Computational Neuroscience: Why and How?

Péter Érdi
perdi@kzoo.edu

Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, MI, USA
http://www.kzoo.edu/physics/ccss

Computational Neuroscience Group
KFKI Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences
http://www.rmki.kfki.hu/biofiz/cneuro

with the help of Gábor Csárdi and Tamás Kiss
Single cells: deterministic and stochastic models
The McCulloch–Pitts neuron

A neuron performs a computational operation on its $x_i$ inputs:

$$y = \begin{cases} 
1 & \text{if } \sum_{i=1}^{N} w_i x_i > \theta, \quad w_i > 0 \\
0 & \text{otherwise}
\end{cases}$$

Long cycles in networks of formal neurons:
Elemér Lábos

---

The linear neuron

Output is the sum of weighted inputs:

$$y = \sum_{i=1}^{N} w_i x_i$$
Single cell models II. – The Hodgkin–Huxley model

\[ C \frac{dV(t)}{dt} = I_{Na}(t) + I_{K}(t) + I_{l}(t) \]

\[ I_{Na}(t) = g_{Na}(t) (V(t) - E_{Na}) \]

\[ I_{K}(t) = g_{K}(t) (V(t) - E_{K}) \]

\[ I_{l}(t) = g_{l}(V(t) - E_{l}) \]

\[ g_{Na}(t) = \overline{g}_{Na} m^3(t) h(t) \]

\[ g_{K}(t) = \overline{g}_{K} n^4(t) \]

\[ \frac{ds(t)}{dt} = \alpha_{s} (V(t)) (1 - s(t)) - \beta_{s} (V(t)) s(t), \quad s \in \{m, n, h\} \]
Single cell models III. – The FitzHugh–Nagumo equations

- General description of ionic flow through the cell membrane:

\[ C_m \frac{dV(t)}{dt} = -\frac{V(t) - V_{eq}}{R} + I_{appl}(t) \]

- Excitable membranes: Hodgkin–Huxley equation: 4D ODE for excitability and action potential (signal) generation

- FitzHugh (1955) and Nagumo (1961) dimension reduction. He observed that
  - separation of time scales \((m\text{ gate is fast relative to } h \text{ and } n)\)
  - the sum \(n + h\) is approximately constant in time

- 4D \(\rightarrow\) 2D reduction technique
  - \(m\) is substituted by its steady state value
  - \(h\) and \(n\) is contracted into a \textit{slow variable} (sodium gating): \(w\)
  - \textit{fast variable} (membrane potential): \(v\)
The FitzHugh–Nagumo model I.

One version of the FitzHugh-Nagumo equation (based on van der Pol (1926) and Bonhoffer (1948))

\[
\begin{align*}
\frac{dv(t)}{dt} &= v(t)(v(t) - \alpha)(1 - v(t)) - w(t) + I \\
\frac{dw(t)}{dt} &= \epsilon(v(t) - \gamma w(t))
\end{align*}
\]

This is a *singularly perturbed system*: one variable evolves much faster than the other variable.

*I*: external excitation (control parameter)
Properties of the FitzHugh–Nagumo model I.

Excitability

For $I = 0.1$ a stable fixed point exists

- small perturbation $\rightarrow$ rapid relaxation
- larger-than-threshold perturbation $\rightarrow$ “large excursion”, action potential generation
Properties of the FitzHugh–Nagumo model II.

Oscillation

Stable fixed point: approaching resting state \((I = 0.15)\)

Unstable fixed point: oscillation \((I = 0.25)\)
Properties of the FitzHugh–Nagumo model III.

The FitzHugh–Nagumo equation does not provide a very accurate description of action potential shapes (ignores Ohms law and explicit reference to the Na and K equilibrium potentials), it does provide mathematical insight into the nature of neuronal excitability. Spikes are generated when $I$ becomes strong enough to destabilize the equilibrium state.
Properties of the FitzHugh–Nagumo model IV.

A “type II” membrane

The fixed point looses stability at $I = I_{bif}$ giving birth of an (unstable) limit cycle. This scenario is called the (subcritical) Hopf bifurcation.
Realistic models – Towards a computational neuropharmacology

- Detailed morphological description using multicompartmental modeling techniques
- Extending the Hodgkin–Huxley formalism by new channel types
- Ion channel densities vary among compartments
- Accounting for intracellular calcium concentration dynamics
- Introducing synaptic channels and gap junctions

These features enable the modeller to specifically describe cellular and synaptic phenomena, creating a helpful mathematical tool for drug discovery.
Realistic models – An example

A CA1 pyramidal cell model

- Model modified from Varona et al. (2000)
- 256 compartment, including soma, axon with Ranvier nodes and axon initial segment
- 8 types of ion channels ($Na$, $K_{DR}$, $K_{M}$, $K_{A}$, $K_{AHP}$, $K_{CT}$, $Ca$, $I_h$)
- Densities of ion channels in different compartment correspond to measured values
- Describing Ca dynamics
- GABA and glutamatergic synapses at specified compartments

Anatomy of a model neuron from Hajós et al. (2004)
Stochastic single cell models I.

