CEPID Neuromat

Complementary Form

2013 - 2015

Contents

1 RIDC NeuroMat Identification

The Research, Innovation and Dissemination Center for Neuromathematics (RIDC NeuroMat) is a center of mathematics whose mission is to develop the new mathematics needed to construct a Theory of the Brain accounting for the experimental data gathered by neuroscience research. Mathematician Antonio Galves coordinates this center. Hosted by the University of São Paulo, the RIDC NeuroMat was established in 2013, with support from the São Paulo Research Foundation (FAPESP), grant 2013/07699-0.

The RIDC NeuroMat has an interdisciplinary team, bringing together researchers in mathematics, computer science, statistics, neuroscience, biology, physics and communication, among other disciplines. RIDC NeuroMat leads a worldwide university network, with ramifications that sprawl to several high-level research institutions in Brazil, Latin America, the United States and Europe. Most research output has had co-authors from more than one country, thus contributing to put NeuroMat at the center of a blossoming international scientific cooperation around Neuromathematics.

Alongside a research team that focuses on the scientific challenges pertaining to Neuromathematics, NeuroMat has active technology-transfer and dissemination teams. The tecnology-transfer effort is concentrated on developing free, open-source computational tools to manage and compile experimental and clinical data. This development team is part of a joint effort to create an international open database for neuroscientific data. The dissemination-team effort includes a nonstatic web portal (Creative Commons license), open multimedia productions and training projects with public-school teachers. A distinctive feature of this effort is that it relies on web-2.0 media tools as a means of communicating on-the-go scientific endeavors as well as involving a scientific and broad community around bridging the high-level science that this RIDC develops and general audiences.

A.C. R. da Silva Filho (USP), C. D. Vargas (UFRJ), E. Hamburger (USP), J. Stolfi (UNICAMP), L. R. Batistella (USP), P. A. Ferrari (USP), and Y. Kohayakawa (USP) remain co-principal investigators, along with PI A. Galves (USP). B. Bollobás (Cambridge/Memphis), C. Newman (NYU), D. Brillinger (UC Berkeley), L. Cohen (National Institute of Neurological Disorders and Stroke), and W. Szpankowski (Purdue/NSF CSoI) take part in NeuroMat's International Advisory Board.

NeuroMat's main laboratory and offices are located on a three-story building, with approximately 800 square meters, at 1171 Prof. Luciano Gualberto Avenue, at USP's central campus, in São Paulo. NeuroMat's main building will soon go through an extension (+175 square meters) and renovation to support new laboratory facilities; the construction cost is estimated at BRL R$ 1,603,339 and is fully covered by USP, MaCLinC grant (recipient: Antonio Galves). NeuroMat is currently setting up a High-Performance Computational Center at the USP's Ribeirão Preto campus. NeuroMat's administrative staff team is composed of two administrative assistants and an IT professional; an executive manager will join the team at the end of October. These positions are supported by USP.

2 Summary of the scientific activity of NeuroMat

The main goal of FAPESP's Research, Innovation and Dissemination Center for Neuromathematics (RIDC NeuroMat) is to build the new mathematical, statistical and computational framework which is necessary to address the challenges of neurobiology.

Activities presented in this report strictly relate to the goals announced in the document submitted to FAPESP in November 2012, in the third and final step of the selection process ("Goals for the first two years"). In this document, the scientific goals for the first two years are described in the subsection *A new mathematical framework for neuroscience*. They are the following:

- Development of a new class of stochastic processes;
- Development of the statistical tools required by this new class of stochastic processes.

Concerning the first goal, in the first two years of activities NeuroMat has advanced in the development of a new class of stochastic processes to model systems of interacting neurons. The starting point of this development is the model introduced in the article written by Galves and Löcherbach "Infinite Systems of interacting chains with memory of variable length - a stochastic model for biological neural nets, Journal of Statistical Physics, **151**, pp 896-921, 2013". It is worth mentioning that the new direction of research initiated by this initiative to model systems of interacting neurons already gave rise to six PhD dissertations.

Main achievements in the development of this new class of stochastic processes were threefold.

