AFRL-AFOSR-VA-TR-2016-0285

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing

Aurel Lazar
THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
116TH AND BDWY
NEW YORK, NY 10027

08/04/2016
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/RTA2

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* 30/07/2016	**2. REPORT TYPE** Final Report	**3. DATES COVERED** *(From - To)* 01/03/2012-30/04/2016

4. TITLE AND SUBTITLE Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER FA9550-12-1-0232
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Aurel A. Lazar	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Columbia University New York, NY 10027	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 875 N. Randolph Street Arlington, VA 22203	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION A: Distribution approved for public release.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During the research period we: (i) devised pathbreaking algorithms for the functional identification and evaluation of non-linear dendritic processing, and (ii) released a groundbreaking open source platform for the isolated and integrated emulation of the fruit fly brain on multiple GPUs (Neurokernel). We established that identifying a single dendritic stimulus processor (DSP) is mathematically dual to decoding of stimuli encoded by a population of neurons with a bank of DSPs. Building on this key duality property, we (i) demonstrated that the evaluation of the functional identification methodology can be effectively and intuitively performed in the stimulus space, and (ii) characterized the effect of noise parameters on the precision of the functional identification of feedforward, feedback and cross-feedback neural circuits with DSP/biological spike generator neuron models. We demonstrated the power of Neurokernel's model integration by combining independently developed models of the retina and lamina neuropils in the fly's visual system and by demonstrating their neuroinformation processing capability.

15. SUBJECT TERMS

dendritic stimulus processors, biological spike generators, duality between decoding and identification, emulation of the fruit fly brain.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU		**19b. TELEPHONE NUMBER** *(Include area code)*

Reset

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18
Adobe Professional 7.0

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing

Final Report *

Aurel A. Lazar
Department of Electrical Engineering
Columbia University, New York, NY 10027

July 28, 2016

Abstract

During the 2012-2016 research period we (i) devised pathbreaking algorithms for the functional identification and evaluation of non-linear dendritic processing (Channel Identification Machines), and (ii) released a groundbreaking open source platform for emulating the fruit fly brain on multiple GPUs (Neurokernel).

We pioneered a multi-input multi-output neural circuits architecture for *non-linear* processing and encoding of stimuli in the spike domain. The basic neuron model in these circuits consists of multi-sensory spatio-temporal dendritic stimulus processors (DSPs) cascaded with a biophysical spike generator (BSG). Multi-sensory spatio-temporal DSPs arise in the encoding of color visual fields and the mixing of natural visual fields and auditory scenes. We established that identifying a single dendritic stimulus processor is mathematically dual to decoding of stimuli encoded by a population of neurons with a bank of dendritic stimulus processors. Building on this key duality property, we (i) demonstrated that the evaluation of Channel Identification Machines can be effectively and intuitively performed in the stimulus space, and (ii) characterized the effect of noise parameters on the precision of the functional identification of feedforward, feedback and cross-feedback neural circuits with DSP/BSG neuron models.

Neurokernel is a groundbreaking open-source platform for the isolated and integrated emulation of fruit fly brain model neural circuits (e.g., sensory and locomotion systems), their connectivity patterns, and other parts of the fly's nervous system on clusters of GPUs. By defining mandatory communication interfaces that specify how data is transmitted between models of each of these modules regardless of their internal design, Neurokernel explicitly enables multiple researchers to collaboratively model the fly's entire brain by integration of their independently developed models of its constituent processing units. We demonstrated the power of Neurokernel's model integration by combining independently developed models of the retina and lamina neuropils in the fly's visual system and by demonstrating their neuroinformation processing capability.

---

[*] The work reported here was supported by AFOSR under Grant/Contract Number: FA9550-12-1-0232

We have been instrumental in organizing and chairing the CNS workshops on Methods of Information Theory in Computational Neuroscience (2012-2014), Methods of System Identification for Studying Information Processing in Sensory Systems (2012-2015) and Open Collaboration in Computational Neuroscience (2015) during the Computational Neuroscience Meetings (CNS). We also chaired and organized the highly acclaimed Columbia Workshop on Brain Circuits, Memory and Computation (2015-2016). Finally, we co-edited a Special Issue on Biological Applications of Information Theory Claude Shannons Centennial of the IEEE Transactions on Molecular, Biological and Multi-Scale Communications, 2016 (to appear).

The report enclosed consists of the detailed yearly progress reports of the entire research period 2012-2016.

Keywords: dendritic stimulus processors, biological spike generators, duality between decoding and identification, emulation of the fruit fly brain.

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing
First Year Report [*]

Aurel A. Lazar
Department of Electrical Engineering
Columbia University, New York, NY 10027

April 30, 2013

Abstract

During the first research period we made substantial progress in three areas: (i) channel identification machines, (ii) open architecture for the massively parallel emulation of the fruit fly brain on multiple GPUs, and (iii) functional identification and evaluation of massively parallel neural circuits.

The channel identification methodology that we investigated employs test signals that are neither white, nor have stationary statistics (e.g., Gaussian with a fixed mean/variance). This is a radical departure from the widely employed nonlinear system identification methods, including the spike-triggered average and the spike-triggered covariance methods. We carried out the channel identification using input signals that belong to Reproducing Kernel Hilbert Spaces (RKHSs), and in particular, spaces of bandlimited functions, i.e., functions that have a finite support in the frequency domain. The latter signals are extensively used to describe sensory stimuli in biological systems and to model signals in communications. We showed that for such signals the channel identification problem becomes mathematically tractable. Furthermore, we demonstrated that the choice of the input signal space profoundly effects the type of identification results that can be achieved.

We have developed an open software architecture for emulating neural circuit modules in the fly brain and their responses to recorded or simulated input stimuli on multiple Graphics Processing Units (GPUs). A key feature of this architecture is its support for integrating instances of different neural circuit models potentially developed by independent researchers by requiring that the models implementations provide interoperable interfaces that adhere to the specification prescribed by the architecture. We refer to the architecture as a Neurokernel because it provides object classes essential to the emulation of the entire fruit fly brain that are analogous to those provided by an operating system kernel: (1) it serves as an extended machine that provides access to neural circuit primitives needed to construct and interconnect models of neural circuit modules in the fly brain; and (2) it serves as a resource allocator that scalably and transparently assigns GPU resources to emulated neural circuit models without manual specification by the researcher.

---

[*] The work reported here was supported by AFOSR under Grant/Contract Number: FA9550-12-1-0232

We substantially expanded our code base for representing information in the time domain and for spike processing with a novel class of algorithms for identifying neural circuits and architectures. Called Channel Identification Machines, these algorithms are based on a novel duality between neural encoding and functional identification of neural circuits.

1 Introduction

During the past twelve months we focused on the identification of massively parallel neural circuits and the emulation of the fruit fly brain. We also focussed on the implementation and demonstration of associated algorithms on a large scale GPU cluster.

2 Channel Identification Machines

Signal distortions introduced by a communication channel can severely affect the reliability of communication systems. If properly utilized, knowledge of the channel response can lead to a dramatic improvement in the performance of a communication link. In practice, however, information about the channel is rarely available a priori and the channel needs to be identified at the receiver. A number of channel identification methods have been proposed for traditional clock-based systems that rely on the classical sampling theorem. However, there is a growing need to develop channel identification methods for asynchronous nonlinear systems, of which Time Encoding Machines (TEMs) are a prime example.

We investigated the following *nonlinear* identification problem: given both the input signal $\mathbf{u}$ and the time sequence $(t_k)_{k\in\mathbb{Z}}$ at the output of a TEM, what is the channel filter? System identification problems of this kind are key to understanding the nature of neural encoding and processing, process modeling and control and, more generally, methods for constructing mathematical models of dynamical systems.

Identification of the channel from a time sequence is to be contrasted with existing methods for rate-based models in neuroscience. In such models the output of the system is taken to be its instantaneous response rate and the nonlinear generation of a time sequence is not explicitly modeled. Furthermore, in order to fit model parameters, identification methods for such models typically require the response rate to be known. This is often difficult in practice since the same experiment needs to be repeated a large number of times to estimate the response rate. Moreover, the use of the same stimulus typically introduces a systematic bias during the identification procedure.

The channel identification methodology that we investigated employs test signals that are neither white, nor have stationary statistics (e.g., Gaussian with a fixed mean/variance). This is a radical departure from the widely employed nonlinear system identification methods, including the spike-triggered average and the spike-triggered covariance methods. We carry out the channel identification using input signals that belong to Reproducing Kernel Hilbert Spaces (RKHSs), and in particular, spaces of bandlimited functions, i.e., functions that have a finite support in the frequency domain. The latter signals are extensively used to describe sensory stimuli in biological systems and to model signals in communications. We showed that for such signals the channel identification problem becomes mathematically tractable. Furthermore, we demonstrated that the choice of the input signal space profoundly effects the type of identification results that can be achieved. In our paper

- Lazar, A.A. and Slutskiy, Y.B., "Channel Identification Machines", *Journal of Computational Intelligence and Neuroscience*, Volume 2012, July 2012, 20 pages.

we introduced three application-driven examples and formally stated the channel identification problem. We presented the single-input single-output (SISO) channel identification machine (CIM) for the finite-dimensional input signal space of trigonometric polynomials. Using analytical methods and simulations, we demonstrated that it is possible to identify the projection of the filter onto the input space loss-free and showed that the SISO CIM algorithm can recover the original filter with arbitrary precision, provided that both the bandwidth and the order of the input space are sufficiently high. We extended our methodology to multidimensional systems and presented multi-input single-output (MISO) CIM algorithms for the identification of vector-valued filters modeling the channel. We generalized our methods to classes of RKHSs of input signals and worked out in detail channel identification algorithms for infinite-dimensional Paley-Wiener spaces. Furthermore, we discussed extensions of our identification results to noisy systems, where additive noise is introduced either by the channel or the asynchronous sampler.

3 An Open Architecture for the Massively Parallel Emulation of the Fruit Fly Brain on Multiple GPUs

We have developed an open software architecture for emulating neural circuit modules in the fly brain and their responses to recorded or simulated input stimuli on multiple Graphics Processing Units (GPUs). A key feature of this architecture is its support for integrating instances of different neural circuit models potentially developed by independent researchers by requiring that the models implementations provide interoperable interfaces that adhere to the specification prescribed by the architecture. We refer to the architecture as a Neurokernel because it provides object classes essential to the emulation of the entire fruit fly brain that are analogous to those provided by an operating system kernel:

- it serves as an extended machine that provides access to neural circuit primitives needed to construct and interconnect models of neural circuit modules in the fly brain; and
- it serves as a resource allocator that scalably and transparently assigns GPU resources to emulated neural circuit models without manual specification by the researcher.

In order to provide these features, the Neurokernel architecture comprises several planes of abstraction that separate its application, control, and computing aspects. Models of brain function implemented using the architecture use the application planes API to access neural circuit primitives without directly specifying which GPU resources to use. The architectures control plane automatically partitions and maps circuits to available GPU resources, and manages communication between multiple GPUs hosted locally or remotely. Storage methods used to efficiently represent large networks of neurons and synapses with feedback connections in GPU memory and numerical methods used to update neuron and synapse states are implemented in the computing plane. We implemented key elements the Neurokernel software using the Python programming language and the PyCUDA interface to NVIDIAs CUDA GPU programming environment to avail ourselves of the increasingly powerful ecosystem of scientific computing Python software and make the architecture accessible to other researchers in the neuroscience community.