A classical paper:
Gerstein GL, Mandelbort B: Random walk models for the spike activity of a single neuron, Biophys J. 1964, 71:41–68

Membrane potential as a Brownian particle: continuous time continuous state space stochastic process (diffusion process).

\[
dX(t) = \mu[X(t), t]dt + \sigma[X(t), t]dW(t) \tag{1}
\]

\[
T = \inf_{t \geq t_0} \{ t : X(t) > S(t) | X(t_0) = x_0 \} \tag{2}
\]

\[
g[S(t), t | x_0, t_0] = \frac{\partial}{\partial t} P\{T \leq t\} \tag{3}
\]
Stochastic single cell models II.


\[ A_1(x) = \mu, \quad \mu \in \mathbb{R} \quad (4) \]
\[ A_2(x) = \sigma^2, \quad \sigma \in \mathbb{R}^+ \quad (5) \]

Drift: from Wiener to Uhlenbeck-Ornstein process.

\[ A_1(x) = -\frac{x}{\vartheta} + \mu \quad (6) \]
\[ A_2(c) = \sigma^2 \quad (7) \]

Nonstationary version of the OU model:

\[
\begin{align*}
\left\{ 
A_1(x) &= -\frac{x}{\vartheta} + \alpha(t)e + \beta(t)i \\
A_2(x) &= \alpha_i(t)e^n + \beta(t)i^n \quad (n = 2, 3, \ldots)
\end{align*}
\]
Stochastic single cell models III.

First passage time problem (L. Ricciardi and C.E. Smith)

When the voltage at a particular place on a neuron reaches a threshold, an action potential (nerve impulse) is produced. Many point processes in biology have similar origins as “first passage times”; that is, they occur when some underlying process first reaches a critical level or threshold.

Even for simple models of the underlying process (1-dimensional stochastic differential equations), very few analytical results are available for first passage times. Through simulation and heuristic approximation methods, several different types of behavior have been identified. The main current research activities are further development of approximation methods.
Stochastic single cell models IV.

Time to first spike in stochastic Hodgkin-Huxley systems
Henry C. Tuckwell and Frederic Y.M. Wan, 2004

- Hodgkin–Huxley equation with Gaussian white noise (with $\sigma$ amplitude) approximating Poisson trains of excitatory and inhibitory post-synaptic potentials:

$$C_m \frac{dV(t)}{dt} = I_{\text{membrane}}(t, V) + \mu + \sigma dW$$

- Transition probability function $p(v, x, y, z, t; \hat{v}, \hat{x}, \hat{y}, \hat{z}, \hat{t})$ of the 4–dimension process $(V, X, Y, Z)$ ($V$ – membrane pot., $X$ – potassium activation, $Y$ – sodium activation, $Z$ – sodium inactivation) satisfies the backward Kolmogorov equation:

$$-\frac{\partial p}{\partial \hat{t}} = \hat{L}p,$$

where $\hat{v}, \hat{x}, \hat{y}, \hat{z}, \hat{t}$ are backward variables and $\hat{L}$ is a suitable differential operator for the noisy HH equations.
Stochastic single cell models V.

Comparision of the 4D and the 2D version

Numerical simulations of the 4D HH model and a 2D version analogous to the FitzHugh–Nagumo system, shows similar time to first spike distributions

Solving the PDE using the method of characteristics for the 4D and the 2D model

\[
\begin{array}{c|c|c|c}
\mu & \text{Calc. 2D} & \text{Sim. 2D} & \text{Sim. 4D} \\
\hline
5 & 3.757 & 3.76 & 4.47 \\
10 & 1.676 & 1.68 & 1.79 \\
15 & 1.137 & 1.14 & 1.19 \\
20 & 0.866 & 0.87 & 0.89 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\mu & \text{Calc. 2D} & \text{Sim. 2D} & \text{Sim. 4D} \\
\hline
4 & 3.425 & 3.92 & 3.70 \\
5 & 3.382 & 3.42 & 3.30 \\
10 & 2.750 & 1.76 & 1.78 \\
\end{array}
\]
Let \( \{X(t), \ t \geq 0\} \), \( X(0) = x \) be a temporally homogeneous diffusion process with Ito stochastic differential

\[
dX = \alpha(X)dt + \beta(X)dW
\]

where \( \{W(t), \ t \geq 0\} \) is a standard Wiener process

The \( T(x, y) \) random variable is the first time when \( X \) hits the \( Y \) moving barrier:

\[
T(x, y) = \inf\{t | X(t) = Y(t) | X(0) = x < y = Y(0)\}
\]
Stochastic single cell models VII.