Firstly, the NeuroMat research team has addressed the issue of hydrodynamic limits for this class of stochastic processes. This is a crucial and unavoidable step to understand the time evolution of mesoscopic and macroscopic behavior of the system. As we argued in (25) *"EEG as well as fMRI data describe the collective behavior of huge subpopulations of neurons. This makes it reasonable to consider a space-time rescaling of the microscopic system reminiscent of what is usually done for interacting particle systems under the name of hydrodynamical limits."* The hydrodynamic limit has been studied in the original model as well as in a spatially structured version (see (2), (22) and (27)). In the above paragraph and in what follows, our articles are quoted by the reference numbers that were used in the extended report of activities, Section 2.1, available at neuromat.numec.prp.usp.br/relatorio/2015/relatorio.pdf.

Secondly, the long term behavior of the stochastic processes in this class of models was studied in (5), (22) and (25). This is an important step to understand the qualitative features displayed by the stochastic process during its time evolution. A central question is to understand whether the initial history of the system affects its long term behavior. From the point of view of neurobiology this theoretical study makes it possible to predict how the system will behave asymptotically in time.

Thirdly, we presented new perfect simulation methods for this class of stochastic processes. Perfect

simulation is related to the issue of long term behavior. Already studied in the original paper by Galves and Löcherbach (2013), it has been extended to the continuous time frame in (26) and in (28). All the aforementioned issues have been discussed and put in perspective in (25).

Development of the statistical tools required by this new class of stochastic processes was the second goal we announced, and this issue has been pursued in papers (2), (20), (31), (32), (35), (43), (44), (47), (74), (75) and (77).

In (75), we address a basic issue in the scientific project of NeuroMat, conjecturing that the brain assigns models to external stimuli by a procedure which is reminiscent of statistical model selection. The experimental protocol consists in exposing volunteers to rhythmic sequences of beats generated by a random source while the EEG signal is recorded. The research question is whether or not the brain is able to identify the context tree which characterizes the random source producing the stimuli. In order to tackle this question, we had to introduce the notion of stochastic processes driven by a context tree model. This new mathematical framework has enabled us to find empirical evidence supporting the conjecture that the brain does statistical model selection. This is a main breakthrough we achieved this year.

The conjecture that the brain performs statistical model selection is not new — it can be traced back to the pioneer work of von Helmholtz who introduced the notion of "unbewusster Schluss" ("unconscious inference"). However, our approach to the conjecture, introduced in (75), is entirely new. First of all, we introduce a new experimental protocol using context tree random sources. Secondly, we translate the problem of structure identification by the brain as a context tree selection procedure. Lastly, EEG data being functions that map time into a multidimensional real space, we face the challenge of testing hypotheses in infinite dimensional spaces.

An interesting extension of this work is presently under development. The experimental protocol is based on a computer game, the *Goalkeeper Game*, developed by the NeuroMat team. In this new experimental protocol, instead of EEG data as in the article (75), the system records the sequence of guesses of the player, who must identify the context tree model used by a virtual football player performing a sequence of penalty kicks. As in (75) the scientific question is whether the brain of the player is able to identify the context tree model used to kick the penalties. A beta version of the game in available at game.numec.prp.usp.br (a version for smartphones is planned).

The model introduced in Galves and Löcherbach (2013) has as one of its components a graph of interactions between regions in the brain in several scales. The features characterizing the graph of interactions at different scales are one of the most important open questions in science. Addressing this question is a priority of the NeuroMat project. For this reason we are organizing the workshop "Random Graphs in the Brain," from November 23-7, coordinated by NeuroMat team members C. Pouzat, A. Galves and C. Vargas, along with R. van der Hofstad.

An issue which is common to all statistical problems addressed by NeuroMat is the fact that they deal with very high dimensional data. This is somehow present in most papers that the NeuroMat research team has published, noteworthily in (2), (20), (21), (31), (32), (45), (46) and (49).

The establishment of NeuroMat's High-Performance Computational Center will provide the team with an important new experimental tool. In fact, the simulation of large-scale network models remains a key activity to test analytical results, and NeuroMat's High-Performance Computational Center will be a laboratory allowing to make such tests.