We presented our initial results as follows:

- L. E. Givon and A. A. Lazar, “Neurokernel: An Open Architecture for Massively Parallel Emulation of Drosophila Brain Models on Multiple GPUs”, Workshop on Insect Vision: Cells, Computation, and Behavior, Janelia Farms Campus, March 3-6, 2013.
- L. E. Givon and A. A. Lazar, “An Open Architecture for the Massively Parallel Emulation of the Drosophila Brain on Multiple GPUs”, Computational Neuroscience Meeting, Atlanta/Decatur, GA, July 21-24, 2012.

4 Functional Identification and Evaluation of Massively Parallel Neural Circuits

Parameter estimation and its evaluation are at the core of functional identification of neural circuits. How are estimates of neural circuits impacted by neuronal biophysics and particular stimuli employed in experiments? What is a suitable metric to assess the goodness of the identified circuit parameters? How can this metric be extended from a single neuron to a circuit consisting of thousands of neurons? We obtained new insights by asking these questions in the rigorous context of channel identification/time encoding and the massively parallel models of biophysical circuits of the early visual system.

Our approach builds on the observation that during an experiment, an entire neural circuit is projected onto the space of input signals, with the circuit projection intuitively determined by the way the input stimuli explore the functional characteristics of that circuit. Assuming the general outline of the circuit architecture is known, we identified the projection of the encoding circuit whose receptive fields and parameters of individual neurons are unknown. We demonstrated that the projection (i.e., each identified parameter) is stimulus-dependent and derive a quantitative description of this dependence in the parameter space. We provided a rigorous theoretical analysis and derived conditions under which the distance between the original and identified parameters can be made arbitrarily small.

After identification, one typically evaluates the estimated parameters by how well individual model neurons predict recorded responses. Although such evaluation in the spike-train domain (space) can be useful, it is often insufficient, especially in a population encoding setting. We argued that it is beneficial to evaluate the identified circuit in the stimulus space by looking at the amount of information that the spike train ensemble retains about the stimulus. Specifically, by employing a nonlinear decoder, we reconstructed novel stimuli sensed by the circuit (i.e., not necessarily the original stimuli). In contrast to linear decoders applied to firing rates of a small population of neurons, our nonlinear decoder works directly with spike times produced by thousands of neurons and is capable of high-fidelity reconstruction of the sensed stimulus. Thus, evaluation of the entire identified neural circuit is reduced to intuitive comparisons in the stimulus space, thereby augmenting the usual neuron-by-neuron comparison in the spike-train space. For images and video, such an evaluation can be visually performed.

We presented extensive simulation results for massively parallel neural circuits that were identified using natural video with the aid of GPU computing. First, we showed that for a faithful encoding model, high quality reconstruction of visual stimuli from spike trains of biophysical neurons is possible. The quality of the decoded video improves with the length of stimuli used in identification, implying a higher accuracy of neural circuit identification. By comparing the reconstructed stimulus with the original, we then demonstrated an algorithm for determining if additional experiments need to be carried out. Finally, we explored the use of the latter algorithm in the case of unobservable/missing circuit outputs and as a means of exploring the circuit function.

We will present our results at CNS*2013:

- A.A. Lazar, Y.B. Slutskiy and Y. Zhou, “Functional identification and evaluation of massively parallel neural circuits”. In *Computational Neuroscience Meeting, Paris, France, July 13-18*, 2013.

5 Code for Reproducible Research

We substantially expanded our code base for representing information in the time domain and for spike processing with a novel class of algorithms for identifying neural circuits and architectures. Called Channel Identification Machines, these algorithms are based on a novel duality between neural encoding and functional identification of neural circuits. Our progress is documented below.

This section is organized as follows. In section 5.1 we describe progress with the implementation of the Time Encoding and Decoding (TED) Toolbox. In section 5.2 we describe a new toolbox for Channel Identification Machines (CIM). The scikits.cuda set of high level programming utilities are described in section 5.3. Finally, the state of the Internal Code Repository is presented in section 5.4.

5.1 Time Encoding and Decoding Toolbox

We continued to release revisions of the Time Encoding and Decoding (TED) Toolbox, adding more functions and in the process substantially increasing the computational efficiency of our algorithms.

- Bionet Group, Time Encoding and Decoding (TED) Toolbox, `http://www.bionet.ee.columbia.edu/code/ted`.

Here we highlight some of the key novel features:

- Merged Python bionet.utils toolbox with TED toolbox to facilitate user installation.

5.2 Chanel Identification Machines Toolbox

The Channel Identification Machines (CIMs) Toolbox provides a set of algorithms for the functional identification of spiking neural circuits, as well as a step-by-step demonstration of the accompanying source code. The neural circuits initially considered consist of a linear dendritic processing filter in cascade with a spiking neuron model. The CIM algorithms identify the dendritic processing filter and reconstruct its kernel with arbitrary precision.

- Bionet Group, Channel Identification Machines (CIM) Toolbox, `http://www.bionet.ee.columbia.edu/code/cim`.

5.3 scikits.cuda

Novel features of the above set of utilities include:

- Added bug fixes and various contributions by users.
- Improved installation via pip.
- Improved CUBLAS version detection to avoid affecting performance by surreptitiously creating temporary contexts to run certain diagnostic functions in CUBLAS.
- Added support for specifying streams when creating FFT plans.
- Rewritten CUBLAS bindings to add support for CUBLAS v2 API (which permits explicit management of GPU contexts used by CUBLAS).
- Added initial support for some CUBLAS 5 functions.

5.4 Internal Code Repository

Progress on the Internal Code Repository was mainly focused on three key developments. First, we developed a GPU computational model for simulating the Local Processing Units (LPUs) arising in fly vision. Second, we developed GPU simulations for the identification of spatio-temporal receptive fields arising in massively parallel neural circuits. Lastly, we upgraded the PyCUDA package for Video Time Encoding and Decoding. The old package includes simple demonstrations of the Video TEM with Gabor or center-surround receptive fields and IAF neurons. A release of the old package is under preparation. The upgraded package has the following features:

- Modularized and more object oriented.
- A new design of spike storage files that store all necessary information for decoding.
- Biophysical spiking neuron models are now supported in an unified way, making it simple to add new neuron models.
- Parent classes of spatio-temporal or spatial receptive fields allow the simple creation of new types of receptive fields.
- Vectorized input videos are now supported, such as color and 3D videos.

6 Workshops Organized

Following up on our past involvement with the Computational Neuroscience Meeting, we have organized the 2012 workshop

- CNS*2012 Workshop on Methods of Information Theory in Computational Neuroscience, July 25-26, 2012, Atlanta, GA. `http://www.bionet.ee.columbia.edu/workshops/cns/methods/2012/information`

Overview

Methods originally developed in Information Theory have found wide applicability in computational neuroscience. Beyond these original methods there is a need to develop novel tools and approaches that are driven by problems arising in neuroscience.

A number of researchers in computational/systems neuroscience and in information/communication theory are investigating problems of information representation and processing. While the goals are often the same, these researchers bring different perspectives and points of view to a common set of neuroscience problems. Often they participate in different fora and their interaction is limited.

The goal of the workshop is to bring some of these researchers together to discuss challenges posed by neuroscience and to exchange ideas and present their latest work.

The workshop is targeted towards computational and systems neuroscientists with interest in methods of information theory as well as information/communication theorists with interest in neuroscience.

Participants (Invited)

Bruno B. Averbeck, Laboratory of Neuropsychology, NIMH/NIH.
Ehud Kaplan, The Neuroscience Department, Mount Sinai School of Medicine.
Alex G. Dimitrov, Department of Mathematics, Washington State University - Vancouver.
Tatyana O. Sharpee, Computational Neurobiology Laboratory, The Salk Institute.
Rava Azeredo da Silveira, Laboratoire de Physique Statistique, Ecole Normale Superieure.
Gabriel A. Silva, Systems Neural Engineering and Theoretical Neuroscience, UCSD.
Marius Buibas, Scientist, Brain Corporation, San Diego, California.
Gaute T. Einevoll, Department of Mathematical Sciences, Norwegian University of Life Sciences.
Robert Sinclair, Mathematical Biology Unit, Okinawa Institute of Science and Technology, Japan.
Conor Houghton, Department of Mathematics, Trinity College Dublin.
Sanggyun Kim, Department of Bioengineering, University of California, San Diego.
Christopher J. Rozell, School of Electrical and Computer Engineering, Georgia Tech.
William B. Levy, Laboratory of Systems Neurodynamics, University of Virginia.
Ilya Nemenman, Department of Physics, Emory University.
Jianing Shi, Department of Electrical and Computer Engineering, Rice University.

We have also organized the following workshop during the 2012 Computational Neuroscience Meeting:

- CNS*2012 Workshop on Methods of Information Theory in Computational Neuroscience, July 25, 2012, Atlanta/Decatur, GA. `http://www.bionet.ee.columbia.edu/workshops/cns/methods/2012/identification`

Overview

A functional characterization of an unknown system typically begins by making observations about the response of that system to input signals. The knowledge obtained from such observations can then be used to derive a quantitative model of the system in a process called system identification. The goal of system identification is to use a given input/output data set to derive a function that maps an arbitrary system input into an appropriate output.

In neurobiology, system identification has been applied to a variety of sensory systems, ranging from insects to vertebrates. Depending on the level of abstraction, the identified neural models vary from detailed mechanistic models to purely phenomenological models.

The workshop will provide a state of the art forum for discussing methods of system identification applied to the visual, auditory, olfactory and somatosensory systems in insects and vertebrates.

The lack of a deeper understanding of how sensory systems encode stimulus information has hindered the progress in understanding sensory signal processing in higher brain centers. Evaluations of various systems identification methods and a comparative analysis across insects and vertebrates may reveal common neural encoding principles and future research directions.

The workshop is targeted towards systems, computational and theoretical neuroscientists with interest in the representation and processing of stimuli in sensory systems in insects and vertebrates.

Participants (Invited)

Garrett B. Stanley, Department of Biomedical Engineering, Georgia Tech and Emory University.
Vladimir Brezina, Department of Neuroscience, Mount Sinai School of Medicine.
Mikko I. Juusola, Department of Biomedical Science, University of Sheffield,
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Matthieu Louis, Centre for Genomic Regulation, Barcelona.
Daniel Coca, Department of Automatic Control and Systems Engineering, University of Sheffield.
Mikko I. Juusola, Department of Biomedical Science, University of Sheffield.
Fabrizio Gabbiani, Department of Neuroscience, Baylor College of Medicine.
Shy Shoham, Department of Biomedical Engineering, Technion.
Theodore W. Berger, Department of Biomedical Engineering, University of Southern California.

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing
Second Year Report *

Aurel A. Lazar
Department of Electrical Engineering
Columbia University, New York, NY 10027

March 29, 2014

Abstract

During the second research period we made substantial progress by (i) devising pathbreaking algorithms for the functional identification and evaluation of biological neural circuits (Channel Identification Machines or CIMs), and (ii) releasing a groundbreaking open source platform for emulating the fruit fly brain on multiple GPUs (Neurokernel).

