [9]. Figure 1 shows an example of such collective firing behavior, taken from a computer simulation of a group of 1,000 coupled spiking neurons.

The neurons in Fig. 1 are numbered on the y-axis (with neuron 1 at the bottom and neuron 1,000 at the top); time, which runs from zero to 1,000 milliseconds (or 1 second), is on the x-axis. This is therefore a simulation of the activity of this group of 1,000 artificial neurons over a period of 1 second. Every time a neuron fires, a dot is placed on the graph at the appropriate time on a line horizontally drawn from that particular neuron. The dots on the graph can thus be regarded as firing events. In the particular graph shown, because of the plasticity of the neuronal connections, many of the events are centered in four bands, which appear as a pulse (or wave) of spiking events.

The spiking events are indeterminate. They are not predictable in advance but are certainly not random, and, as is the case with the scenario shown in Fig. 1, can be closely correlated. A rhythmic pattern, such as the one pictured in Fig. 1, is connected with the polychronous firing of a particular group of neurons in which the firing of a particular neuron generates a sequence of events. These events stimulate a large number of neurons, which form a closed group in which the connections between the neurons are reinforced through repeated firing of the first neuron. The group of neurons fire not in synchrony but with polychrony. It is the rhythmic nature of these sequences of firing events that have become of great interest to us. Indeed, this phenomenon has been suggestively referred to as *cortical songs* by Ikegaya and colleagues [10].

In this paper, we focus on a spiking neural network (SNN) model developed by Izhikevich [11], which reproduces spiking and bursting behavior of known types of cortical neurons. The model contains N neurons, each of which is described by two dimensionless variables v_i and u_i, where v_i represents the membrane potential (membrane potential is the difference in voltage between the interior and exterior of a cell) of the i^{th} neuron and u_i represents a membrane recovery variable, which provides negative feedback to v_i. The system is then described by the following coupled ordinary non-linear differential equations:

$$\frac{dv_i}{dt} = 0.04v_i^2 + 5v_i + (140 - u_i) + I_i \tag{1}$$

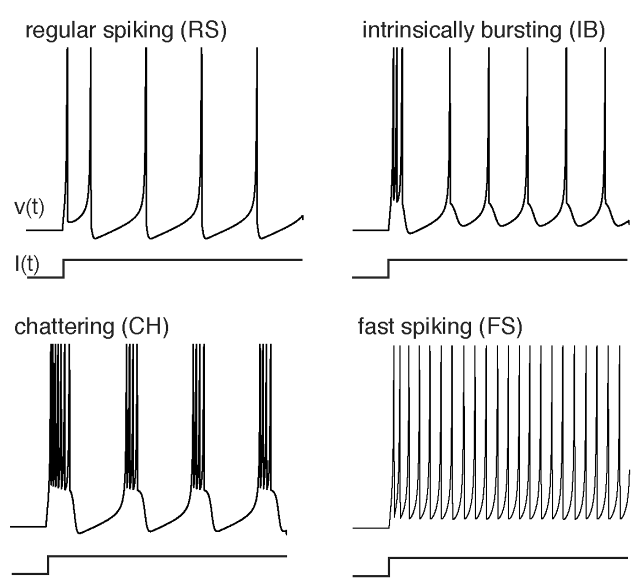
$$\frac{dv_i}{dt} = a(bv_i - u_i) \tag{2}$$

with the following auxiliary condition after spike resetting: if $v_i \geq 30$ millivolts, then $v_i \rightarrow c$ and $u_i \rightarrow (u_i + d)$. Essentially, this means that when a neuron receives a spike input, its membrane potential is immediately reset.

