Some Properties of Memory Retrieval in an Associative Memory Neural Network for Conscious and Unconscious Mental Behavior

Roseli S. Wedemann[1,4]

Raul Donangelo[2,3], Luís Alfredo V. de Carvalho[3]

Henrik J. Jensen[4], Maheen Siddiqui[4]

[1]Universidade do Estado do Rio de Janeiro
at Imperial College until September 2013

[2]Universidad de la República - Uruguay

[3]Universidade Federal do Rio de Janeiro

[4] Imperial College London

Happy Birthday Constantino !

Motivation

- Develop illustrative, schematic, self-organizing, neural-network models to describe mechanisms associated with mental processes.

Neurocomputational Models

- Understand the importance of the capacity for operating on symbols in mental life and in therapy.

- Understand cognitive functions involved in consciousness → artificial consciousness

- Study the topological properties of these models. Concepts and methods from statistical mechanics and complex networks.

Mental Processes

- **Creativity:** Capacity of broadening attention to a wider range of elements, allowing the discovery of unusual associations of ideas (associationists).
- **Delusions:** Statements made in inappropriate contexts. Total and unquestionable certainty implies incorrigibility. A detachment from *reality* (Freud).
- **Disorganized Thought:** Excessively heterogeneous ideas are associated. Subject´s discourse becomes incoherent and unintelligible.

- In schizophrenia, disorganized thought, delusions and hallucinations are considered positive psychotic symptoms and respond well to neuroleptic treatment.

- Psychodynamical theories correlate creativity, psychopathology and unconsciousness. Aspects such as broader, distant or looser association making and unfocusing of attention are common in describing creativity, psychotic thinking and schizophrenia.

- **Neuroses:** as described by Freud, repressions and traumas causing a compulsion to repeat painful (neurotic) symptoms, working-through, conscious / unconscious associations

"The intention is to furnish a psychology that shall be a natural science: that is, to represent psychical processes as quantitatively determinate states of specifiable material particles, thus making those processes perspicuous and free from contradiction."

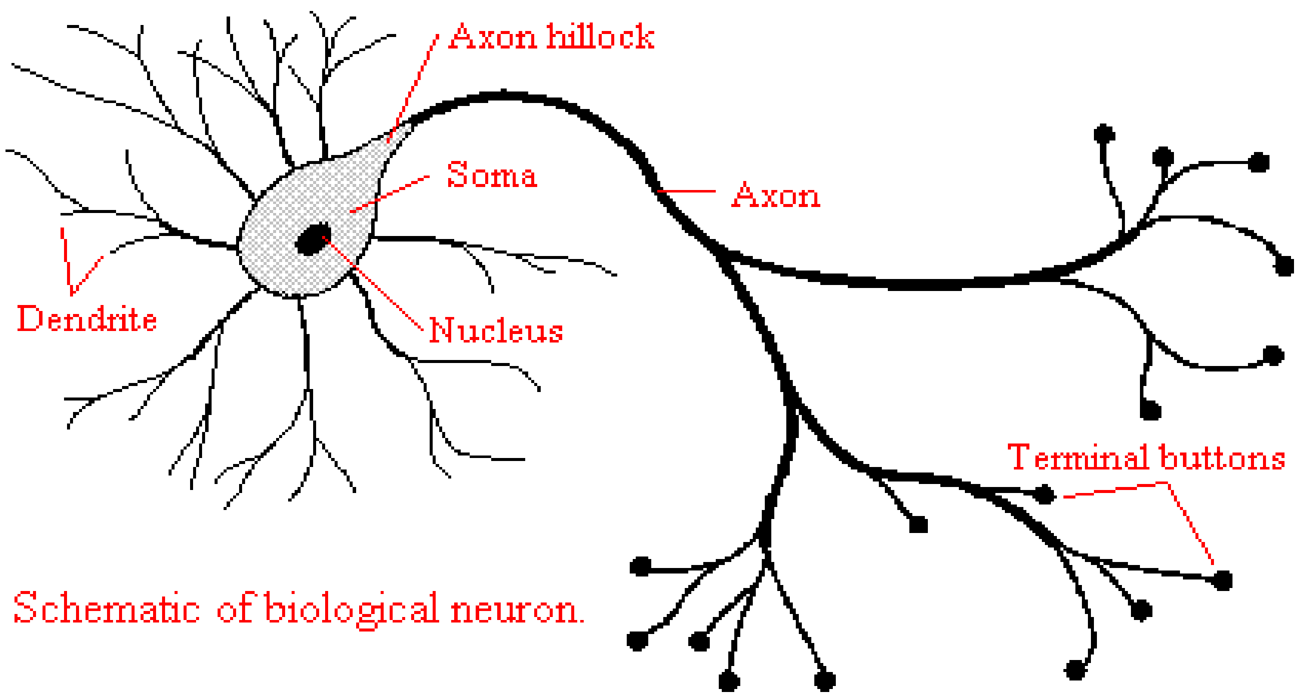
Freud's *Project for a Scientific Psychology*