We have devised a class of pathbreaking CIM algorithms that enable the identification of neural circuits with spiking inputs - no other machine learning algorithms exist today that solve this critical task. The identified neural circuits have been shown to converge to stable phenomenological units - a fundamental characteristic that is lacking in Bayesian approaches such as the generalized linear models. Furthermore, these novel CIMs are highly effective in identifying non-linear dendritic processing - also a first. Finally, the novel theoretical results on CIMs have shown that in order to identify non-linear neural circuits, these have to be deconstructed into smaller operational units. Simply put, repeated nonlinearities in the signal path are algorithmically not manageable. For example, it is necessary to identify the axon hillock (using patch clamping) separately from the dendritic processing unit of the neurons.

Neurokernel is a groundbreaking open-source platform for the isolated and integrated emulation of fruit fly brain model neural circuits (e.g., sensory and locomotion systems), their connectivity patterns, and other parts of the fly's nervous system on clusters of GPUs. Neurokernel also serves as a platform for in vivo functional identification of neural circuits. The Neurokernel project challenges the basic model of knowledge creation in neuroscience. The project employs Neurokernel Requests for Comments documents (Neurokernel RFCs) as a tool both for advancing the design of the architecture and the neural system models built with it. While open-source projects have been proposed in computational neuroscience, they tend to have a monolithic structure that continue to lack strong appeal in the neuroscience community. Neurokernel provides standard APIs among local processing units (modeling neuropils) and tools to build an architecture from components developed by the community of researchers at large. Most importantly, it does not impose a unique computational model - rather it allows the existence of many models that can be comparatively evaluated.

---

[*] The work reported here was supported by AFOSR under Grant/Contract Number: FA9550-12-1-0232

1 Key Publications

During the past twelve months we published 2 pathbreaking papers, and submitted one paper for publication. We also submitted one patent application to the USPTO and are in the process of submitting a paper for publication that is based on the patent application.

1.1 Functional Identification of Spike-Processing Neural Circuits

A number of methods have been proposed to quantify and model neuronal processing. The majority of these methods assume that the input to a neural circuit is continuous and that the output is a sample path of a point process (e.g., Poisson process). In biological neural circuits, however, the inputs for most neurons are spike trains generated by presynaptic cells and the outputs are determined by a multi-dimensional dynamical system. Furthermore, the highly nonlinear nature of spike generation has been shown to produce interactions between stimulus features that profoundly affect the estimation of receptive fields. Hence, there is a fundamental need to develop tractable methods for identifying neural circuits that incorporate biophysical models of spike generation and receive multi-dimensional spike trains as input stimuli.

Here we describe a new methodology for a *complete* neural circuit identification that takes the above-mentioned considerations into account. Our neural circuit models admit multidimensional spike trains as input stimuli. They also incorporate nonlinear dynamical system (Hodgkin-Huxley, Morris Lecar, hard-threshold IAF, etc.) models of the spike-generating mechanism. The nonlinear contribution of a dynamical system such as the Hodgkin-Huxley neuron model is stimulus-driven. It changes from one spike to the next and thus affects receptive field estimation if not properly taken into account.

Our approach builds on the observation that during any experiment, a neural circuit is projected onto a particular space of input signals, with the circuit projection determined by how well the input space explores that circuit. Employing reproducing kernel Hilbert spaces (in particular, spaces of bandlimited functions) to model the input stimuli, we quantitatively described the relationship between the underlying circuit parameters and their projections. We also derived conditions under which these projections converge to the true parameters. In doing so, we achieved the mathematical tractability needed to characterize biophysical spike generators and to identify the multitude of receptive fields in neural circuits with full connectivity. We estimated all model parameters directly from spike times produced by neurons and repeating the same stimulus is no longer necessary.

- Lazar, A.A. and Slutskiy, Y.B., “Functional Identification of Spike-Processing Neural Circuits”, *Neural Computation*, Volume 26 , Number 2 , MIT Press , pp. 264-305 , February 2014.

1.2 Multisensory Encoding, Decoding, and Identification

We investigated a biophysically-grounded spiking neural circuit and a tractable mathematical methodology that together allow one to study the problems of multisensory encoding, decoding, and identification within a unified theoretical framework. Our neural circuit is comprised of a bank of multisensory receptive fields in cascade with a population of neurons that implement stimulus multiplexing in the spike domain. The circuit architecture is quite flexible in that it can incorporate complex connectivity and a number different spike generation models.

Our approach is grounded in the theory of sampling in Hilbert spaces. Using this theory, we showed that signals of different modalities, having different dimensions and dynamics, can be faithfully encoded into a single multidimensional spike train by a common population of neurons. Some benefits of using a common population include (a) built-in redundancy, whereby, by rerouting, a circuit could take over the function of another faulty circuit (e.g., after a stroke) (b) capability to dynamically allocate resources for the encoding of a given signal of interest (e.g., during attention) (c) joint processing and storage of multisensory signals/stimuli (e.g., in associative memory tasks).

More formally, we demonstrated that stimuli of different dimensions can be faithfully multiplexed and encoded in the spike domain and derived tractable algorithms for decoding each stimulus from the common pool of spikes. We also showed that the identification of multisensory processing in a single neuron is dual to the recovery of stimuli encoded with a population of multisensory neurons, and proved that only a projection of the circuit onto input stimuli can be identified. We provided an example of multisensory integration using natural audio and video and discuss the performance of the proposed decoding and identification algorithms.

- Lazar, A.A. and Slutskiy, Y.B., “Multisensory Encoding, Decoding, and Identification”, *Advances in Neural Information Processing Systems 26*, C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K.Q. Weinberger , December 2013 , Lake Tahoe, Nevada, United States.

1.3 Channel Identification Machines for Multidimensional Receptive Fields

The motivation for multidimensional CIMs is provided by the concept of a receptive field that is well established in neuroscience. Introduced in 1906 by Sherrington to describe as area of the body surface capable of eliciting a reflex in response to a stimulus, the term “receptive field” has been extended to many different sensory modalities and spans many different types of neurons. For example, in the visual system, the receptive field of a photoreceptors is a 3-dimensional cone in space comprising all possible directions in which light can hit the photoreceptor. In the auditory system, receptive fields can correspond to certain spectral regions of audio stimuli. More broadly, the receptive field is that part of the the sensory space that can evoke a neuronal response. Spatial and spatiotemporal receptive fields have been successfully used in vision to model retinal ganglion cells in the retina as well as neurons in the lateral geniculate nucleus and the visual cortex. Similarly, spectrotemporal receptive fields have been used to describe responses of auditory neurons, neurons in cochlear nuclei and neurons in the auditory cortex.

More formally, until now all CIMs and their associated methods have been considered only in the context of one-dimensional signals, e.g., functions of time $u_1(t)$, $t \in \mathbb{R}$. Here we discuss multidimensional channel identification machines that allow one to identify signal transformations applied to multidimensional signals $u_n(x_1, ..., x_n)$, $n \in \mathbb{N}$, where x_n typically designates the time variable. A few examples of multidimensional CIMs include (i) spatial CIMs, where the input signal $u_2(x, y)$ is a function of a two-dimensional space, describing, e.g., images; (ii) spectrotemporal CIMs, where the input signal $u_2(\nu, t)$ is a function of spectrum and time, describing, e.g., auditory signals; (iii) spatiotemporal CIMs, where the input signal $u_3(x, y, t)$ is a function of space and time, describing, e.g., video signals.

We presented algorithms for identifying multidimensional receptive fields directly from spike trains produced by biophysically-grounded neuron models. We demonstrated that only a projection of a receptive field onto the input stimulus space may be identified and derived conditions under which this identification is possible. We also provided simulation results for neural circuits incorporating spatiotemporal and spectrotemporal receptive fields.

- Lazar, A.A. and Slutskiy, Y.B., “Channel Identification Machines for Multidimensional Receptive Fields”, *Frontiers in Computational Neuroscience*, Special issue on the role of mathematics and theory in understanding the brain, March 2014, submitted.

1.4 Massively Parallel Neural Circuits for Stereoscopic Color Vision: Encoding, Decoding and Identification

In this paper we obtained two key results.

First, we extended Video Time Encoding Machines to encode videos that have multiple color components. Two direct applications are encoding color videos and encoding 3D stereoscopic videos. For the case of color videos, each neuron is associated with a color receptive field. Each component receptive field defines a linear operation on one of the color channels (*e.g.*, R, G, or B), respectively. The output of each of the color channels of three receptive field components is aggregated and encoded by a single neuron. Note the mixing/encoding of signals from different color channels by the same neurons. We proved that the jointly encoded signals can be de-mixed during the reconstruction process. Modeling of the encoding process took place in the space of vector-valued RKHSs, an extension to the previously used scalar-valued RKHSs. In addition, using the duality between decoding and identification, we demonstrated that under appropriate conditions the color receptive fields can be identified using natural (bandlimited) color videos.

Second, we demonstrated that the evaluation of the quality of functional identification of a massively parallel neural circuit can (and should) be performed in the stimulus space rather than in the spike train space. The duality of the formalism of Time Decoding Machines (TDMs) and Channel Identification Machines (TEMs) provides us with an unique capability for

1. functionally identifying the entire neural circuit,
2. using the identified parameters for decoding novel stimuli encoded with the original massively parallel neural circuit,
3. comparing the decoded stimuli with the original stimuli in the input space.

Such a novel evaluation method provides simple insight into the operation of encoding neural circuits. If a neural circuit can be identified with high precision, the encoded signals can be decoded with a high degree of faithfulness.

We also provided detailed examples of encoding, decoding and identification of color videos with a massively parallel neural circuit. One of the example circuits consists of 30,000 neurons. We systematically studied the quality of identification by using different number of spikes during the identification process. The quality of the evaluation was performed by decoding stimuli with the identified circuits. We showed an increasing SNR in decoding when increasing the number of spikes in the identification process. The simulations were all performed on a GPU platform; it took about 3 months for a massively parallel circuit with 30,000 neurons to be identified.

- Lazar, A.A., Slutskiy, Y.B. and Zhou, Y., Systems and Methods for Time Encoding and Decoding Machines", *patent application* submitted to USPTO on March 17, 2014.
- Lazar, A.A., Slutskiy, Y.B. and Zhou, Y., "Massively Parallel Neural Circuits for Stereoscopic Color Vision: Encoding, Decoding and Identification", in preparation for submission to *Neural Networks*.

2 Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain

To enable the development of an accurate model of the entire fly brain, we released during March 2014 the design of a groundbreaking open source Python platform called Neurokernel that enables collaborative development of comprehensive fly brain models and their execution and testing on multiple Graphics Processing Units (GPUs). Neurokernel also empowers researchers to leverage additional GPU resources and future improvements in GPU technology to accelerate model execution to the same time scale as a live fly brain; this will enable in vivo validation of Neurokernel-based models against real-time recordings of live brain activity.