Firing time of a model neuron in the moving barrier scheme

The displacement $X(t)$ of a nerve cell's electrical potential from its resting value has been described by a diffusion approximation:

$$dX = (\alpha - X)dt + \beta dW, \quad X(0) = x$$

- constant threshold: not plausible $\rightarrow$ threshold declines after absolute refractory period

- time-dependent thresholds: eg. oscillatory (Wilbur and Rinzel, 1983), exponential (Weiss, 1964) $\rightarrow$ moments of first-passage (firing) times can be calculated

- the model can be extended with jump components:

$$dX = \alpha(X)dt + \beta(X)dW + \gamma(X,u)\nu(dt \times du)$$
Stochastic resonance I.

The phenomena

Stochastic resonance is a mechanism, where noise plays a beneficial role in amplifying weak signals arriving to some nonlinear system.

Consider a double–welled container where the highly damped ball is periodically forced and some noise is also added to the system.

From: http://www.agnld.uni-potsdam.de/~zaikin/html_test/SRintro.html
The model consists of two layers of integrate and fire cells. The first layer is driven by a periodic and noisy signal.

For strong synaptic coupling the signal to noise ratio (SNR) is high in the low noise regime, increasing the noise weaker coupling is desirable. → Synaptic depression sets coupling strength.
Stochastic resonance III.

The effect of synaptic depression on stochastic resonance
László Zalányi, Fülöp Bazsó, Péter Érdi (2001)
Neural code: rate vs. temporal I.

For many years, ever since the experiments of Lord Adrian two-thirds of a century ago, neurobiologists agreed that neurons communicate by means of a frequency (or rate) code, in which nerve impulse frequency is used to specify the variable in question for example, light intensity. But recently a number of neurobiologists have, based on new observations, proposed that the frequency code is insufficient, and that other properties of spike (action-potential) trains are used to encode information.
Neural code: rate vs. temporal II.

Rate as a Spike Count (Average over Time)

rate = average over time 
(single neuron, single run)

\[
\nu = \frac{n_{sp}}{T}
\]

Definition of the mean firing rate via a temporal average.

Neural networks
From neurons to networks I.

Hierarchy of dynamics

- Neuron
  - soma (cell body)
  - dendrites
  - axon

- Synapse
  - $S_{ij}$

- Network
  - afferent neurons
  - interneurons
  - efferent neurons
  - $I(t)$
From neurons to networks II.

Neuron and synapse

Neuron

- complex morphology
- membrane potential dynamics
- signal generating and processing device
- dendrites
- axon
- soma (cell body)

Synapse

- structural and functional connection between cells
  - specialized for transferring information
- neurochemically mediated synapses (mostly)
- excitatory and inhibitory
- synaptic strength (synaptic weight, synaptic efficacy)
- $\frac{d}{dt} s_{ij}(t)$ ↔ learning

Alfred Rényi Math. Inst. 9/12, 2004
From neurons to networks III.

Network

\[ \frac{d}{dt} s_{ij}(t) = 0 \]  
fixed wiring

\[ \frac{d}{dt} s_{ij}(t) \neq 0 \]  
pattern formation problems

pattern recognition problems

afferent neurons  
interneurons  
efferent neurons
How dynamic systems can process continuously varying inputs?

How to preserve (if at all) the power of computation with attractors?
Computation with attractors: Scope and limits II.

Conventional assumptions and rules, implications and difficulties I.

- attractors exist and they are fixed points
- learning stage precedes the recall process
- learning is described by a static “one-shot” rule
Conventional assumptions and rules, implications and difficulties II.

• the recall dynamics is given as

\[ \dot{a}_i(t) = f(a(t), S, I); \quad a(0) = a_0 \]

(a – activity, S – connectivity, I – input)

– the inputs and mostly their temporal character are neglected
– the mathematical objects to be classified are the (static) initial values of the rapidly varying activity variables.