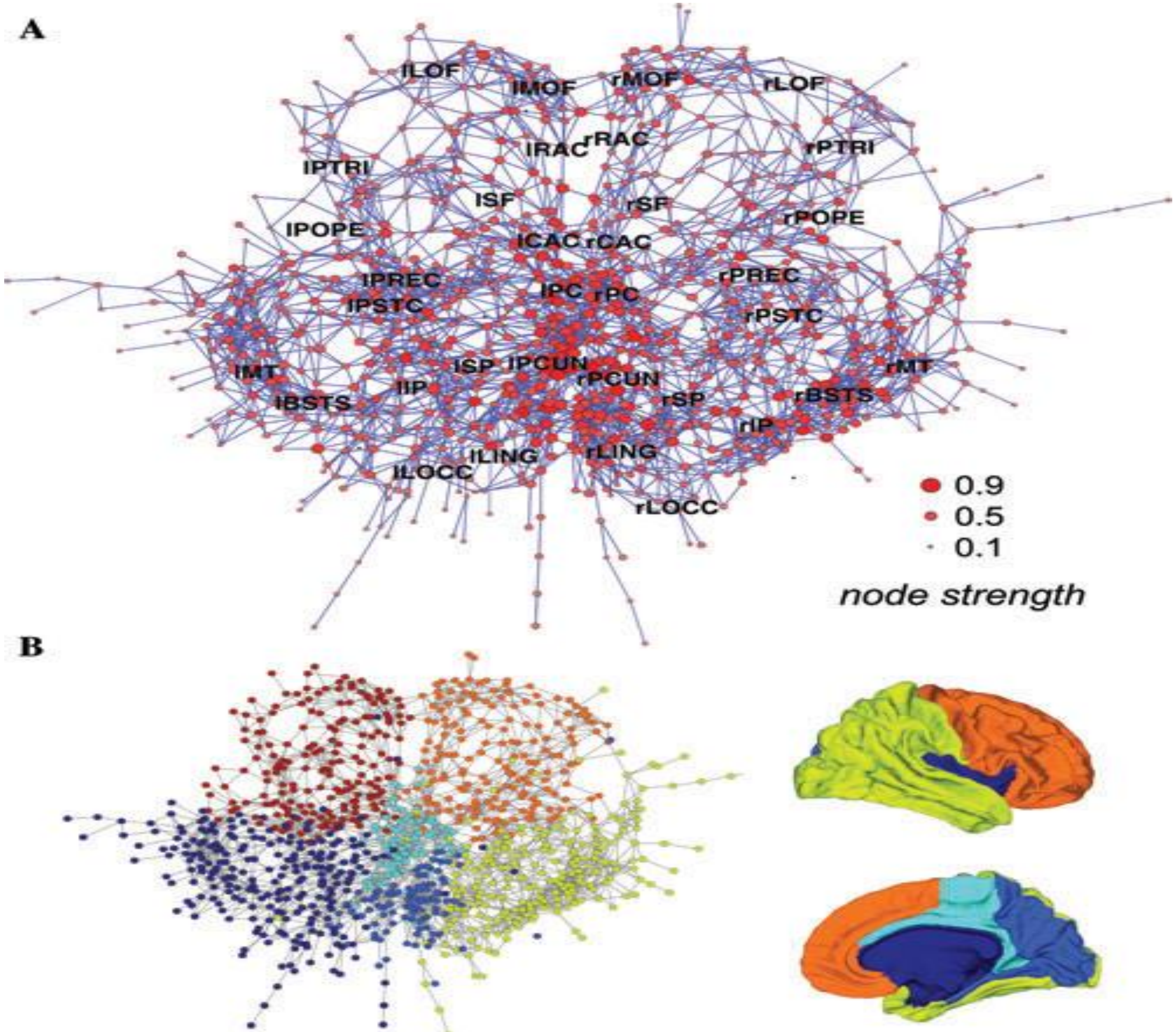
late 1890's, Vol. 1 of The Complete Psychological Works of Sigmund Freud, first published only in 1950.

Assumptions

- Mental states result from the global cooperation of the distributed neural cell activity in the brain. A global emergent state generates a bodily response, an *act*.
- Memory is encoded in the architecture of the neural net of the brain. Information is recorded by reconfiguring the net, learning.
- Memory traces are retrieved through an associative memory mechanism.
- Each brain state represents only one mental state. Each symbol is associated to one meaning.