The scope of Neurokernel's goals and the need to support the revision of models in light of new data requires a structured means of advancing the framework's design. To this end, the Neurokernel project employs Requests for Comments documents (RFCs) as a tool both for advancing the design of Neurokernel architecture and the LPU models built with it. Neurokernel RFCs are publicly accessible design proposals for major new components or features of the fly brain and are available on the project website.

Neurokernel's model support architecture described in RFC #1 was motivated by the organization of the fly brain into fewer than 50 functional modules called local processing units (LPUs) that are each characterized by a unique population of local neurons. By defining communication interfaces that specify how spikes and neuron membrane states are transmitted between LPUs, Neurokernel enables collaborative development and refinement of whole-brain emulations by integration of independently developed processing units. We have demonstrated Neurokernel's module interfacing feature by using it to integrate independently developed models of olfactory and vision LPUs based upon experimentally obtained connectivity information.

Vision is critical for the fly for avoiding predators, flight control, mating, etc.. In order to gain insights into the general principles of neural processing of the fruit fly visual system, we modeled in the Neurokernel RFC #2 the first optic neuropil, the Lamina. Our working hypothesis was that, since the Lamina only contains a very small number of neuron types and has spatially structured neural circuits, by characterizing the I/O behavior of the Lamina it may become possible to uncover some of the information processing transformations that take place in the first stage of visual processing in the fly and, possibly, other organisms.

The anatomical structure of the Lamina has been studied extensively and most of the cell types have been identified. A recent study mapped out most of the synaptic connections (i.e., chemical synapses). In addition, genetic techniques have helped to identify some of the neurotransmitters in the optic lobe. Moreover, imaging and recordings paired with genetic manipulation has led to the understanding of some of the functional properties of the cells in the Lamina.

The objective of the Neurokernel RFC #2 then, was to combine all the currently available information regarding the Lamina LPU and to construct a model that can be used in the Neurokernel environment. Towards this goal, we identified the key components in modeling the neural circuits of the Lamina including (i) cartridges as a canonical neural circuit abstraction, and (ii) composition rules among cartridges.

- Givon, L.E. and Lazar, A.A., "Neurokernel: An Open Scalable Architecture for Emulation and Validation of Drosophila Brain Models on Multiple GPUs", Neurokernel Request for Comments, Neurokernel RFC #1, February 2014.
- Lazar, A. A., Ukani, N. H. and Zhou, Y., The Cartridge: A Canonical Neural Circuit Abstraction of the Lamina Neuropil – Construction and Composition Rules, Neurokernel Request for Comments, Neurokernel RFC #2, January 2014.

3 Code for Reproducible Research

We have been continuously expanding our code base for information representation in the time domain, for spike processing and for identifying neural circuits and architectures. In section 3.1 we describe progress with the implementation of the Time Encoding and Decoding (TED) Toolbox. In section 3.2 we describe the Channel Identification Machines (CIM) toolbox. The CUDA Scikit set of high level programming utilities are described in section 3.3.

3.1 Video Time Encoding Machines (Video TEMs)

We released the Video Time Encoding Machines package for encoding and decoding natural and synthetic visual scenes (videos). The Video TEM circuit consist of Gabor or center surround receptive fields followed by Integrate-and-Fire neurons. The decoder supports both the pseudoinverse algorithm and the recurrent neural networks method.

3.2 Channel Identification Machines (CIM) Toolbox

The Channel Identification Machines (CIMs) Toolbox provides a set of algorithms for the functional identification of spiking neural circuits, as well as a step-by-step demonstration of the accompanying source code. The neural circuits initially considered consist of a linear dendritic processing filter in cascade with a spiking neuron model. The CIM algorithms identify the dendritic processing filter and reconstruct its kernel with arbitrary precision.

3.3 CUDA Scikit

CUDA Scikit (toolkit for SciPy) provides Python interfaces to a subset of the functions in the CUDA, CUDART, CUBLAS, and CUFFT libraries distributed as part of NVIDIA's CUDA Programming Toolkit, as well as interfaces to select functions in the basic and premium versions of the CULA Toolkit. In contrast to most existing Python wrappers for these libraries (many of which only provide a low-level interface to the actual library functions), this package uses PyCUDA to provide high-level functions comparable to those in the NumPy package.

4 Workshops Organized

Following up on our past involvement with the Computational Neuroscience Meeting, we have organized the 2013 workshop

- CNS*2013 Workshop on Methods of Information Theory in Computational Neuroscience, July 17-18, 2013, Paris, France.

Overview

Methods originally developed in Information Theory have found wide applicability in computational neuroscience. Beyond these original methods there is a need to develop novel tools and approaches that are driven by problems arising in neuroscience.

A number of researchers in computational/systems neuroscience and in information/communication theory are investigating problems of information representation and processing. While the goals are often the same, these researchers bring different perspectives and points of view to a common set of neuroscience problems. Often they participate in different fora and their interaction is limited.

The goal of the workshop is to bring some of these researchers together to discuss challenges posed by neuroscience and to exchange ideas and present their latest work.

The workshop is targeted towards computational and systems neuroscientists with interest in methods of information theory as well as information/communication theorists with interest in neuroscience.

Participants (Invited)

Dmitri B. Chklovskii, HHMI Janelia Farm, Ashburn, VA.
Pier Luigi Dragotti, Imperial College.
J. Leo van Hemmen, Technical University of Munich.
Lubomir Kostal, Institute of Physiology, Academy of Sciences of the Cech Republic.
Simon Laughlin, Department of Zoology, University of Cambridge.
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Jean-Pierre Nadal, Laboratoire de Physique Statistique, Ecole Normale Supérieure, Paris.
Israel Nelken, ICNC and the Edmond and Lily Safra Center for Brain Sciences, Hebrew University.
Alex Pouget, University of Geneva.
Simon R. Schultz, Department of Bioengineering, Imperial College.
Tatyana O. Sharpee, Computational Neurobiology Laboratory, The Salk Institute.
Lawrence C. Sincich, University of Alabama.
Taro Toyoizumi, Riken Brain Sciences Institute.

We have also organized the following workshop during the 2013 Computational Neuroscience Meeting:

- CNS*2013 Workshop on Methods of System Identification for Studying Information Processing in Sensory Systems, July 17, 2013, Paris, France.

Overview

A functional characterization of an unknown system typically begins by making observations about the response of that system to input signals. The knowledge obtained from such observations can then be used to derive a quantitative model of the system in a process called system identification. The goal of system identification is to use a given input/output data set to derive a function that maps an arbitrary system input into an appropriate output.

In neurobiology, system identification has been applied to a variety of sensory systems, ranging from insects to vertebrates. Depending on the level of abstraction, the identified neural models vary from detailed mechanistic models to purely phenomenological models.

The workshop will provide a state of the art forum for discussing methods of system identification applied to the visual, auditory, olfactory and somatosensory systems in insects and vertebrates.

The lack of a deeper understanding of how sensory systems encode stimulus information has hindered the progress in understanding sensory signal processing in higher brain centers. Evaluations of various systems identification methods and a comparative analysis across insects and vertebrates may reveal common neural encoding principles and future research directions.

The workshop is targeted towards systems, computational and theoretical neuroscientists with interest in the representation and processing of stimuli in sensory systems in insects and vertebrates.

Participants (Invited)

Alex Borst, Max Planck Institute for Neurobiology, Martinsried.
Sonja Grün, Institute of Neuroscience & Inst. for Advanced Simulation, Jülich Research Centre.
Vivek Jayaraman, Janelia Farm, Ashburn, VA.
Mikko I. Juusola, Department of Biomedical Science, University of Sheffield.
Holger G. Krapp, Department of Bioengineering, Imperial College.
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Brian D. McCabe, Department of Pathology and Cell Biology, Columbia University.
Karim G. Oweiss, Department of ECE and Department of Neuroscience, Michigan State Univ.
Jean-Pierre Rospars, INRA Versailles.

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing
Third Year Report *

Aurel A. Lazar
Department of Electrical Engineering
Columbia University, New York, NY 10027

March 14, 2015

Abstract

During the third research period we pioneered a multi-input multi-output neural circuits architecture for *nonlinear* processing and encoding of stimuli in the spike domain. The basic neuron model in these circuits consists of a nonlinear dendritic stimulus processor (DSP) cascaded with a biophysical spike generator (BSG). We established that identifying a single dendritic stimulus processor is mathematically dual to decoding of stimuli encoded by a population of neurons with a bank of dendritic stimulus processors. Applying the duality to neural circuits with arbitrary connectivity, we were able to achieve fundamental sweeping results that we did not expect at the beginning of the grant period.

For single-input multi-output neural circuit models with feedforward, feedback and cross-feedback DSPs cascaded with BSGs, we theoretically analyzed the effect of noise sources on stimulus decoding. Building on the key duality property, we characterized the effect of noise parameters on the precision of the functional identification of the complete neural circuit with DSP/BSG neuron models. Again, these results are fundamental in nature and unique in the theoretical neuroscience literature.

A significant benefit of the above results was our ability to formulate the problem of stimulus encoding with a neural circuit with DSPs as one of generalized sampling (taking noisy measurements). This is of critical interest to both experimentalists and theoreticians alike. The understanding that the problem of neural decoding and functional identification are dual to each other is key to building on either or both. Knowing how many repetitive experiments need to be performed for a precise quantitative identification of Volterra kernels is of great value in neurophysiology. A further qualitative insight of our work is that for neural circuits with arbitrary connectivity, feedforward kernels are typically easier to estimate than feedback kernels. Finally, our finding that key non-linear neural circuits would be tractable for noise analysis was completely unexpected.

We also redesigned the Neurokernel Open API, developed a novel parallel programming model and devised a graph-based platform for constructing and querying fly brain model circuits.

---

* The work reported here was supported by AFOSR under Grant/Contract Number: FA9550-12-1-0232

1 Key Publications

Combined with the already established and well-understood biophysics of action potential generation, past and recent studies suggest that in single neurons, results of nonlinear signal processing within the dendritic tree are encoded into action or graded potentials by a nonlinear *dynamical* system. The resulting phenomenological model of the fundamental building block of the brain – the neuron – is therefore, to a first approximation, Dynamic Nonlinear-Nonlinear (DNN). Such a model takes into account (i) the dynamic nonlinear analog-to-analog stimulus transformation upstream of the spike initiation zone as well as (ii) the dynamic nonlinear analog-to-discrete transformation within the spike initiation zone that maps the aggregate current produced by the tree into a discrete time sequence, or spikes. Both transformations are dynamic in the sense that their outputs depend on both present and past values of their respective inputs.

For a novel DNN neuron model that arises in sensory systems we discovered a key duality result between two closely related decoding and identification problems. The basic neuron model consists of a nonlinear dendritic stimulus processor (DSP) cascaded with a biophysical spike generator (BSG). The nonlinear dendritic processor is modeled as a set of nonlinear operators that are assumed to have a Volterra series representation. Biophysical point neuron models, such as the Hodgkin-Huxley neuron, are used to model the spike generator. We proved that the identification of the dendritic stimulus processor of a single neuron and the decoding of stimuli encoded by a population of neurons with a bank of dendritic stimulus processors are dual to each other. This duality result enabled us to derive lower bounds on the number of experiments to be performed and the total number of spikes that need to be recorded for identifying a neural circuit. For arbitrary architectures we also addressed in detail the question of how intrinsic noise sources affect the precision in encoding and decoding of sensory stimuli and the functional identification of their constituent circuits. Our main results are briefly described in sections 1.1 and 1.2.