The neurons are coupled to one another through a matrix of synaptic connection weights. Synaptic connection weights are given by the matrix $S = (s_{ij})$, such that the firing of the j^{th} neuron instantaneously changes the variable v_i by s_{ij}. Variations on this rule can make the updating of the weights more biologically informed, for example, by including plasticity and axonal conduction delays. Synaptic currents or injected DC-currents (currents that come from either other neurons or from sensory information) are encompassed within variable I. Variables a, b, c and d are parameters whose effects are summarized in Fig. 2. Basically, a describes the time scale of the recovery variable u_i and b describes the sensitivity of u_i to the fluctuations of the membrane potential v_i. The parameter c corresponds to the after-spike reset value of v_i and d represents the after-spike reset value of u_i. Different values for these parameters produce different individual intrinsic neuron firing patterns such that complex spiking, bursting or chattering of cortical and thalamic neurons can be simulated. Izhikevich has suggested typical values for these variables and the types of patterns they engender [12]. Figure 3 illustrates four examples of such patterns, referred to as regular spiking (RS), intrinsically bursting (IB), chattering (CH) and fast spiking (FS).

All of the parameters in the model can be used to control the time of firing for each neuron. Every time a neuron fires, a dot such as the one shown in Fig. 1 is placed in a graph, and we use these events as temporal points at which sounds are produced.

Fig. 3. Examples of spiking patterns that can be produced by the SNN model. Each shows the response of the model to a step of DC-current I(t) = 10. Note I(t) in the figure corresponds to the variable I in differential equations (1) and (2). (Screen grab ©Eugene Izhikevich. Electronic version of this figure and reproduction permissions are freely available at <www.izhikevich.com>.)



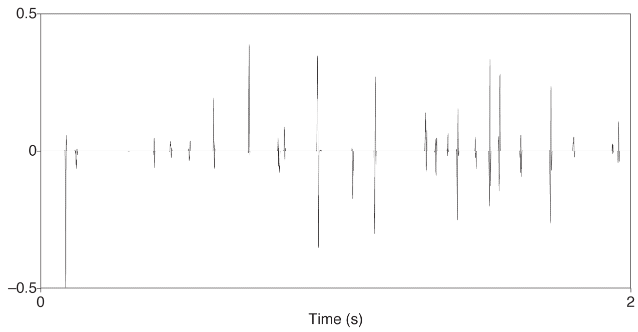


Fig. 4. An example produced by the neurogranular sampler using the simple SNN model with four excitatory neurons and one inhibitory neuron, all of regular spiking type. (© Eduardo Miranda)

The Neurogranular Sampler

The neurogranular sampler works by taking short segments, or sound grains, from sound files and triggering them when any of the neurons fire. These sound grains are packets of sound within a range of duration of approximately 10–100 milliseconds. When a neuron in the network fires at time *t*, a sound grain of predetermined length and amplitude is taken from the recorded sample of sound, convoluted within a Hanning envelope [13] and played. Networks with synchronized firing of neurons produce a periodic pulse, whilst networks containing only a few firing neurons produce sparse rhythmic sequences.

We have implemented two versions of the neurogranular sampler using two variations of the SNN model proposed by Izhikevich: a simple model, where the neurons are instantaneously updated with firing information with static connectivity (that is, the architecture of the network does not change; the connections between the neurons remain static), and an advanced model, wherein axonal conduction delays are built into the system. The axonal conduction delays replicate the time that spike waves take to travel along the axons to the synapses of the post-synaptic neurons. In this case, the elements of the matrix of connections are updated according to a spike-timing dependent plasticity (STDP) algorithm. In this algorithm, connections for which pre-synaptic spikes cause post-synaptic firing are potentiated and those for which pre-synaptic spikes arrive after post-synaptic firing has occurred are depressed. At present, the neuronal network remains isolated from sensory information; that is, it is not stimulated from outside the network. Rather, the network is driven by noise such that the initial current and voltage parameters within a proportion of neurons will be of a high enough value to drive them to fire. It is interesting in itself that random (noisy) inputs can produce synchronous rhythms, a phenomenon well understood within the dynamic systems community.

In the implementation of the neurogranular sampler using the simple SNN model, the connections are geometrically noisy, in the sense that the matrix *S* is a random matrix with all-to-all connections and all current inputs are noise. When all *N* neurons are in regular spiking mode, a variety of musically interesting results can be produced by having either rather few (up to 10) or many (over 500) neurons. The result with up to 10 neurons sounds very sparse, but it is possible to hear rhythmic patterns emerge and then transiently die away. Figure 4 shows a didactic simple example, which includes only four excitatory neurons and one inhibitory neuron (which all have associated negative elements in the matrix *S*); all have regular spiking behavior. The sound sample was taken from the recording of a single note played on a harmonium. In this case, the lengths of the grains were between 250 and 500 samples (with sampling frequency at 44,100 Hz) and had randomly assigned amplitudes. Synchronous behavior takes place if all the neurons are identical and if there are more than 500 of them. This sounds like a very gritty pulse, especially if the selected grain size is short.