• \[ E := \frac{1}{2} \sum_{i,k=1}^{n} a_i S_{ik} a_k + \sum_{i=1}^{n} a_i \Theta_i \]

• random asynchronous update.
Computation with attractors: Scope and limits IV.

Modifications of the classic scenario,
Fixed points vs. Strange attractors

Not only fixed points (A), but also limit cycles (B) or strange attractors (C) might be involved.

- neurophysiological experiments, theoretical studies.
- structural conditions for the possible generation of rhythms and chaos can be given based on the notions of qualitative stability and instability.
Computation with attractors: Scope and limits V.

Modifications of the classical scenario, 
Input influence

Distortion of the basin can also be obtained by tuning the external input (Tank and Hopfield)

The behavior of autonomous systems (generally) strongly depends on initial values,

\textbf{BUT}

- the violation of the Lipshitz condition eliminates this dependence (Zak)

- input variables were identified with initial values in a degenerate case (Baird)

- \textit{constant inputs} are classified by specific networks (Grossberg and Cohen, Kelly)
Activity equations are supplemented by dynamic equations for the connections (‘learning rule’)

Learning might cause the distortion of attractor basins which might even lead to qualitative changes in the nature of the attractors.
Two level dynamics (constant input):

\[ \dot{a}(t) = f(a(t), S(t), I); \quad a(0) = a_0 \]
\[ \dot{S}(t) = g(a(t), S(t)); \quad S(0) = S_0 \]

\textit{infinite slow} learning (bifurcation)

the classical interpretation of the activity dynamics can be preserved using the notation of “dynamic categories”.

\textit{finite (> 0) velocity}

during recall process: frozen connection strength

initial values can be \textit{RE}assigned → the classical classification problem exists

\textbf{WHAT TO DO?}
Computation with attractors VIII.

*Initial values versus time-variable inputs: what to classify?*

Nonautonomous system: the environment has a direct influence on the change of state.

Two different sources of nonautonomy:

- the parameters depend explicitly on time
  
  (It can be eliminated at the cost of increasing the dimension of the state space by introducing differential learning rule, new *interpretation* of the state space.)

- (most often additive) time-dependent term, representing a continuously time-variable input.

\[
\dot{a}(t) = f(a(t), S(t), I(t, \beta)); \quad a(0) = a_0 \\
\dot{S}(t) = g(a(t), S(t)); \quad S(0) = S_0
\]

(continuous learning, time-dependent input)
Computation with attractors: Scope and limits VIII.

1. Which are the objects to be classified?

2. How to define the classes (categories)?

1. In the general case the set of input functions should be classified.

   More specific case: the form of the input functions is fixed, and the set of parameters \((\beta)\) is the object of classification.

2. Autonomous ODE \(\Rightarrow\) fixed number of ‘innate’ categories \(\Leftrightarrow\) number of attractors

   Nonautonomous ODE \(\Rightarrow\) even the existence of attractors is doubtful.
Computation with attractors: Scope and limits IX.

Under certain conditions the classification of *inputs* in NA systems can be reduced to the classification of *initial values*.

One trivial elimination of nonautonomy:
Proper transformation: input parameters $\rightarrow$ initial values

\[
\begin{align*}
\text{Nonautonomous: } & \dot{x}(t) = F(x(t), I(t, \beta)); \quad x(\tau) = \xi \\
\downarrow \\
\text{Autonomous: } & \dot{z}(t) = F^*(z(t)); \quad z(0) = k(\beta)
\end{align*}
\]
# Computation with attractors: Scope and limits

<table>
<thead>
<tr>
<th></th>
<th>Constant $I(t)$</th>
<th>Time-dependent $I(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant $S_{ij}$</td>
<td>classification of initial values or at most constant input $\Downarrow$</td>
<td>pattern classification $\Leftarrow$ e.g. $\Downarrow$</td>
</tr>
<tr>
<td></td>
<td>classical scenario</td>
<td>oscillatory $\Downarrow$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-oscillatory $\Downarrow$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Li-Hopfield olfactory bulb model</td>
</tr>
<tr>
<td>Time-dependent $S_{ij}$</td>
<td>Dynamic categories $\Uparrow$</td>
<td>nonautonomous $\Downarrow$</td>
</tr>
<tr>
<td></td>
<td>$\Downarrow$ synaptic modification induced transition</td>
<td>autonomous</td>
</tr>
<tr>
<td></td>
<td>$\Downarrow$ no real classification problems</td>
<td>simulation methods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>classification or time dependent input during learning</td>
</tr>
</tbody>
</table>
Neural Networks II: Small world and scale-free networks?

- The nervous system of the nematoda worm *Caenorhabditis elegans* forms a small-world network.