1.1 Spiking Neural Circuits with Dendritic Stimulus Processors: Encoding, Decoding, and Identification in Reproducing Kernel Hilbert Spaces

We developed a general multi-input multi-output (MIMO) neural circuit architecture for *nonlinear* processing and encoding of stimuli in the spike domain with DNN neuron models. Both the architecture and the associated methods of analysis are grounded in the rigorous context of the well-established theory of time encoding machines (TEMs). Previously investigated TEMs included various circuit architectures in which linear time-invariant filters appeared in series with nonlinear asynchronous samplers, such as point neuron models. Input signals were thus linearly processed and encoded into a sequence of spikes by these nonlinear samplers.

We devised our results for MIMO neural circuit architectures that exhibit *nonlinear* processing and encoding of stimuli in the spike domain. In these architectures a bank of dendritic stimulus processors implements nonlinear transformations of multiple temporal or spatio-temporal signals such as spike trains or auditory and visual stimuli in the analog domain. Dendritic stimulus processors may act on both individual stimuli and on groups of stimuli, thereby executing complex computations that arise as a result of interactions between concurrently received signals. The results of the analog-domain computations are then encoded into a multi-dimensional spike train by a population of spiking neurons modeled as nonlinear dynamical systems. We investigated general conditions under which such circuits faithfully represent stimuli and demonstrated algorithms for (i) stimulus recovery, or decoding, and (ii) identification of dendritic stimulus processors from the observed spikes.

DSPs were modeled using Volterra series. Volterra series have been used for analyzing nonlinear neuronal responses in many contexts, and have been applied to the identification of single neurons in many of sensory areas. Volterra dendritic processors can model a wide range of nonlinear effects commonly seen in sensory systems. We successfully extended the theory of time encoding to *Volterra TEMs* that employ a bank of dendritic stimulus processors to implement nonlinear computations in the analog domain. We investigated conditions under which such circuits faithfully represent stimuli and presented *Volterra TDMs* that can decode the original stimuli, given the spikes at the output of a Volterra TEM. We also developed *Volterra Channel Identification Machines* (Volterra CIMs), an extension of traditional CIMs that allow one to identify nonlinear dendritic processing in neural circuits.

- Lazar, A.A. and Slutskiy, Y.B., "Spiking Neural Circuits with Dendritic Stimulus Processors: Encoding, Decoding, and Identification in Reproducing Kernel Hilbert Spaces", Journal of Computational Neuroscience, Volume 38, Number 1, pp. 1-24, February 2015.

1.2 Volterra Dendritic Stimulus Processors and Biophysical Spike Generators with Intrinsic Noise Sources

We investigated two intrinsic noise sources arising (i) in the active dendritic trees underlying the DSPs, and (ii) in the ion channels of the BSGs. Noise in dendritic stimulus processing arises from a combined effect of variability in synaptic transmission and dendritic interactions. Channel noise arises in the BSGs due to the fluctuation of the number of the active ion channels. Using a stochastic differential equations formalism we showed that encoding with a neuron model consisting of a nonlinear DSP cascaded with a BSG with intrinsic noise sources can be treated as generalized sampling with noisy measurements.

For single-input multi-output neural circuit models with feedforward, feedback and cross-feedback DSPs cascaded with BSGs we theoretically analyzed the effect of noise sources on stimulus decoding. Building on a key duality property, the effect of noise parameters on the precision of the functional identification of the complete neural circuit with DSP/BSG neuron models was given. We demonstrated through extensive simulations the effects of noise on encoding stimuli with circuits that include neuron models that are akin to those commonly seen in sensory systems, e.g., complex cells in V1.

Our nonlinear dendritic stimulus processors describe functional I/O relationships between the dendritic outputs in the first stage and inputs that are either sensory stimuli or spikes generated by BSGs. DSPs were again modeled using Volterra series. Here, in addition, we introduced nonlinear interactions between the spike trains of arbitrary neuron pairs. This gives rise to interesting neural processing capabilities directly in the spike domain, *e.g.*, coincidence detection.

We systematically investigated how the strength of noise sources degrades the faithfulness of stimulus representation and the quality of functional identification for our proposed class of neural circuit models. Furthermore, since the representation is based on spike timing, we characterized the ways in which spike timing variability affects the precision in representing the amplitude information of sensory stimuli.

- A. A. Lazar and Y. Zhou, "Volterra Dendritic Stimulus Processors and Biophysical Spike Generators with Intrinsic Noise Sources", Frontiers in Computational Neuroscience, Volume 8, Number 95, pp. 1-24, September 2014.

1.3 Reconstructing Natural Visual Scenes from Spike Times

As part of our grant, we developed in the previous years a formal theory of Video Time Encoding Machines (Video TEMs) with simplified spiking neuron models, such as Leaky-Integrate-and-Fire (LIF) and Threshold-and-Fire neurons with feedback. We now extended these architectures for encoding natural visual scenes with more complex (but tractable) models of dendritic stimulus processing and biophysical spike generation (BSG).

We put forth a biologically relevant neural circuit architecture that consists of ON and OFF component BSGs, that are inspired by the ON and OFF cells present in the early visual system. The ON-OFF BSGs with self- and cross-feedback that we formally analyzed are of key interest in building neuromorphic silicon retinas. Event driven silicon components allow for high dynamic range in imaging natural visual environments whose light intensity can vary several orders of magnitude in a short spatial range. Finally, for the extended Video TEM architectures that we proposed, we demonstrated that neural encoding can readily be interpreted as generalized sampling.

The decoding approach of visual scenes encoded with ON-OFF BSGs is based on using phase response curves (PRCs). With the help of a manifold of PRCs, decoding is tractable even for a wide input signal dynamic range while the I/O equivalence properties of spike generators is preserved. As a result, bias currents that were essential in the encoding process formulated previously can largely be reduced or eliminated. Taken together, these extensions relax or even eliminate the need of bias currents in biophysical spike generators and further reduce the power consumption in silicon-based retinas.

We also provided Nyquist-type rate conditions on visual signal reconstruction. For noisy neural circuits we demonstrated a graceful degradation of signal recovery. Finally, we presented examples of massively parallel encoding of natural visual scenes with high performance in recovery. Such decoding methods provide an indicator of how well the information is represented in the spike domain and whether visual scenes can be reconstructed unambiguously from spike times.

- A. A. Lazar and Y. Zhou, "Reconstructing Natural Visual Scenes from Spike Times", Proceedings of the IEEE, Volume 102, Number 10, pp. 1500-1519, October 2014.

1.4 Neural Circuits for Encoding with Complex Cells: Rank Minimization for Decoding and Identification

We investigated the capability of a population of complex cells to encode space-time varying stimuli. Similar to simple cells, complex cells in the primary visual cortex (V1) are selective to oriented edges/lines over a limited region of the visual field. While simple cells respond maximally to a particular phase of the edge, complex cells are largely phase invariant. Therefore, the receptive fields of complex cells cannot be simply subdivided into excitatory and inhibitory regions. Receptive fields of simple cells can be modeled as linear Gabor filters while processing in complex cells can be modeled with a quadrature pair of Gabor filters followed by squaring.

We showed that processing in complex cells has the complexity of Volterra models briefly described above. General Volterra based models, however, run into the curse of dimensionality. We have demonstrated, however, that complex cells have a processing structure that can be exploited and decoding and identification problems can be efficiently solved. The reconstruction algorithms were formulated as rank minimization problems. We (i) implemented an algorithm that reconstructs the visual stimuli based on the spike times generated by neural circuits with widely employed complex cells models, and (ii) developed an mechanistic algorithm for functionally identifying the processing in complex cells using the spike times they generate.

These algorithms are based on the key observation that functional identification of processing in complex cells can be viewed as the dual problem of decoding stimuli that are encoded by an ensemble of complex cells. We have also shown how to identify energy processing in complex cells and demonstrated that our identification algorithms substantially outperform spike-triggered covariance based algorithms. Finally, we evaluated our identification algorithms by reconstructing novel stimuli in the input space. Currently a massive evaluation of these algorithms is under way.

- A.A. Lazar, N.H. Ukani and Y. Zhou, "Neural Circuits for Encoding with Complex Cells Rank Minimization for Decoding and Identification", in preparation for submission.

1.5 Second Year Report Papers Published

We followed through with the submission and publication of the following papers that were highlighted in our "Second Year Report". They are included here for completeness:

- Lazar, A.A. and Slutskiy, Y.B., "Channel Identification Machines for Multidimensional Receptive Fields", *Frontiers in Computational Neuroscience*, Special Issue on the Role of Mathematics and Theory in Understanding the Brain, Volume 8, Number 117, September 2014.
- Lazar, A.A., Slutskiy, Y.B. and Zhou, Y., "Massively Parallel Neural Circuits for Stereoscopic Color Vision: Encoding, Decoding and Identification", *Neural Networks*, Volume 63, pp. 254-271, March 2015.

2 Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain

Successfully implementation of fly brain models *in silico* requires integration and execution of local processing units (LPUs) and tract models developed independently by multiple researchers. Software that enables such integration must also afford researchers the freedom to specify the parameters of LPU models that may employ disparate internal designs. Ongoing efforts to obtain increasingly accurate data regarding the fly's brain structure also demand that such software be sufficiently flexible to facilitate revision of the models in a straightforward manner even if completely new components must be added. Existing tools for structured specification of neural models provide limited means for querying model data required by emulation engines responsible for efficient model execution. Moreover, changes to an LPU model's internal design may complicate its reimplementation by requiring that support for accessing new design elements be retrofitted into representations of the existing model.

Towards this end we (i) redesigned, implemented and published a new Open API between LPUs, (ii) developed a novel parallel programming model for designing LPUs, and (iii) devised a graph-based platform called Neuroarch for constructing and querying fly brain model circuits.

2.1 Neurokernel Open API

We redesigned our initial implementation of the groundbreaking Neurokernel open-source platform to enable enhanced collaborative development of comprehensive fly brain models and their substantially faster execution and testing on multiple Graphics Processing Units (GPUs). The Neurokernel open-source code is freely available at the Github website: http://neurokernel.github.io/index.html.

By defining a standard communication interface (API) that all neuropil models (LPUs) must use, Neurokernel enables now independently developed neuropil and connectivity pattern models with compatible interfaces to be integrated into a single executable model irrespective of their internal design. To reduce interference between model development and addition of new features to the software, Neurokernel's architecture separates its application plane (which provides support for neuropil model and connectivity pattern specification tools) from its control and compute planes (which respectively provide resource management and GPU-based digital machinery required by neuropil models). As before, the development of the Neurokernel Project is based upon the well-known and highly successful collaborative model of the IETF's Requests for Comments.