The scientific achievements we have introduced here remain the expression of a cohesive, collaborative global network of research and exchange. This network brings together scientists from universities in Europe —France (3), Italy (3), Netherlands (3), United Kingdom (1)—, Latin America —Argentina (3), Brazil (8), Uruguay (1)— and the United States (8). A vast proportion of research productions, as presented on Section 2 of this report, has had co-authors from more than one country, thus contributing to put FAPESP's NeuroMat as a relevant player in the global field of Neuromathematics. An illustration of the international network that NeuroMat is embedded in and is fostering is provided on Figure 1.

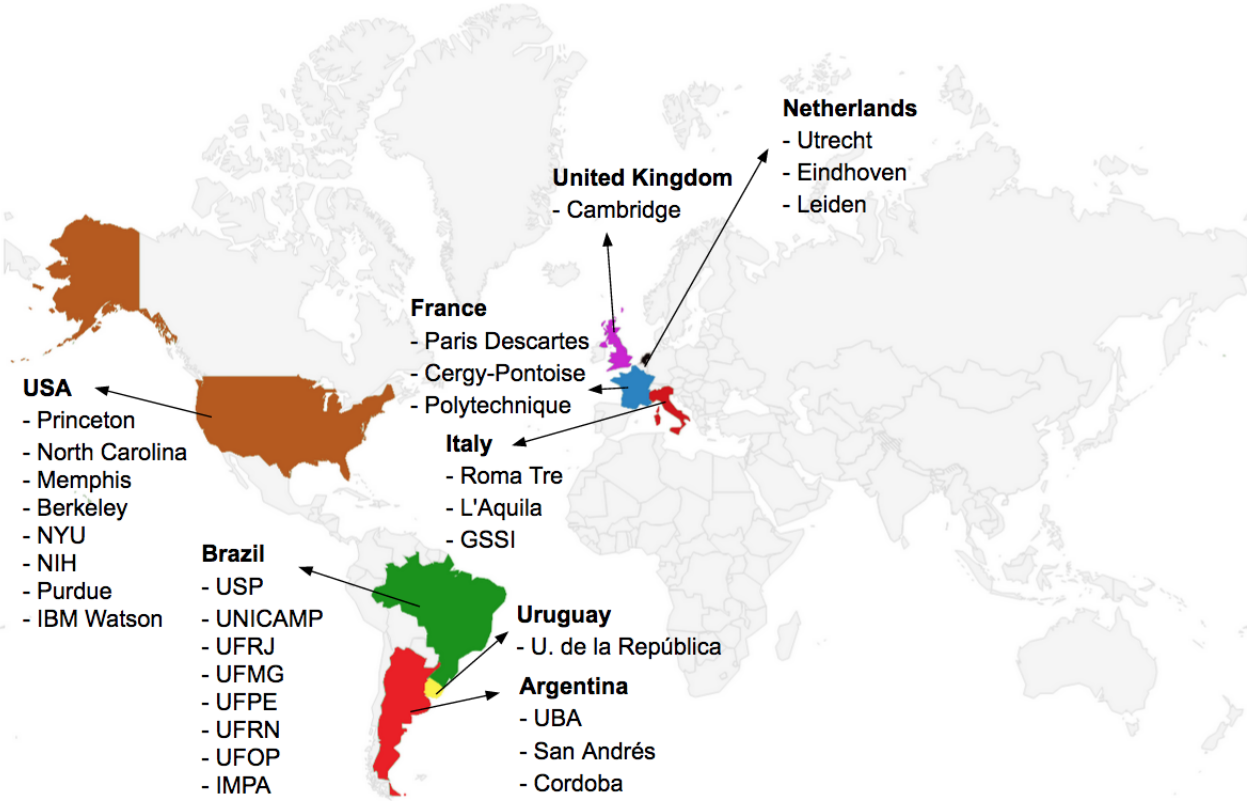

Figure 1: Illustration of FAPESP's NeuroMat global network of scientific, institutional affiliations.

To conclude is worth mentioning that more than 60 papers were already published or accepted for publication since the inception of the RIDC NeuroMat. Besides the articles already published or accepted for publication, we presently have around 20 submitted articles (the list can be downloaded from

neuromat.numec.prp.usp.br/submittedpapers.