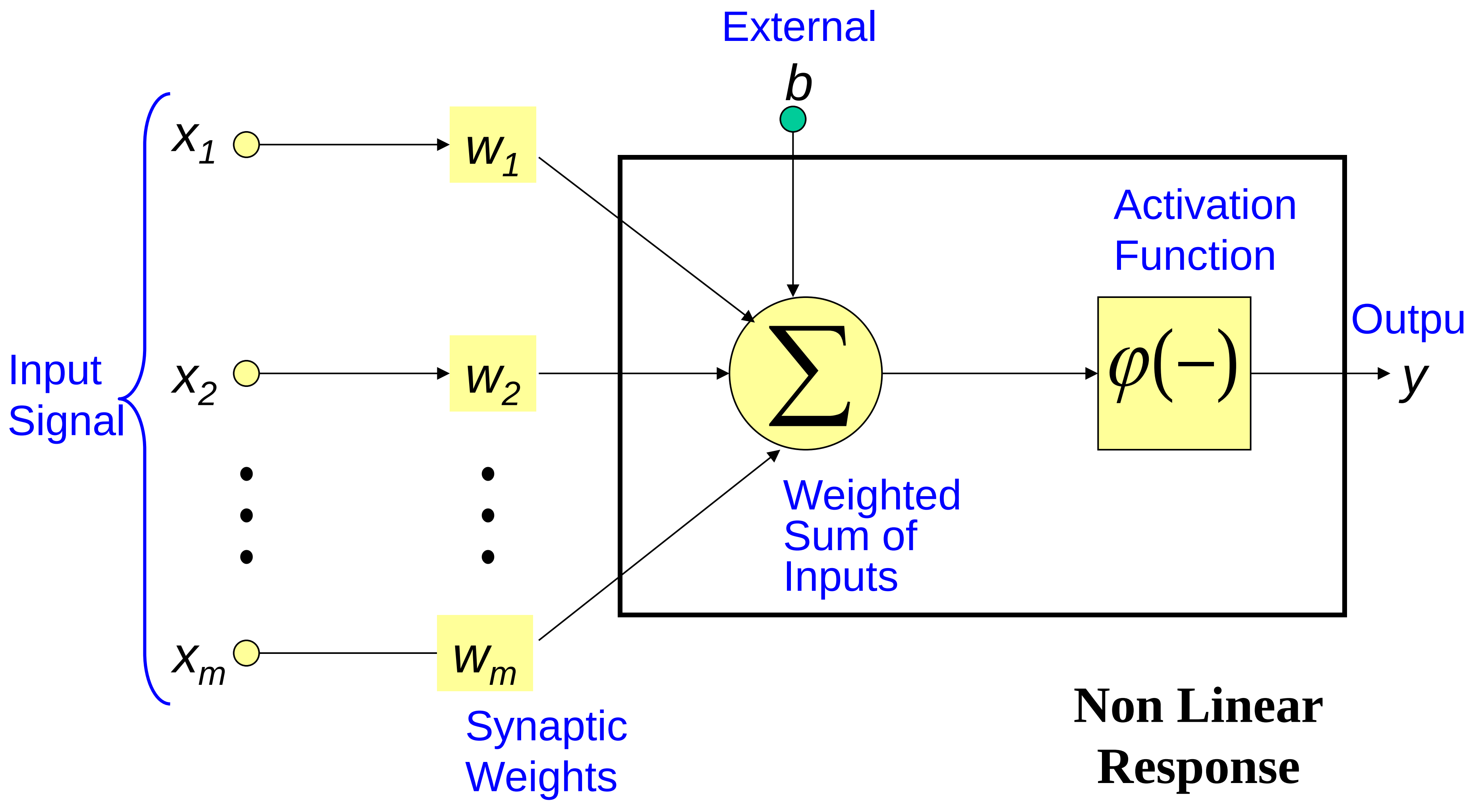
Biological Neuron





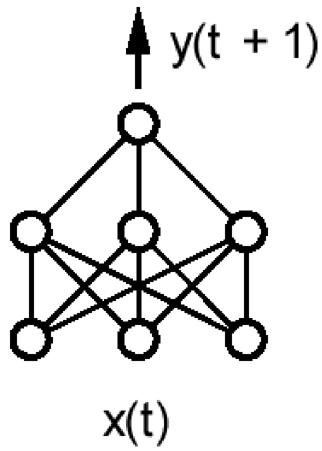
Olaf Sporns, The human connectome: a complex network, *Ann. N.Y. Acad. Sci.* 1224 (2011) 109–125 c 2011, New York Academy of Sciences.

Computacional Model of Neuron

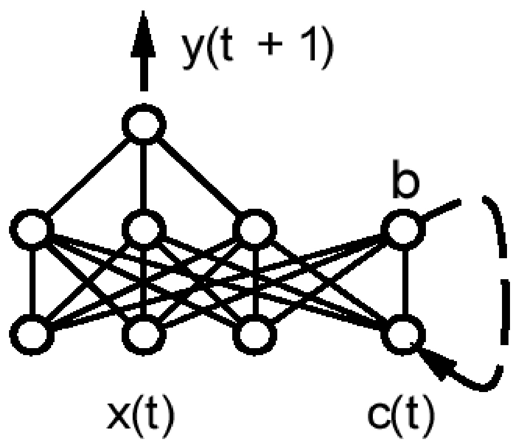


Non Linear Response

Models of Neural Networks



(a) Feedforward network



(b) Recurrent network

Complex Systems

Freud and Neuroses[1]

- Traumatic or repressed memories are knowledge which is present in the subject, but which is momentarily or permanently inaccessible to his consciousness: unconscious knowledge.

- Neurotic patients systematically repeat symptoms in the form of ideas and impulses: compulsion to repeat, related to the repressed memories.

- Neurotics have obtained relief and cure from strong neurotic symptoms through a mechanism called working-through: constructing conscious knowledge of the repressed and understanding and changing the compulsion to repeat through transference → creativity.