With the new API implementation, Neurokernel has been used to connect and execute proof-of-concept models of select neuropils in the fly's olfactory and vision sensory systems that have been independently developed by different teams. Implementations of these models and the core Neurokernel source code are both available online. Neurokernel's support for interfacing independently developed neuropil models and executing them on commodity parallel hardware provides neuroscience researchers with a fundamentally collaborative platform upon which to join forces to achieve the ultimate goal of building the fly brain *in silico*. Researchers worldwide started to take notice and expressed interest in collaborating with us including the research group around Prof. Daniel Coca at the University of Sheffield, Great Britain and Prof. Chung-Chuan Lo at the Institute of Systems Neuroscience and Department of Life Sciences, National Tsing Hua University, Taiwan.

- L.E. Givon, A.A. Lazar, K. Psychas, N.H. Ukani, C.-H. Yeh, and Y. Zhou, “Neurokernel: Building an in Silico Fruit Fly Brain”, IEEE EMBS BRAIN Grand Challenges Conference, IEEE, November 2014.

2.2 A Parallel Programming Model of LPUs in the Fruit Fly Brain

To identify the I/O transformations taking place in LPUs, we developed a parallel programming model for exploration of high level LPU algorithms. The programming model takes into account existing anatomical observations of the underlying neuronal circuits of the fly brain neuropils and has been demonstrated on a range of sensory LPUs, including those in vision and in olfaction. The programing model comprises canonical neural circuit abstractions (cNCAs) and composition rules (CRs) among them. These are briefly described below.

cNCAs are fundamental computational units in many LPUs. For example, cartridges and columns are cNCAs in the lamina and medulla LPUs, respectively, of the early visual system, and channels are cNCAs in the antennal lobe of the early olfactory system. Each LPU comprises tens, if not hundreds of its respective cNCAs. The cNCAs, by themselves, implement a particular algorithm that performs local computations. Such an algorithm may utilize multiple neurons that each perform, in turn, elementary operations. Using cNCAs rather than individual neurons as computational units highlights the preeminence of circuit building blocks underlying neural computation over elementary neuronal operations and it is akin to using threads in parallel programming.

Communication among cNCAs is achieved by composition rules. CRs are global algorithms that are performed asynchronously and are implemented by a few neurons. By enabling interaction among cNCAs, the CRs facilitate the design of algorithms that can use locally processed information to achieve computation on different spatial brain scales. For example, spatially restricted information that is individually processed by cNCAs can be compared using CRs to implement motion detection algorithms in which computation between spatially displaced visual areas is essential. Thus, CRs are critical for the diversity and functionality of algorithms that can be realized using the programming model proposed here.

- A.A. Lazar, N.H. Ukani, C.-H. Yeh, and Y. Zhou, “A Parallel Programming Model of Local Processing Units in the Fruit Fly Brain”, Frontiers in Neuroinformatics, Number 24, August 2014, Leiden, The Netherlands.

2.3 Neuroarch: A Graph-Based Platform for Constructing and Querying Models of the Fruit Fly Brain Architecture

We have developed a Python package called Neuroarch for representation and storage of LPU-based models of the fly brain. It provides researchers and software applications with a common interface for defining, querying, and manipulating integrated model data. Neuroarch’s representation of the fly’s brain distinguishes between the connectivity architecture linking its LPUs and the design of the individual LPUs previously identified; the former is modeled as graphs of communication ports exposed by each LPU and the connections between them while the latter comprises graphs of the internal modeling elements required by specific LPU models. Neuroarch’s flexibility stems from its storage of all modeling elements comprised by the connectivity architecture and different LPU designs (including synaptic model instances) as graph nodes; edges are exclusively used to represent relationships between modeling elements. Neuroarch stores all model data in a graph database to accelerate those queries typically performed to determine which modeling elements must be updated simultaneously during model execution.

Both LPU design and inter-LPU connectivity model data is exclusively accessed through Neuroarch’s object-relational mapping (ORM) of modeling elements to their internal database representations. A key feature of this ORM is its support for multimodal views of query results over stored model data that may be passed as operands to other graph operations. This enables model data to be accessed or modified either as a subgraph (to facilitate graph-based queries) or a tensor (to facilitate tabular or relational queries). Neuroarch's I/O layer extends and invokes the ORM to load model data expressed in several specification formats such as CSV, GEXF, and NeuroML. Models may also be constructed from basic circuit motifs (specified either as small graphs or tensors) using Neuroarch's graph composition operators.

We have used Neuroarch to drive fly brain emulations of a prototype multisensory coincidence detection system that integrates 4 independently developed LPUs in the fly vision and olfactory systems executed using the Neurokernel framework; model data for some of these LPUs is explicitly specified, while the remaining LPUs are constructed using Neuroarch's graph operators to compose canonical circuits identified in those LPUs.

- L.E. Givon, A.A. Lazar, and N. H. Ukani, "Neuroarch: A Graph-Based Platform for Constructing and Querying Models of the Fruit Fly Brain Architecture", Frontiers in Neuroinformatics, Number 42, August 2014, Leiden, The Netherlands.

3 Workshops Organized

3.1 CNS*2014 Workshop on Methods of Information Theory in Computational Neuroscience

Following up on our past involvement with the Computational Neuroscience Meeting, we organized the 2014 workshop

- CNS*2014 Workshop on Methods of Information Theory in Computational Neuroscience, July 30-31, 2014, Québec City, Canada.

Overview

Methods originally developed in Information Theory have found wide applicability in computational neuroscience. Beyond these original methods there is a need to develop novel tools and approaches that are driven by problems arising in neuroscience.

A number of researchers in computational/systems neuroscience and in information/communication theory are investigating problems of information representation and processing. While the goals are often the same, these researchers bring different perspectives and points of view to a common set of neuroscience problems. Often they participate in different fora and their interaction is limited.

The goal of the workshop is to bring some of these researchers together to discuss challenges posed by neuroscience and to exchange ideas and present their latest work.

The workshop is targeted towards computational and systems neuroscientists with interest in methods of information theory as well as information/communication theorists with interest in neuroscience.

Participants (Invited)

Chris DiMattina, Department of Psychology, Florida Gulf Coast University.
Maurice Chacron, Faculty of Medicine, McGill University.
Alexander G. Dimitrov, Department of Mathematics, Washington State University - Vancouver.
Jean F. Lienard, Oregon Hearing Research Center, Oregon Health & Science University.
Andre Longtin, Physics Department, University of Ottawa.
Lav R. Varshney, Department of Electrical and Computer Engineering, UIUC.
Mike DeWeese, Physics Department and the Helen Wills Neuroscience Institute, UC Berkeley.
Byron Yu, Electrical, Computer and Biomedical Engineering, Carnegie Mellon University.
Kechen Zhang, Department of Biomedical Engineering, Johns Hopkins University.

3.2 CNS*2014 Workshop on Methods of System Identification for Studying Information Processing in Sensory Systems

We have also organized the following workshop during the 2014 Computational Neuroscience Meeting:

- CNS*2014 Workshop on Methods of System Identification for Studying Information Processing in Sensory Systems, July 30, 2014, Québec City, Canada.

Overview

A functional characterization of an unknown system typically begins by making observations about the response of that system to input signals. The knowledge obtained from such observations can then be used to derive a quantitative model of the system in a process called system identification. The goal of system identification is to use a given input/output data set to derive a function that maps an arbitrary system input into an appropriate output.

In neurobiology, system identification has been applied to a variety of sensory systems, ranging from insects to vertebrates. Depending on the level of abstraction, the identified neural models vary from detailed mechanistic models to purely phenomenological models.

The workshop will provide a state of the art forum for discussing methods of system identification applied to the visual, auditory, olfactory and somatosensory systems in insects and vertebrates.

The lack of a deeper understanding of how sensory systems encode stimulus information has hindered the progress in understanding sensory signal processing in higher brain centers. Evaluations of various systems identification methods and a comparative analysis across insects and vertebrates may reveal common neural encoding principles and future research directions.

The workshop is targeted towards systems, computational and theoretical neuroscientists with interest in the representation and processing of stimuli in sensory systems in insects and vertebrates.

Participants (Invited)

Claude Desplan, Department of Biology, NYU.
Mark A. Frye, Department of Integrative Biology and Physiology, UCLA.
Mikko I. Juusola, Department of Biomedical Science, University of Sheffield.
Arvind Kumar, Bernstein Center Freiburg, University of Freiburg.
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Brian D. McCabe, Department of Pathology and Cell Biology, Columbia University.
Stefan Mihalas, Allen Institute for Brain Science.
Johannes Reisert, Monell Chemical Senses Center, Philadelphia, PA.
Tatyana O. Sharpee, The Computational Neurobiology Laboratory, Salk Institute.
Glenn C. Turner, Cold Spring Harbor Laboratory.

3.3 Columbia Workshop on Brain Circuits, Memory and Computation

Finally, during March 2015 I organized the following workshop at Columbia University:

- Columbia Workshop on Brain Circuits, Memory and Computation (BCMC 2015), March 16-17, 2015, New York, NY.

Overview

The goal of the workshop is to bring together researchers interested in developing executable models of neural computation/processing of the brain of model organisms. Of interest are models of computation that consist of elementary units of processing using brain circuits and memory elements. Elementary units of computation/processing include population encoding/decoding circuits with biophysically-grounded neuron models, non-linear dendritic processors for motion detection/direction selectivity, spike processing and pattern recognition neural circuits, movement control and decision-making circuits, etc. Memory units include models of spatio-temporal memory circuits, circuit models for memory access and storage, etc. A major aim of the workshop is to explore the integration of various computational sensory and control models.

Participants (Invited)

Yoshi Aso, Ph.D., Rubin Lab, Janelia Research Campus, Ashburn, Virginia, USA.
Dmitri "Mitya" B. Chklovskii, Simons Center for Data Analysis, Simons Foundation.
Damon A. Clark, Department of Molecular, Cellular, and Developmental Biology, Yale University.
Daniel Coca, Department of Automatic Control and Systems Engineering, University of Sheffield.
Fabrizio Gabbiani, Dept. of Neurosci., Baylor College of Med., and Applied Math, Rice University.
Charles Randy Gallistel, Rutgers Center for Cognitive Science, Rutgers University.
Charles D. Gilbert, Laboratory of Neurobiology, Rockefeller University.
Vivek Jayaraman, Janelia Research Campus, Ashburn, VA.
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Anthony Leonardo, Janelia Research Campus, Ashburn, VA.
Wolfgang Maass, Institute of Computer Science, Graz University of Technology, Graz, Austria.
Gary F. Marcus, Department of Psychology, New York University.
Brian D. McCabe, Department of Pathology and Cell Biology, Columbia University.
Stefan Mihalas, Allen Institute for Brain Science, Seattle, WA.
Barbara Webb, School of Informatics, University of Edinburgh.

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing
Fourth Year Report *

Aurel A. Lazar
Department of Electrical Engineering
Columbia University, New York, NY 10027

April 8, 2016

Abstract

During the fourth research period we devised a motion detection algorithm based on local phase information and demonstrated a fast, parallel algorithm for its real-time implementation. Our results suggest that local spatial phase information may provide an efficient alternative to perform many visual tasks *in silico* as well as *in vivo* biological vision systems.

We released the first stable Neurokernel open-source platform during the Fruit Fly Brain Hackathon that took place at Columbia University on March 17, 2016. We also designed a graph database, called NeuroArch, for codifying knowledge about fruit fly brain circuits. The NeuroArch database provides open fly brain datasets for designing executable circuit models that can be evaluated with the Neurokernel open-source platform.