The neurogranular sampler using the advanced SNN model included axonal conduction delays and the STDP algorithm. By way of comparison with the output of the sampler using the simple SNN model shown in Fig. 4, Fig. 5 shows the output using the advanced SNN model with four excitatory neurons and one inhibitory neuron, with an axonal delay of up to 10 milliseconds, including STDP. Note that there is much more firing activity here and that this firing appears in a much more correlated fashion. As with the example shown in Fig. 4, the sound grains in Fig. 5 were also taken from a single note played on a harmonium. The lengths of the grains were between 250 and 500 samples (with sampling frequency at 44,100 Hz) and the grains had randomly assigned amplitudes.

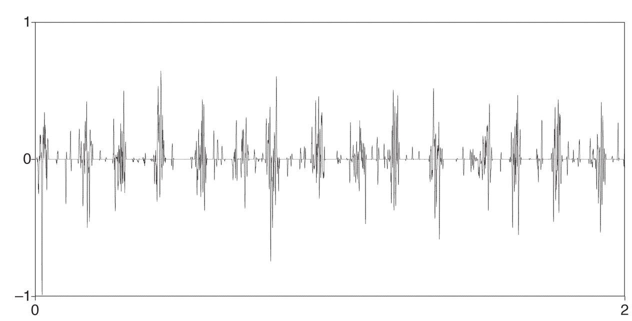
The axonal delays, along with the STDP algorithm, encourage particular pathways in the network to become established, which leads to more events and more regular frequency of neuronal firing. In the absence of sensory input to the network, these pathways have a transient lifetime, but we would expect these pathway lifetimes (and similarly the correlations in the audio output) to increase substantially if there were repeated correlated sensory input to the network.

Concluding Remarks

In this paper we have outlined a novel approach to sound synthesis based on neuronal functional paradigms. We have presented the neurogranular sampler, which uses spiking neuronal networks (SNN) models to control the triggering of sound grains taken from a given recorded sound.

Recently, artist Jane Grant used the neurogranular sampler to produce the sonic elements within a sound and video installation called *Threshold* [14]. In this work, the sound of both voice and breath are recorded and then reconfigured via the neurogranular sampler in real time in order to merge the voice or breath with the patterns and rhythms occurring in the neuronal network. The work uses the advanced version of the sampler, including spike-timing dependent plasticity, and a new interface for the instrument was developed in order to complete the work [15]. The neurogranular sampler also forms the basis of a new work entitled *The Fragmented Orchestra* by Jane Grant, John Matthias and Nick Ryan [16], in which a real-time

Fig. 5. Output from the neurogranular sampler using the advanced SNN model with four excitatory neurons and one inhibitory neuron, all of regular spiking type, including an axonal delay of up to 10 milliseconds and spike-timing dependent plasticity (STDP). (© Eduardo Miranda)



version of the instrument is spatially distributed around 24 sites in the U.K. and at the FACT Gallery in Liverpool. The work recently won the PRS New Music Award 2008 [17]. We have also exploited the neuronal firing events to trigger signals for performers via flashing LED lights in a new work, *Cortical Songs* by John Matthias and Nick Ryan [18] for solo violin and string orchestra. *Cortical Songs* has been released as an album, on the Nonclassical record label, which also included 11 remixes by artists and musicians such as Radiohead's Thom Yorke and the Verve's Simon Tong [19].

Technically, we found that the system has a very wide variety of temporal patterns and behaviors, which can be controlled according to the parameters of the model. Different sounds can be obtained by varying parameters such as:

a. the number and type of neurons
b. the geometry of the connectivity
c. the parameters a, b, c and d, which determine the intrinsic properties of the neurons
d. the "sensory" input I
e. the nature of the sample sound source
f. the duration of the triggered sample.