  Freely talking, analyzing dreams, etc...

[1]**S. Freud, "*Remembering, Repeating and Working-Through*," *The Complete Psychological Works of Sigmund Freud*, Standard ed. The Hogarth Press, London, 1953, Vol. 12, 1st German ed. in 1914.**

J. Lacan, *O Seminário, Livro 8: A Transferência* Jorge Zahar, Rio de Janeiro, Brazil, 1992, 1st French ed. in 1991.

Functional Model for Neuroses

Neuroses manifest themselves as an associative memory process: network returns a stored pattern, when it is shown another input pattern sufficiently similar to the stored one.

Compulsion to repeat: neurotic symptom is acted when the subject is presented with a stimulus which resembles, at least partially, a repressed or traumatic memory trace, $\hat{S}$.

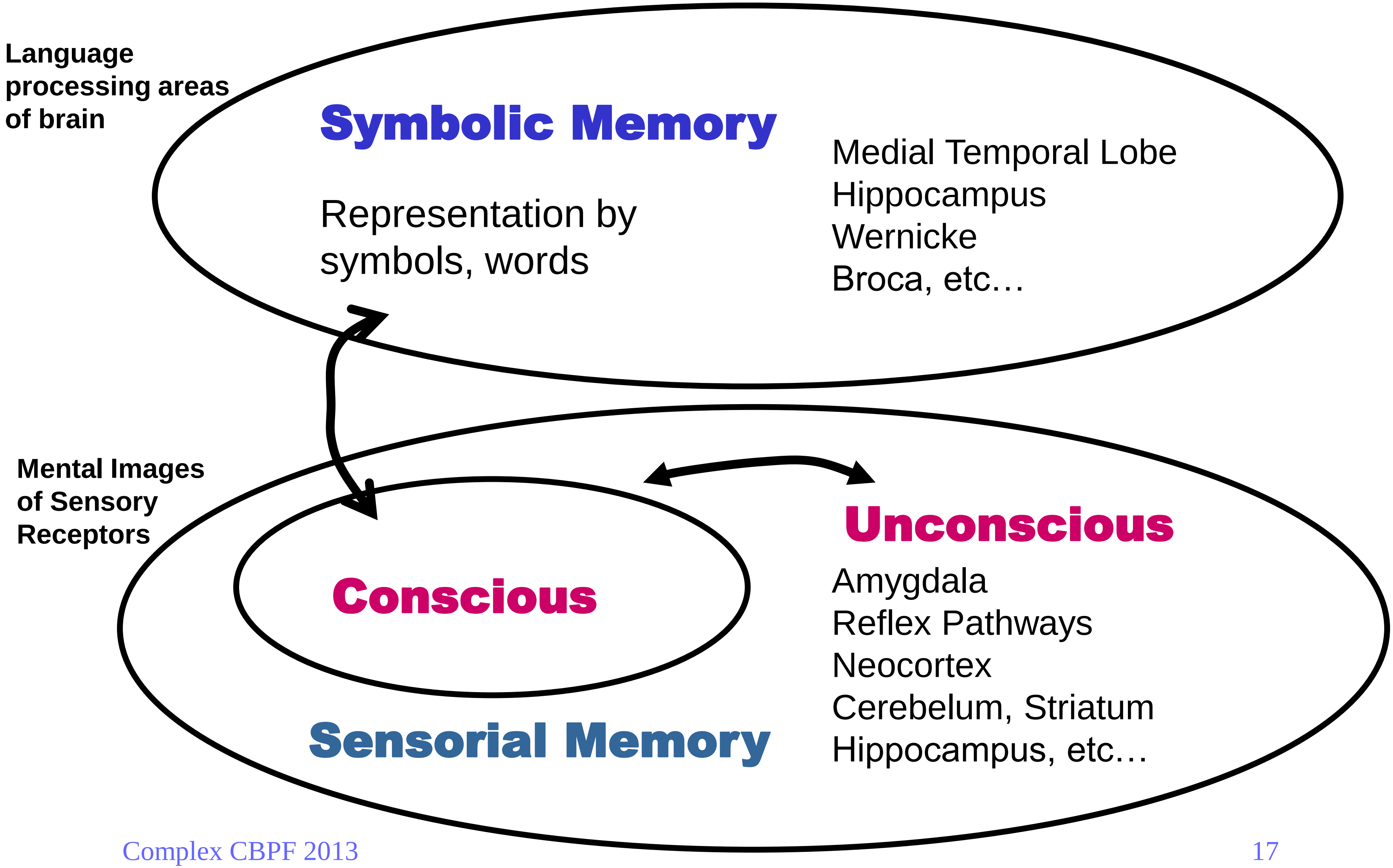
stimulus → **net stabilizes on $\hat{S}$** → **neurotic act**

Neurotic behavior: the act isn´t a result of the stimulus as a new situation, but a response to **Ŝ**.

Psychoanalytic working-through:

linguistic, symbolic associative process, language → **reinforcing synapses among memory traces in brain (declarative memory, consciousness)**

Architecture for Conscious / Unconscious Processes



- A trace in sensorial memory may “become conscious” if associated to a pattern in symbolic memory.