---

[*] The work reported here was supported by AFOSR under Grant/Contract Number: FA9550-12-1-0232

1 Key Publications

The key publications during the fourth research period pertain to (i) the first stable release of the Neurokernel platform and some of its extensions, and (ii) a novel class of motion detection algorithms that opens new directions for biologically inspired computing algorithms.

1.1 A Motion Detection Algorithm Using Local Phase Information

In this paper, we revealed the power of phase in representation and processing of visual stimuli. We first presented an algorithm that uses local phase information of an image to perfectly reconstruct the image. The computation of phase is of highly nonlinear nature, yet the reconstruction can be formulated in a similar fashion as in the Video Time Decoding Machine except that it operates in the analog domain without spikes.

After demonstrating that local phase can be used for the representation of visual stimuli, we took a step forward and argue that not only can phase information represent information, but it can be utilized in the processing of visual stimuli, such as in detecting local motion. We proposed a local phase-based motion detection algorithm that consists of two building blocks. The first building block measures the temporal change of the local phase. The temporal derivative of the local phase is shown to exhibit the structure of a second order Volterra kernel with two normalized inputs, and thereby closely related to biological motion detector such as the Reichardt motion detector. We provide an efficient, FFT-based algorithm for implementing the change of the local phase. The second processing building block implements the detector. The local phase-based motion detection algorithm compares the maximum of the Radon transform of the local phase derivative with a chosen threshold.

We demonstrated examples of applying the local phase-based motion detection algorithm on several video sequences. The proposed algorithm was shown to outperform the Reichardt motion detector and Barlow-Levick motion detector while it can be efficiently implemented on parallel hardware in a way comparable to the other two detectors. In particular, we showed that the local phase-based motion detection algorithm perform much better in lower contrast and luminance conditions. We also showed how the locally detected motion can be used for segmenting moving objects in video scenes and compare our local phase-based algorithm to segmentation achieved with a widely used optic flow algorithm.

The results in the paper open up a few new avenues for future investigation. First, our results suggest that local spatial phase information may provide an efficient alternative to perform many visual tasks. Phase processing machines can be an effective strategy that sensory systems employ. Second, the motion detection algorithm can be mapped to the motion detection circuit in the fruit fly visual system.

- Lazar, A.A., Ukani, N.H. and Zhou, Y., "A Motion Detection Algorithm Using Local Phase Information", Journal of Computational Intelligence and Neuroscience, Volume 2016, January 2016.

1.2 Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain

We have developed an open software platform called Neurokernel for collaborative development of comprehensive models of the brain of the fruit fly *Drosophila melanogaster* and their execution and testing on multiple Graphics Processing Units (GPUs). Neurokernel provides a programming model that capitalizes upon the structural organization of the fly brain into a fixed number of functional modules to distinguish between these modules' local information processing capabilities and the connectivity patterns that link them. By defining mandatory communication interfaces that specify how data is transmitted between models of each of these modules regardless of their internal design, Neurokernel explicitly enables multiple researchers to collaboratively model the fly's entire brain by integration of their independently developed models of its constituent processing units. We demonstrate the power of Neurokernel's model integration by combining independently developed models of the retina and lamina neuropils in the fly's visual system and by demonstrating their neuroinformation processing capability. We also illustrate Neurokernel's ability to take advantage of direct GPU-to-GPU data transfers with benchmarks that demonstrate scaling of Neurokernel's communication performance both over the number of interface ports exposed by an emulation's constituent modules and the total number of modules comprised by an emulation.

- L.E. Givon and A.A. Lazar "Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain", PLOS ONE, January 2016.

1.3 Identifying Multisensory Dendritic Stimulus Processors (Invited Paper)

Functional identification is a key methodology for uncovering the logic of neuroinformation processing of brain circuits. For neural circuits modeling sensory systems as dynamical systems, the identification algorithm largely depends on the number of stimuli used. Neurons in these circuits consist of dendritic stimulus processors modeling the dendritric tree and biological spike generators modeling the axon hillock. Once the dendritic stimulus processor of a biological neuron is identified, its deviation from the "ground thruth" can not be evaluated as the latter is typically not available. As a remedy, the deviation of the response of the identified neuron from the response of its biological counterpart is often measured in the space of spikes. Measures defined in the space of spikes, however, lack intuition, are error prone and are difficult to interpret.

In this paper we review the identification of multi-sensory spatio-temporal dendritic stimulus processors that arise in the encoding of color visual fields and the mixing of natural visual fields and auditory scenes. We demonstrate the fundamental duality between the identification of the dendritic stimulus processor of a single neuron and the decoding of stimuli encoded by a population of neurons with a bank of dendritic stimulus processors. The duality enables us to reconstruct the originally encoded stimuli from all the generated spikes by using the identified neural circuit. The reconstruction leads to a simple and intuitive evaluation of the identified dendritic stimulus processors in the space of stimuli.

- A.A. Lazar and Y. Zhou, "Identifying Multisensory Dendritic Stimulus Processors, Invited Paper", IEEE Transactions on Molecular, Biological and Multi-Scale Communications, Special Issue on Biological Applications of Information Theory - Claude Shannon's Centennial, 2016 (submitted).

2 Neurokernel Requests For Comments

2.1 A Parallel Processing Model of the Drosophila Retina

In this RFC, we presented a detailed description of the algorithms required for a full-scale parallel emulation of the retina. We first describe the mapping of the visual field onto the photoreceptors, including the geometry of the retina with respect to the visual field and the modeling of input to photoreceptors as a function of the visual field. We then characterized the I/O of single photoreceptors based on a transduction model that closely follows the state-of-the-art in the literature. Finally, we presented the parallel information processing of the entire retina and visually evaluated the overall response of the retina.

The full-scale parallel emulation of the retina is critical for the spatio-temporal I/O characterization of retina pre-processing. We argue that the visual evaluation provides a highly intuitive way to explore parallel visual processing in the retina and similarly in the investigation of processing in subsequent neuropils. Based on the RFC, we submitted an abstract to COSYNE 2016 and presented a poster in the conference.

Moreover, we made available with the RFC a GPU implementation that is fully compatible with the Neurokernel API.

- A.A. Lazar, K. Psychas, N.H. Ukani and Y. Zhou, "A Parallel Processing Model of the Drosophila Retina", Neurokernel Request for Comments, Neurokernel RFC #3, August 2015.

2.2 Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain

We have developed an open software platform called Neurokernel for collaborative development of comprehensive models of the brain of the fruit fly *Drosophila melanogaster* and their execution and testing on multiple Graphics Processing Units (GPUs). Neurokernel provides a programming model that capitalizes upon the structural organization of the fly brain into a fixed number of functional modules to distinguish between these modules' local information processing capabilities and the connectivity patterns that link them. By defining mandatory communication interfaces that specify how data is transmitted between models of each of these modules regardless of their internal design, Neurokernel explicitly enables multiple researchers to collaboratively model the fly's entire brain by integration of their independently developed models of its constituent processing units. We demonstrate the power of Neurokernel's model integration by combining independently developed models of the retina and lamina neuropils in the fly's visual system and by demonstrating their neuroinformation processing capability. We also illustrate Neurokernel's ability to take advantage of direct GPU-to-GPU data transfers with benchmarks that demonstrate scaling of Neurokernel's communication performance both over the number of interface ports exposed by an emulation's constituent modules and the total number of modules comprised by an emulation.

- L. E. Givon, A. A. Lazar, "Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain", Neurokernel Request for Comments, Neurokernel RFC #4, October 2015.

2.3 NeuroArch: A Graph dB for Querying and Executing Fruit Fly Brain Circuits

NeuroArch is a database for codifying knowledge about fruit fly brain circuits. It is designed with two user communities in mind: (i) neurobiologists interested in querying the database to address questions regarding neuroanatomy, neural circuits, neurons, synapses, neurotransmitters, and gene expression, and (ii) computational neuroscientists interested in the instantiation of models of neural circuits and architectures, their program execution, and validation of hypotheses regarding brain function. A key aim of NeuroArch is to provide an interface between the concerns of these two communities. To this end, NeuroArch defines a data model for representation of both biological data and model structure and the relationships between them within a single graph database. When coupled with a powerful interface for querying both types of data within the database in a uniform manner, this representation enables neurobiologists to benefit from high-level organization of fruit fly data, while model designers can capitalize upon the integration of biological data from multiple sources. This document describes the requirements and design of NeuroArch, and details how it can be used with the Neurokernel framework to accelerate collaborative development of models of the fruit fly brain. A brief discussion of future development plans is also provided.

- L. E. Givon, A. A. Lazar, and N. H. Ukani, "NeuroArch: A Graph dB for Querying and Executing Fruit Fly Brain Circuits", Neurokernel Request for Comments, Neurokernel RFC #5 , December 2015.

2.4 Generating an Executable Model of the Drosophila Central Complex

The central complex (CX) is a set of neuropils in the center of the fly brain that have been implicated as playing an important role in vision-mediated behavior and integration of spatial information for locomotor control. In contrast to currently available data regarding the neural circuitry of neuropils in the fly's vision and olfactory systems, comparable data for the CX neuropils is relatively incomplete; many categories of neurons remain only partly characterized, and the synaptic connectivity between CX neurons has yet to be experimentally determined. Successful modeling of the information processing functions of the CX neuropils therefore requires a means of easily constructing and testing a range of hypotheses regarding both the high-level structure of their neural circuitry and the properties of their constituent neurons and synapses. This document demonstrates how NeuroArch and Neurokernel may be used to programmatically construct and evaluate executable neural circuit models of the CX neuropils and their interconnects based upon currently available information regarding the geometry and polarity of the arborizations of identified local and projection neurons in the CX.

- L. E. Givon, A. A. Lazar, "Generating an Executable Model of the Drosophila Central Complex", Neurokernel Request for Comments, Neurokernel RFC #6, April 2016.

2.5 Neurokernel RFC for the Fruit Fly Medulla LPU - Work in Progress

Having completed the first two neuropils of the fruit fly visual system, we started to investigate the medulla in search for "the holy grail" of fly motion vision - the medulla circuit that leads to motion detection.

Using the EM data of 7 columns of medulla that was made available by the Janelia Research Campus, we performed initial analysis of the connections, including 1) the renaming of neurons based on standard convention, 2) identification/classification of a few unidentified neuron types, 3) visualization of each type of neurons and 4) high level diagram of the major pathway in the motion detection circuit. We also developed several tools to investigate circuits in the medulla data. For visualization, we automated the SWC file generation to visualize individual neurons with their pre- and post-synaptic sites. We also automated the selection of files such that the input and output circuit of a neuron can be easily visualized. In addition, we developed automated graph visualization to investigate composition rules. This tool allowed us to intuitively identify the input pattern of the motion sensitive T4 neurons.

Since the medulla data only contains neurites of neurons within 7 columns, many tangentially arborizing neurons were not fully reconstructed. We plan to study some of these neurons with the help of meso-scale data from the Flycircuit database and some more recent papers that outline the structure and tiling of these non-columnar neurons.

The characterization of the medulla circuit has been preliminarily summarized in the "Fruit Fly Brain Medulla: From Structure to Function", RFC #9 (draft). We plan to substantially expanded the draft RFC # 9 and finish its implementation in the next few months.