International conferences and training programs have been at the core of NeuroMat's engagement in the collaborative global network around Neuromathematics. These activities include:

- The First Workshop of FAPESP's Center for Neuromathematics, from January 20 to 25, 2014, in São Paulo, Brazil.
- The workshop "Mathematics and Neuroscience: a Dialogue", from September 3 to 5, 2014, in Uthrect (the Netherlands).
- The three-day-long training "Spike sorting: What is it? Why do we need it? Where does it come from? How is it done? How to interpret it?," at the University of São Paulo, from November 25 to 27, 2014, with NeuroMat member Christophe Pouzat (CNRS/France).
- The First NeuroMat Young Researchers Workshop in São Paulo, from May 5 to May 7, 2015.
- The Workshop on Stochastic Modeling of Neural Nets and Spike Sorting, in São Paulo, from May 26 to May 28.
- The previously mentioned workshop "Random Graphs in the Brain," from November 23 to November 27 in São Paulo.
- The VI Latin American School on Computational Neuroscience (LASCON), from January 3 to 29, in São Paulo (forthcoming).

3 List of innovation and transfer activities

NeuroMat's technology-transfer team has been developing open-source software tools in order to enable scientists to store their data in databases, in a standardized and secure manner, and to provide retrieval facilities to the users who want access these data. This high-quality scientific data that is being collected and generated in the scope of the project, after the anonymization of the sensitive information, will feed an open database that will be made publicly available via a Web portal. While public scientific data repositories are common in other knowledge areas, they are still rare in neuroscience; thus, the NeuroMat technology-transfer effort aims at contributing to change this scenario. The emphasis on devising computational tools that are akin to spreading open data, a pillar of Open Science, remains a distinctive feature of NeuroMat's software-development activities.

3.1 Conceptual database model for neuroscience experimental data

As the initial step in the building of the NeuroMat computational data and software resources, we have identified and characterized the data gathered in experiments conducted by the researchers of the project. Gathering a large team of researchers and professionals, we have created a conceptual database model for neuroscience experimental data. The NeuroMat database has been designed to store both experimental data and their provenance information (i.e., experimental protocol data and other orthogonal information). There are different types of experiments in neuroscience, e.g., behavioral, cognitive, electrophysiological and neuroimaging. The experimental protocol comprises all the definitions of an experiment, including the statement of the objectives, the description of the groups of subjects to be tested, the experimental conditions to which groups will be submitted, the types of data acquisition that will be carried out, and the equipment settings used in the data acquisition. The data acquisition can be, for example, signals or images captured from the experiment subjects and digitally recorded by equipments such as electroencephalography (EEG) and magnetic resonance imaging (MRI), or measurements or manual notes on the observed behavior of subjects. In order to accommodate all these diverse data and facilitate the evolution of the database, we divided the database structure in modules: Organizational Structure, Experiment Protocol, Electrophysiological Data Acquisition, Behavioral Data Acquisition, Documents.

3.2 Developing the Neuroscience Experiments System – NES

The NeuroMat development team has next devoted its efforts to the design and development of free software for the management of clinical and neuroscience experimental data. The software under development was named Neuroscience Experiments System, or NES. The system aims to assist neuroscience research laboratories in routine procedures for data collection. The model of the NeuroMat database supports the reproducibility of experiments, enables comparison of data across studies, and keeps data provenance. In addition, it promotes standardized formats for experiments and analyses reporting. The

use of a Database Management System (DBMS) in the implementation of the database improves the efficiency and security of the data storage. All data managed by NES are stored in a local database of the lab, with restricted access. NES helps to record quality data, which can later be analyzed, shared or reused easily.

3.3 Quality and security to data collected through questionnaires using LimeSurvey

A considerable part of the data analyzed in neuroscience experiments is collected by means of paper-based questionnaires. These ad-hoc questionnaires are generally designed by the scientists specifically for the studies that they are conducting. To deal with this problem and also to provide more quality and security to data collected through questionnaires, we use an electronic questionnaires management system. In the project, we have adopted LimeSurvey, which is an open-source, web-server-based software that enables users to develop and publish on-line surveys and collect responses. In order to facilitate the administration of questionnaires and to centralize the access to the clinical and experimental data of a laboratory, NES is integrated with LimeSurvey.