- Symbolic memory areas associated with language → we can associate a word (symbol) such as “red” to the visual sensation of seeing a red object.

- Access to symbolic memory represents Freud’s concept of conscious / unconscious mental processes (preconscious) and role of language in psychoanalysis. Importance of representation, symbolization in mental phenomena.

- Similar to ideas and models obtained from neurophysiology and cognition (Cleeremans[1], Changeux[2], Edelman[3], Baars[4]).

[1] A. Cleeremans, B. Timmermans, A. Pasquali, *Consciousness and metarepresentation: A computational sketch*, Neural Networks 20, 2007, pp. 1032–1039.

[2] J. P. Changeux, *The molecular biology of consciousness investigated with genetically modified mice*, Phil. Trans. R. Soc. B, 2006, 361, pp. 2239.

[3] G. M. Edelman, *Wider than the Sky, a Revolutionary View of Consciousness*, Penguin Books, London, 2005.

[4] B. J. Baars, In the Theatre of Consciousness: Global Workspace Theory, A Rigorous Scientific Theory of Consciousness. *Journal of Consciousness Studies*, **4**, No. 4, 1997, pp. 292-309.

Cleeremans et al.

"Metarepresentations inform the agent about its own internal states, making it possible for it to develop an understanding of its own workings. And this, we argue, forms the basis for the contents of conscious experience, provided of course – which cannot be the case in any contemporary artificial system - that the system has learned about its representations by itself, over its development, and provided that it cares about what happens to it, that is, provided its behaviour is rooted in emotion-laden motivation (to survive, to mate, to find food, etc.)."

Cleeremans, A., Timmermans, B., Pasquali, A.: Consciousness and metarepresentation: A computational sketch. Neural Networks 20, 1032–1039 (2007)

Computational Model

We developed **Algorithm Neuroses**[1,2] to ilustrate these ideas.

- **Hierarquical Clustering Algorithm**: generates clustered hierarchical topology in memory networks, based on competitive and cooperative biological mechanisms: neural growth factors and Hebbian learning. Networks store neurotic traces.
- **Memory Access Mechanism**: Simulated annealing mechanisms on the network to reach stable states of the neural net → stored memory traces.
- **Working-through Algorithm**: based on Hebbian learning mechanism that regulates synaptic plasticity and reconfigures connectivity of network topology. → New stable memories.

[1]R. S. Wedemann, R. Donangelo, and L. A. V. Carvalho, Chaos 19, 015116, 2009.

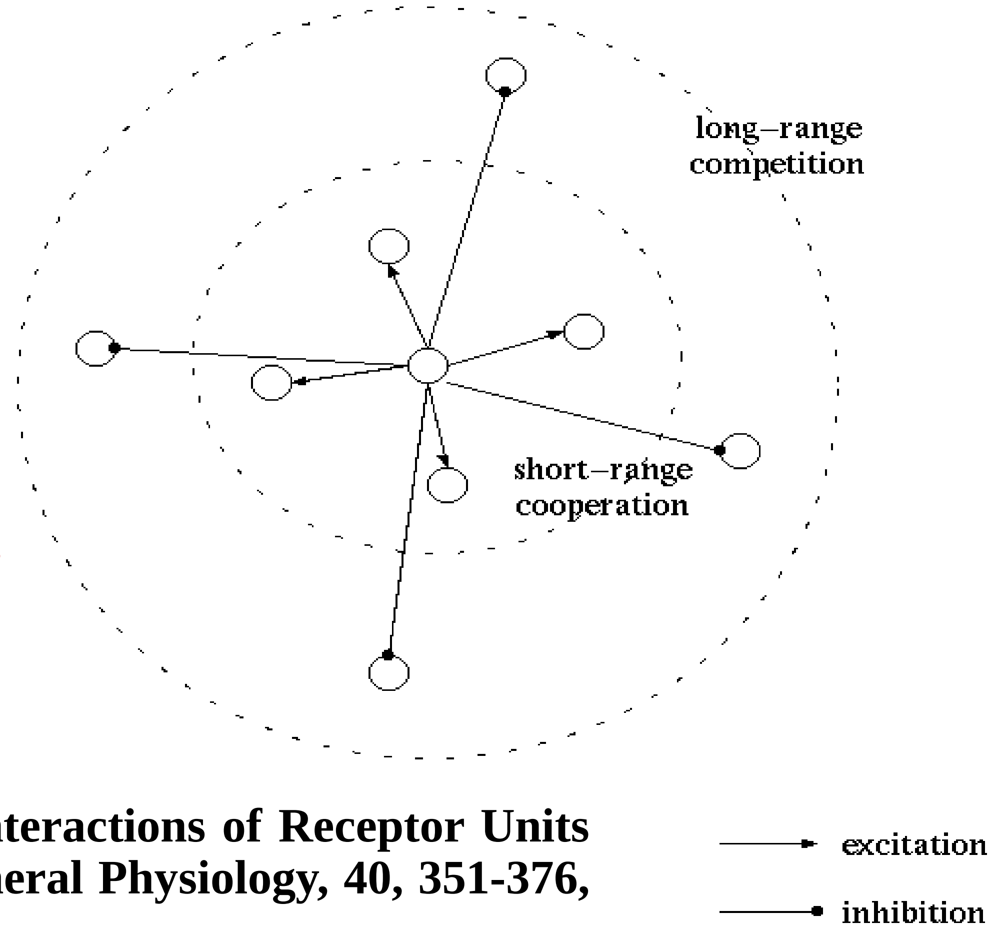
[2]R. S. Wedemann, L. A. V. Carvalho, and R. Donangelo, Neurocomputing, 71, 3367–3371, 2008.