Most of the settings above can be interpolated during sound production, allowing for changes on the fly. Generally speaking, increasing the number of neurons in the model means more firing and therefore more sonic texture. However, when the network exhibits synchronous behavior, increasing the number of neurons tends to lower the frequency of the collective response.

References

Unedited references as provided by the authors.

1. Churchland, P. and Sejnowski, T. 1994. *The Computational Brain*. Cambridge, MA: The MIT Press.

2. O'Reilly, R. and Munakata, Y. 2000. *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: The MIT Press.

3. McKinstry, J.L., Edelman, G.M. and Krichmar, J.L. 2006. "A cerebellar model for predictive motor control tested in a brain-based device." *Proceedings of the National Academy of Sciences of the United States of America* 103(9):3387–3392.

4. Shibata, T., et al. 2001. "Humanoid Oculomotor Control Based on Concepts of Computational Neuroscience." *Proceedings of Humanoid 2001—2nd IEEE-RAS International Conference of Humanoid Robots*. Waseda University, Japan.

5. Miranda, E.R., et al. 2003. "On Harnessing the Electroencephalogram for the Musical Braincap." *Computer Music Journal* 27(2):80–102.

6. Miranda, E R. (2006). "Brain-Computer music interface for composition and performance," *International Journal on Disability and Human Development*, 5(2):119–125.

7. Miranda, E.R. and Brouse, A. (2005). "Interfacing the Brain Directly with Musical Systems: On Developing Systems for Making Music with Brain Signals," *Leonardo*, 38(4):331–336.

8. Coath, M. and Denham, S. 2005. "Robust sound classification through the representation of similarity using response fields derived from stimuli during early experience." *Biological Cybernetics* 93(1): 22–30.

9. Gerstner, W. and Kistler, W.M. 2002. *Spiking Neuron Models*. Cambridge: Cambridge University Press.

10. Ikegaya Y., et al. 2004. "Synfire chains and cortical songs: Temporal modules of cortical activity." *Science* 304:559–564.

11. Izhikevich, E.M. 2003. "Simple model of spiking neurons." *IEEE Transactions on Neural Networks* 14:1469.

12. Izhikevich [11].

13. Abramovitz, M. and Stegun, I. 1972. *Handbook of Mathematical Functions*. New York: Dover.

14. Grant, J. 2008. *Threshold,* ArtSway Gallery UK.

15. Grant, J., Matthias, J., Hodgson, T. and Miranda, E.R. 2008. "Hearing Thinking," Paper presented at *EvoMUSART—Sixth European Workshop on Evolutionary and Biologically Inspired Music, Sound, Art and Design*, Napoli, Italy.

16. <www.thefragmentedorchestra.com/> (accessed 22 July 2008).

17. <www.prsfoundation.co.uk/newmusicaward/index.htm> (accessed 22 July 2008).

18. Matthias, J. and Ryan, N. 2007. "Cortical Songs—Sonic Events from Spiking Neurons Body," Space and Technology Journal—Perspectives 8. <http://people.brunel.ac.uk/bst/vol0701/johnmatthias/home.html> (accessed 22 July 2008).

19. Matthias, J. and Ryan, N. 2008. *Cortical Songs*, Music album CD (NONCLSS004), London, Nonclassical Records.

Manuscript received 19 March 2008.

Eduardo R. Miranda is a professor in computer music and head of the Interdisciplinary Centre for Computer Music Research (ICCMR) at the University of Plymouth, U.K. He co-authored the book New Digital Musical Instruments: Control and Interaction Beyond the Keyboard *(A-R Editions, 2006) and has authored a number of research papers on new technologies for music. His solo album CD* Mother Tongue *(Sargasso, London) includes a selection of works from his research into new possibilities and technologies for composing with voice.*

John Matthias lectures on music and music technology at the Faculty of Arts in University of Plymouth, U.K., and conducts research within the Interdisciplinary Centre for Computer Music Research (ICCMR). He has worked with many recording bands and artists including Radiohead, Matthew Herbert and Coldcut. He currently plays in the band Derailer and co-authored with Nick Ryan the album CD Cortical Songs, *released by Nonclassical, London.*