3.4 Design based on the needs of the Laboratory of Neuroscience and Rehabilitation

The current features available at NES were designed and evaluated based on requirements collected at Laboratory of Neuroscience and Rehabilitation (LNR) of the Institute of Neurology Deolindo Couto (INDC), at Federal University of Rio de Janeiro (UFRJ). But these features were implemented in such a way that they could benefit any neuroscience laboratory. Since version 0.1 was installed in INDC, in November 2014, 72 patients have been registered and 187 sets of answers have been given to 8 questionnaires. Version 0.2.1, which supports the representation of research projects and experimental protocols, was installed in INDC in August 2015.

3.5 Development of software to support the preparation of annual reports – NIRA

Furthermore, in response to a need identified during the preparation of NeuroMat's first Report of Activities, the NeuroMat technology transfer team along with the administrative body have created a parallel work front aiming to develop a piece of software to support the preparation of annual reports. This software, called NeuroMat Individual Report of Activities (NIRA), is designed so that NeuroMat members can register and share information on publications, scientific missions and service requests.

3.6 Used technologies

Finally, we emphasize that NES and NIRA are web applications, available on the NeuroMat GitHub account, where video tutorials and documentation are also provided, implemented using the Django Web framework, and written in Python. We use the Bootstrap framework to generate the application layout and make it responsive, adjusting the web pages dynamically according the device used. The NES and NIRA databases are being implemented as a relational database in the open-source DBMS PostgreSQL.

4 List of dissemination activities

4.1 Communication Activities

4.1.1 Media Exposure and Newsletter

Activities from FAPESP'S RIDC NeuroMat were featured in twenty seven external media outlets since its inception, in 2013. Media exposure includes pieces on top Brazilian general publications, such as Info Exame, Mente e Cérebro and Le Monde Diplomatique, as well as regular coverage on media vehicles by FAPESP and the University of São Paulo. NeuroMat's clipping is available here. The NeuroMat newsletter has had 17 issues since it was first released in February 2014. It runs monthly, generally being sent to subscribers at the end of each month. It is distributed to around 550 people, always in English. NeuroMat's newsletter compilation is available here.

4.1.2 Web portal, Facebook page and online material

The NeuroMat web portal was launched in early February 2014, and brings up all relevant updates in the research, technology transfer and dissemination activities. It is available here. Publications are in English and Portuguese. Around 7,500 different users –47% of whom were identified as non-Brazilian users– have visited NeuroMat's webpage since its inception, with 46,000 page views. The NeuroMat Facebook page brings up short texts and media content on a daily basis. Since its creation (9/20/2014), the page has reached approximately 450 likes. Amidst content that is originally produced and distributed by NeuroMat's dissemination team, it is worth highlighting the release of seven online lectures and 47 presentations that were part of the First NeuroMat Young Researchers Workshop, in May 2015, and the First Workshop of FAPESP's Center for Neuromathematics, in January 2014. NeuroMat has also offered live streaming and video recording of special events, which so far have attracted around 4,000 viewers. NeuroMat videos are available here. Lectures, presentations and videos are available in English.

4.1.3 NeuroCineMat

The NeuroMat dissemination team has released one movie and is currently producing two movies. "Spike Sorting: Ontology Droplet" was released in June 2015, and is available here. This NeuroMat movie was acknowledged as "Media of the Day" (June 10) on Wikimedia Commons and displayed on the front page of over 200 related websites. A short movie on Open Science and a filmed document of an experiment on brain and rhythms should be released in the fall, 2015. These forthcoming movies have under-production versions, respectively here and here; the password is paraaprovação.