[3]R. S. Wedemann, R. Donangelo, L. A. V. Carvalho, and I. H. Martins, Lect. Notes Comput. Sci. 2329, 236 2002.

Hierarquical Clustering Algorithm: generates structure of the topology of each memory

In many animals,
brain cells have
a structure called[1]
on-center / off-surround.

Cooperation / Competition



[1] **H. Hartline, F. Ratcliff, "Inhibitory Interactions of Receptor Units in the Eye of Limulus", Journal of General Physiology, 40, 351-376, 1957.**

- A signaling network is established to control development and plasticity of neuronal circuits.

- Competition / cooperation is controlled by environmental stimulation → this is the way environment represents itself in the brain.

- Formation of neuronal organizations (biological circuits) called maps.

- **Hebbian learning**: synaptic growth among two neurons (or two regions representing memory traces) is promoted by simultaneous stimulation of the pair.

Hierarquical Clustering Algorithm

1. Generate a uniform distribution of neurons on two 2-dimensional memory sheets.

2. Distribute synapses between nodes n_i and n_j and their respective weights, according to gaussian distribution P_G with standard deviation σ. Inspired by solution of diffusion equation of neural growth factors, for simulation of synaptic growth.

3. For each neuron n_k corresponding to a center of a region that represents a symbol generated by a previous SOM model, repeat step 4 until clustered.

4. For each n_j adjacent to n_k ,

$$prob = |\omega_{kj}| \ / \ \Sigma_l \, |\omega_{kl}|$$

If *random_number* < *prob*, increase $|\omega_{kj}|$ by

$$\Delta = \eta \ prob \qquad\qquad \eta \in (0, 1]$$

and decrease Δ from all neighbors ω_{km} , $m \neq j$ by

$$\delta_{km} = (1 - |\omega_{km}| \ / \ (\Sigma_{l \neq j} \, |\omega_{kl}|) \ \Delta$$

for all i,j , $\quad |\omega_{ij}| < \omega_{max}$

- Regulates synaptic *plasticity*, by strengthening synapses within a cluster and reducing synaptic strength between clusters (disconnects clusters).

- A kind of *preferential attachment* with conservation of total synaptic weights.

- Neurons that have received stronger sensorial stimulation (are more strongly connected), will stimulate their neighborhoods and promote still stronger connections. Inspired in Hebbian learning.

- This is a self-organized process, based only on local properties of the network → parallel/distributed process.

5. Establish synapses among clusters (long range synapses) reflecting associations among representations, within and among memories.

→ LANGUAGE.

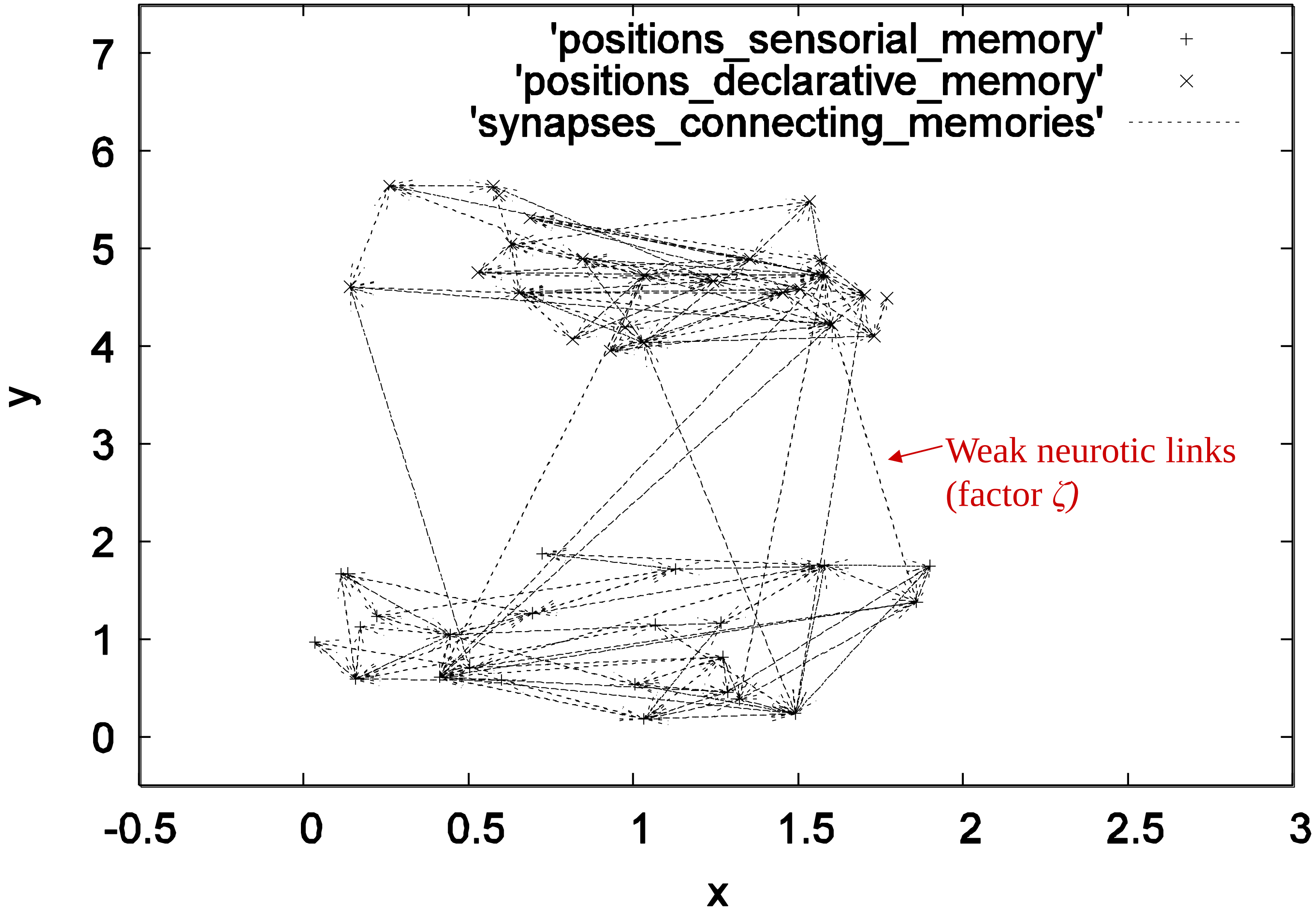
We don´t know this distribution.