- Mammalian cerebral cortex: its network is neither regular nor random.

- The distance of two arbitrarily choosen cortical neurons is 5 (John Szentágothai)
Neural Networks II: Small world and scale-free networks?

Specificity Versus (Quasi-) Randomness in Cortical Connectivity

J. Szentágothai

First Department of Anatomy, Semmelweis University Medical School, H-1450, Budapest, Hungary

NEURAL CONNECTIVITIES: BETWEEN DETERMINISM AND RANDOMNESS

P. Erdi¹,² and J. Szentágothai²

¹ Central Research Institute for Physics, Hungarian Academy of Sciences, Budapest, Hungary
² 1st Department of Anatomy, Semmelweis University Medical School, Budapest, Hungary

Rényi álfréd—szentágothai János
Az ingérületátvitel valószínűsége egy egyszerű konvergens kapcsolású interneuronális synapsys modellemben (Előzetes közlemény)
Neural Networks II: Small world and scale-free networks?

- **Anatomical connectivity:** the set of synaptic connections linking its elements
- **Functional connectivity:** the correlations between spatially remote neurophysiological events.
- **Effective connectivity:** the influence a neuron (or neuronal population) on another. At the neuronal level this is equivalent to the effect pre-synaptic activity has on post-synaptic responses, otherwise known as synaptic efficacy. Models of effective connectivity are designed to identify a suitable metric of influence among interconnected components (or regions of interest) in the brain.
Neural Networks II: Small world and scale-free networks?

Glossary: graph theory ad networks

Adjacency (connection) matrix: The adjacency matrix of a graph is a $n \times n$ matrix with entries $a_{ij} = 1$ if node $j$ connects to node $i$, and $a_{ij} = 0$ is there is no connection from node $j$ to node $i$.

Characteristic path length: The characteristic path length $L$ (also called 'path length' or 'average shortest path') is given by the global mean of the finite entries of the distance matrix. In some cases, the median or the harmonic mean can provide better estimates.

Clustering coefficient: The clustering coefficient $C_i$ of a node $i$ is calculated as the number of existing connections between the node’s neighbors divided by all their possible connections. The clustering coefficient ranges between 0 and 1 and is typically averaged over all nodes of a graph to yield the graph’s clustering coefficient $C$. 
Neural Networks II: Small world and scale-free networks?

Glossary: graph theory ad networks II.

Connectedness: A connected graph has only one component, that is a set of nodes, for which every pair of nodes is joined by at least one path. A disconnected graph has at least two components.

Cycle: A cycle is a path that links a node to itself.

Degree: The degree of a node is the sum of its incoming (afferent) and outgoing (efferent) connections. The number of afferent and efferent connections is also called the 'in-degree' and 'out-degree', respectively.

Distance: The distance between a source node $j$ and a target node $i$ is equal to the length of the shortest path.
Neural Networks II: Small world and scale-free networks?

Glossary: graph theory ad networks III.

**Distance matrix:** The entries $d_{ij}$ of the distance matrix correspond to the distance between node $j$ and $i$. If no path exists, $d_{ij} = \infty$.

**Graph:** Graphs are a set of $n$ nodes (vertices, points, units) and $k$ edges (connections, arcs). Graphs may be undirected (all connections are symmetrical) or directed. Because of the polarized nature of most neural connections, we focus on directed graphs, also called digraphs.

**Path:** A path is an ordered sequence of distinct connections and nodes, linking a source node $j$ to a target node $i$. No connection or node is visited twice in a given path. The length of a path is equal to the number of distinct connections.
Neural Networks II: Small world and scale-free networks?

**Glossary: graph theory ad networks IV.**

**Random graph:** A graph with uniform connection probabilities and a binomial degree distribution. All nodes have roughly the same degree (‘single-scale’).

**Scale-free graph:** Graph with a power-law degree distribution. ‘Scale-free’ means that degrees are not grouped around one characteristic average degree (scale), but can spread over a very wide range of values, often spanning several orders of magnitude.

Neural Networks II: Small world and scale-free networks?

Scale-free structure of brain functional networks

• nodes: voxels.

• edges: functional connectivity.

• connected: those voxels whose temporal evolution is correlated beyond a threshold.

• fMRI experiment: 64x64x36 voxels.

• edge distribution of a realistic network: scale-free.

• large-scale network of human brain might be scale-free.

http://www.apkarianlab.nwu.edu/text/smallworld.pdf
Concluding remarks
The RIKEN project

Understanding the Brain

Protecting the Brain

Creating the Brain