- A.A. Lazar and Y. Zhou, "Fruit Fly Brain Medulla: From Structure to Function", Neurokernel Request for Comments, Neurokernel RFC #9, April 2016 (draft).

3 Workshops Organized

3.1 CNS*2015 Workshop on Methods of System Identification for Studying Information Processing in Sensory Systems

During the 2015 Computational Neuroscience Meeting, I organized jointly with Mikko I. Juusola the:

- CNS*2015 Workshop on Methods of System Identification for Studying Information Processing in Sensory Systems, July 22, 2015, Prague, Czech Republic.

Overview

A functional characterization of an unknown system typically begins by making observations about the response of that system to input signals. The knowledge obtained from such observations can then be used to derive a quantitative model of the system in a process called system identification. The goal of system identification is to use a given input/output data set to derive a function that maps an arbitrary system input into an appropriate output.

In neurobiology, system identification has been applied to a variety of sensory systems, ranging from insects to vertebrates. Depending on the level of abstraction, the identified neural models vary from detailed mechanistic models to purely phenomenological models.

The workshop provided a state of the art forum for discussing methods of system identification applied to the visual, auditory, olfactory and somatosensory systems in insects and vertebrates.

The lack of a deeper understanding of how sensory systems encode stimulus information has hindered the progress in understanding sensory signal processing in higher brain centers. Evaluations of various systems identification methods and a comparative analysis across insects and vertebrates may reveal common neural encoding principles and future research directions.

The workshop was targeted towards systems, computational and theoretical neuroscientists with interest in the representation and processing of stimuli in sensory systems in insects and vertebrates.

Participants (Invited)

Martin Egelhaaf, Department of Neurobiology, University of Bielefeld.
Eugenia Chiappe, Champalimaud Neuroscience Programme, Champalimaud, Lisbon.
C. Giovanni Galizia, Department of Zoology and Neurobiology, University of Konstanz.
Mikko I. Juusola , Department of Biomedical Science, University of Sheffield.
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Chung-Chuan Lo, Department of Life Sciences, National Tsing Hua University, Taiwan.
Dong Song, Department of Biomedical Engineering, USC.
Andrew D. Straw, Research Institute of Molecular Pathology, Vienna.
Thomas Nowotny, Department of Informatics, University of Sussex.

3.2 CNS*2015 Workshop on Open Collaboration in Computational Neuroscience

Following up on my past involvement with the Computational Neuroscience Meeting, I jointly organized with Padraig Gleeson the 2015 workshop

- CNS*2015 Workshop on Open Collaboration in Computational Neuroscience, July 23, 2015, Prague, Czech Republic.

Overview

Building and analyzing biophysically and anatomically detailed neuronal networks is a complex and time consuming task, which ideally involves researchers with a range of backgrounds and technical skills. However, most labs cannot expect to have all of these researchers present at any given time. This can lead to stalled projects, lost data/software and needless repetition of experimental and computational work.

A number of initiatives have been started which address these issues. Some are creating public resources with freely available data to constrain such models. Others are using best practices from open source software development to encourage building and sharing of models in a collaborative environment. This workshop served as an informative introduction to these projects as well as a discussion forum for getting feedback and gathering requirements from the community for the developers of these initiatives.

Participants (Invited)

Jan Antolik, UNIC, UPR CNRS, Paris.
Nicholas Cain, Allen Brain Institute, Seattle.
Matteo Cantarelli, MetaCell Ltd.
Chaitanya Chintaluri, Nencki Institute of Experimental Biology, Warsaw.
Padraig Gleeson, University College London.
Victor Jirsa, CNRS, Marseille.
Aurel A. Lazar, Department of Electrical Engineering, Columbia University.
Bill Lytton, SUNY Downstate Medical Center, Brooklyn, NY.
Eilif Muller, EPFL, Lausanne.
Adrian Quintana, University College London.
Michael Sonntag, Ludwig-Maximilians-Universität (LMU), Munich.
Shreejoy Tripathy, University of British Columbia, Vancouver.

3.3 Fruit Fly Brain Hackathon 2016

I raised funding support from IBM Research for the first Fruit Fly Brain Hackathon. The hackathon was organized by Nikul Ukani and Yiyin Zhou, members of the Bionet Group, jointly with Paul Richmond and Adam Tomkins, University of Sheffield:

- Fruit Fly Brain Hackathon (FFBH 2016), March 17, 2016, Columbia University, New York, NY.

Overview

The goal of the hackathon is to bring together researchers interested in developing executable models of the fruit fly brain. Towards that end we will engage systems and computational neuroscientists in modeling, design, implementation and biological validation of an open-source emulation platform of the whole fruit fly brain. All hackathon participants will be provided with an Amazon Machine Image of the recently developed open-source Neurokernel platform [1] for executable fruit fly brain circuits.

The hackathon is aimed at three main groups of participants: biologists, modelers and software engineers. For biologists, the hackathon focuses on the intuitive modeling and representation of biological data, such as anatomical and recordings data of the fruit fly brain, in the NeuroArch database. For modelers, the hackathon aims at creating/modifying models of neuropils that are compliant with the Neurokernel API. For software engineers, the hackathon focuses on improving the Neurokernel platform and its API, and developing new, proof-of-concept features that are needed by biologists and modelers alike. All hackathon participants will be strongly encouraged to collaborate towards the realization of executable fruit fly brain models.

The Fruit Fly Brain Hackathon is organized in conjunction with the Columbia Workshop on Brain Circuit, Memory and Computation on March 18-19, 2016. Participants of the hackathon are welcome to attend the workshop.

[1] Lev E. Givon and Aurel A. Lazar, Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain, PLOS ONE 11(1): e0146581. doi:10.1371/ journal.pone.0146581, January 2016.

There were 30 participants attending the hackthon. The complete list of participants is available on the hackathon webpage (FFBH 2016).

3.4 Columbia Workshop on Brain Circuits, Memory and Computation

Finally, during March 2016 I was the Organizer and Program Chair of the highly acclaimed:

- Columbia Workshop on Brain Circuits, Memory and Computation (BCMC 2016), March 18-19, 2016, New York, NY.

Overview

The goal of the workshop is to bring together researchers interested in developing executable models of neural computation/processing of the brain of model organisms. Of interest are models of computation that consist of elementary units of processing using brain circuits and memory elements. Elementary units of computation/processing include population encoding/decoding circuits with biophysically-grounded neuron models, non-linear dendritic processors for motion detection/direction selectivity, spike processing and pattern recognition neural circuits, movement control and decision-making circuits, etc. Memory units include models of spatio-temporal memory circuits, circuit models for memory access and storage, etc. A major aim of the workshop is to explore the integration of various computational sensory and control models.

Participants (Invited)

J. Douglas Armstrong, School of Informatics, University of Edinburgh.
Kwabena Boahen, Bioengineering Department, Stanford University.
Alexander Borst, Max Planck Institute of Neurobiology, Martinsried.
Thomas R. Clandinin, Department of Neurobiology, Stanford University.
Chung-Chuan Lo, National Tsing Hua University, Hsinchu, Taiwan.
Michael Hawrylycz, Allen Institute for Brain Science, Seattle, WA.
Stanley Heinze, Department of Biology, Lund University.
Matthieu Louis, Centre for Genomic Regulation, Barcelona.
Gaby Maimon, Laboratory of Integrative Brain Function, Rockefeller University.
Mala Murthy, Department of Molecular Biology, Princeton University.
Michael B. Reiser, Janelia Research Campus, Ashburn, VA.
Vanessa Ruta, Laboratory of Neurophysiology and Behavior, Rockefeller University.
Friedrich T. Sommer, Redwood Center for Theoretical Neuroscience, UC Berkeley.
Glenn C. Turner, Janelia Research Campus, Ashburn, VA.
Marta Zlatic, Janelia Research Campus, Ashburn, VA.

AFOSR Deliverables Submission Survey

Response ID:6606 Data

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

aurel@ee.columbia.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

212-854-1747

Organization / Institution name

Columbia University

Grant/Contract Title
The full title of the funded effort.

Fundamental Mechanisms of NeuroInformation Processing: Inverse Problems and Spike Processing

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-12-1-0232

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

Aurel A. Lazar

Program Manager
The AFOSR Program Manager currently assigned to the award

Dr. James H Lawton

Reporting Period Start Date

03/01/2012

Reporting Period End Date

04/30/2016

Abstract

During the research period we: (i) devised pathbreaking algorithms for the functional identification and evaluation of non-linear dendritic processing, and (ii) released a groundbreaking open source platform for the isolated and integrated emulation of the fruit fly brain on multiple GPUs (Neurokernel). We established that identifying a single dendritic stimulus processor (DSP) is mathematically dual to decoding of stimuli encoded by a population of neurons with a bank of DSPs. Building on the key duality property, we (i) demonstrated that the evaluation of the functional identification methodology can be effectively and intuitively performed in the stimulus space, and (ii) characterized the effect of noise parameters on the precision of the functional identification of feedforward, feedback and cross-feedback neural circuits with DSP/biological spike generator neuron models. We demonstrated the power of Neurokernel's model integration by combining independently developed models of the retina and lamina neuropils in the fly's visual system and by demonstrating their combined neuroinformation processing capability.

Distribution Statement
This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

Lazar Final Report.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

lazar final2016.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

Lazar, A.A., Ukani, N.H. and Zhou, Y.,
"A Motion Detection Algorithm Using Local Phase Information",
Journal of Computational Intelligence and Neuroscience, Volume 2016,
January 2016.

L.E. Givon and A.A. Lazar,
"Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain",
PLOS ONE, January 2016.

A.A. Lazar, K. Psychas, N.H. Ukani and Y. Zhou,
"A Parallel Processing Model of the Drosophila Retina",
Neurokernel Request for Comments, Neurokernel RFC #3, August 2015.

L. E. Givon, A. A. Lazar,
"Neurokernel: An Open Source Platform for Emulating the Fruit Fly Brain",
Neurokernel Request for Comments, Neurokernel RFC #4, October 2015.

L. E. Givon, A. A. Lazar, and N. H. Ukani,
"NeuroArch: A Graph dB for Querying and Executing Fruit Fly Brain Circuits",
Neurokernel Request for Comments, Neurokernel RFC #5 , December 2015.

L. E. Givon, A. A. Lazar,
"Generating an Executable Model of the Drosophila Central Complex",
Neurokernel Request for Comments, Neurokernel RFC #6, April 2016.

2. New discoveries, inventions, or patent disclosures:
Do you have any discoveries, inventions, or patent disclosures to report for this period?

Yes

Please describe and include any notable dates

Invention report submitted to Columbia Technology Ventures
(patent pending)

Robust and efficient visual motion detection based on Fast Fourier
Transform and phase-based algorithms
by
Aurel A. Lazar, Nikul H. Ukani and Yiyin Zhou

Do you plan to pursue a claim for personal or organizational intellectual property?

Yes

Changes in research objectives (if any):

Change in AFOSR Program Manager, if any:

Extensions granted or milestones slipped, if any:

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

Jul 29, 2016 15:45:13 Success: Email Sent to: aurel@ee.columbia.edu