Started with random and study to find something better...

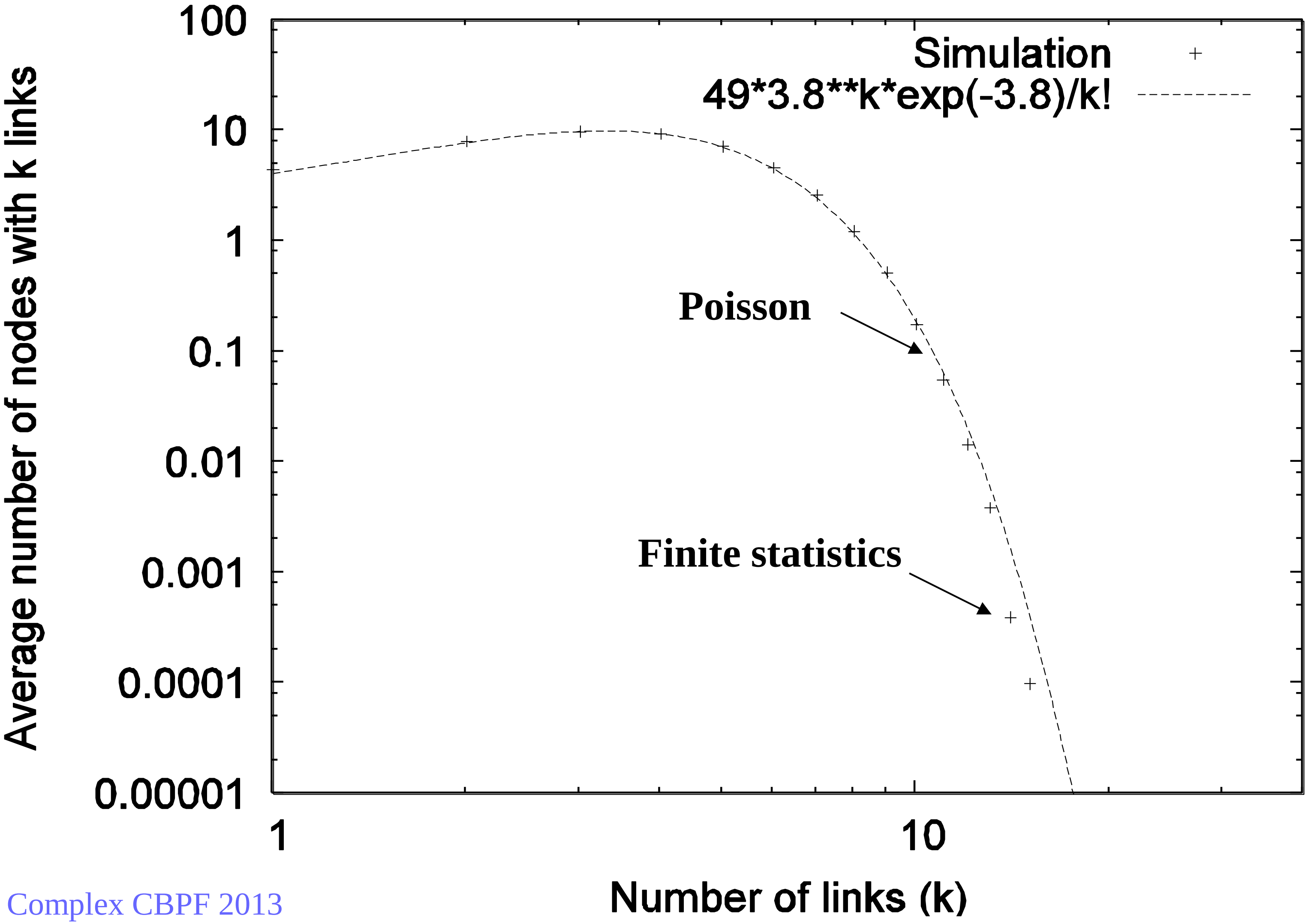
If long range synapse connects neurons in different memories, multiply by $\zeta \in (0, 1]$. → Neurotic network.