4.2 Education Activities

4.2.1 Training courses for teachers and science activities in schools

NeuroMat has developed content and activities to support teachers on dealing with data description and analysis, i.e., a free course to public school teachers at the Municipal School Desembargador Amorim

Lima, in São Paulo, in May 2015. This pilot activity was offered to forty teachers, and acknowledged as a successful educational activity by the University of São Paulo (here). NeuroMat has actively taken part in involving undergraduate students who have been awarded scholarships (PIBID) in education activities that are related to Neuromathematics, aiming at disseminating topics pertaining to this new topic of research through students' engagement. The ongoing network organized around NeuroMat-related science activities has involved four high school teachers, approximately 350 students and 28 PIBID undergraduate students. NeuroMat team members have led one-time discussions in 7 schools and university settings as a means of disseminating Neuromathematics and related topics, including the participation at the last editions of the Semana do Cérebro and the University of São Paulo's Virada Científica. Trainings were organized on topics such as questionnaires, software and data management.

4.2.2 NeuroMat Wikipedia Initiative

The NeuroMat Wikipedia Initiative is a strategic activity and was launched in July 2014, aiming at ensuring that articles on neuromathematics (research, terminology, theory) are up-to-date, complete, and written in a style that is neutral and appropriate for the general public, as well as based on reliable sources. Activities pertaining to this Initiative include: the writing-up of articles on Wikipedia and other Wikimedia projects; the organization of an edit-a-thon on Neuromathematics on May 5 and 6, 2015; a systematic uploading of pictures and movies on Wikimedia Commons; and two scientific articles —one accepted, one submitted— on the use of Wikipedia and other Wikimedia platforms as tools for education and scientific dissemination. This Initiative has been the object of media coverage.

4.3 Exhibition Activities

4.3.1 "Inside the Brain"

"Inside the Brain" is an exhibition at the São Paulo Museum of Anatomy, which is developed as a partnership between NeuroMat, the Laboratory of Cognitive Science (Labcog) and the Museum of Veterinary Anatomy (MAV), with the support of the Pró-Reitoria de Cultura e Extensão of the University of São Paulo, that shows the logic of life as it relates to the nervous system, in order to understand the functioning of the central nervous system. This event is open and aimed at the general public.

5 List of additional funding sources except from FAPESP

Funding Agency / University / Other sources	**Equipment costs (verba de capital)**	**Other direct costs (verba de custeio)**	**Salaries**	**Scholarships**
NUMEC-MaCLinC - USP	R$1.488.925,25	R$503.096,75		R$18.633,84
FINEP PROINFRA	R$330.000,00	R$390.000,00		
FAPERJ HOSPITAIS UNIVERSITARIOS	R$110.000,00	R$90.000,00		
FAPERJ SEDIADAS 210.737/2014	R$33.200,00	R$89.000,00		
Edital Universal MCT/CNPq/14/2012 480108/2012-9	R$60.000,00	R$24.000,00		R$18.240,00
CAPES NUFFIC – Vigência: 2012-2013		R$52.724,15		
IBRO (Intern.Brain Research Organization) – V LASCON 2014		EURO 14.000,00 (R$53.726,40 bcb 18/08/15)		
Edital Universal MCT/CNPq 14/2014-459335/2014-6	R$17.000,00	R$20.000,00		
PROBRAL CAPES/DAAD - Proc. 430/15		R$33.748,00		
CAPES (PAEP 8621/2013-32) V LASCON 2014		R$ 30.000,00		
Edital Universal MCT/CNPq 14/2012 478537/2012-3	R$9.000,00	R$12.000,00		R$12 960,00

CNPq (Eventos Nacionais 456821/2013-9) - V LASCON 2014		R$ 9.000,00		
CNPq-BPP				R$963.600,00
CNPq				R$390.000,00
CAPES				R$287.000,00
USP			R$4.175.305,55	
UNICAMP			R$1.594.688,72	
UFRN			R$586.812,00	
IMPA			R$267.160,56	
UFRJ			R$ 200.000,00	
UFABC			R$198.302,52	
UFSCar			R$93.871,20	
UFMG			R$91,524.54	
UFPE			R$11.533,24	
UFOP			R$7.627,02	
TOTAL	**R$2.048.125,25**	**R$1.307.295,30**	**R$7.226.825,35**	**R$1.690.433,84**

6 Publications / citations URL

MyResearcherID: http://www.researcherid.com/rid/J-2428-2015

Google Scholar: https://scholar.google.com.br/citations?user=OaY57UIAAAAJ&hl=pt-BR