- Long range costs more energy and is less frequent.

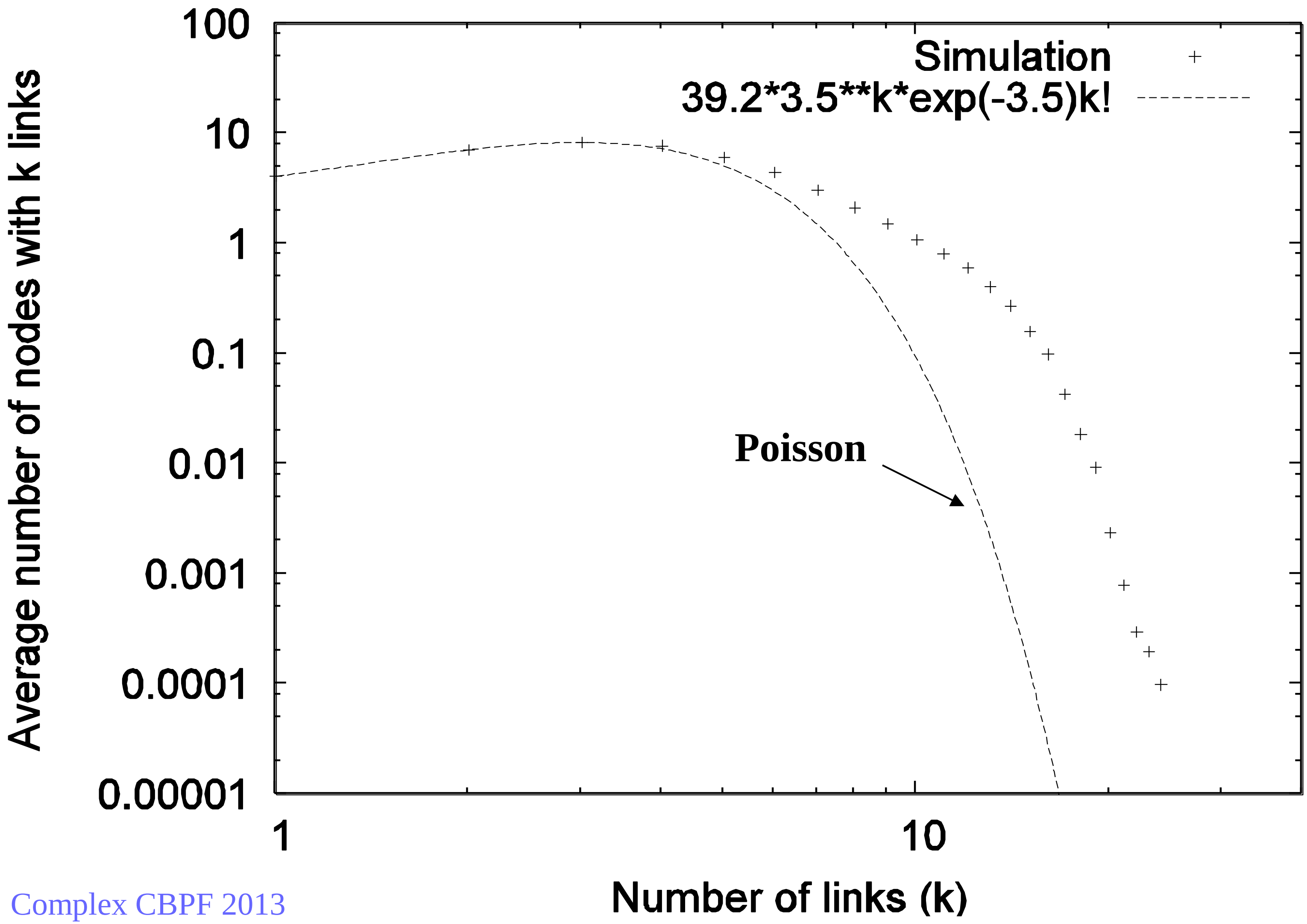
Network Topology with Long Range Synapses:
N = 50, σ = 0.58



Node Degree Distribution: Gaussian Distribution of Links (Until Step 2) N = 50, σ = 0.58



Node Degree Distribution with Clustering and Long Range: N = 50, σ = 0.58



- Random graphs follow a Poisson distribution.
  So does the Gaussian distribution of links (step 2).

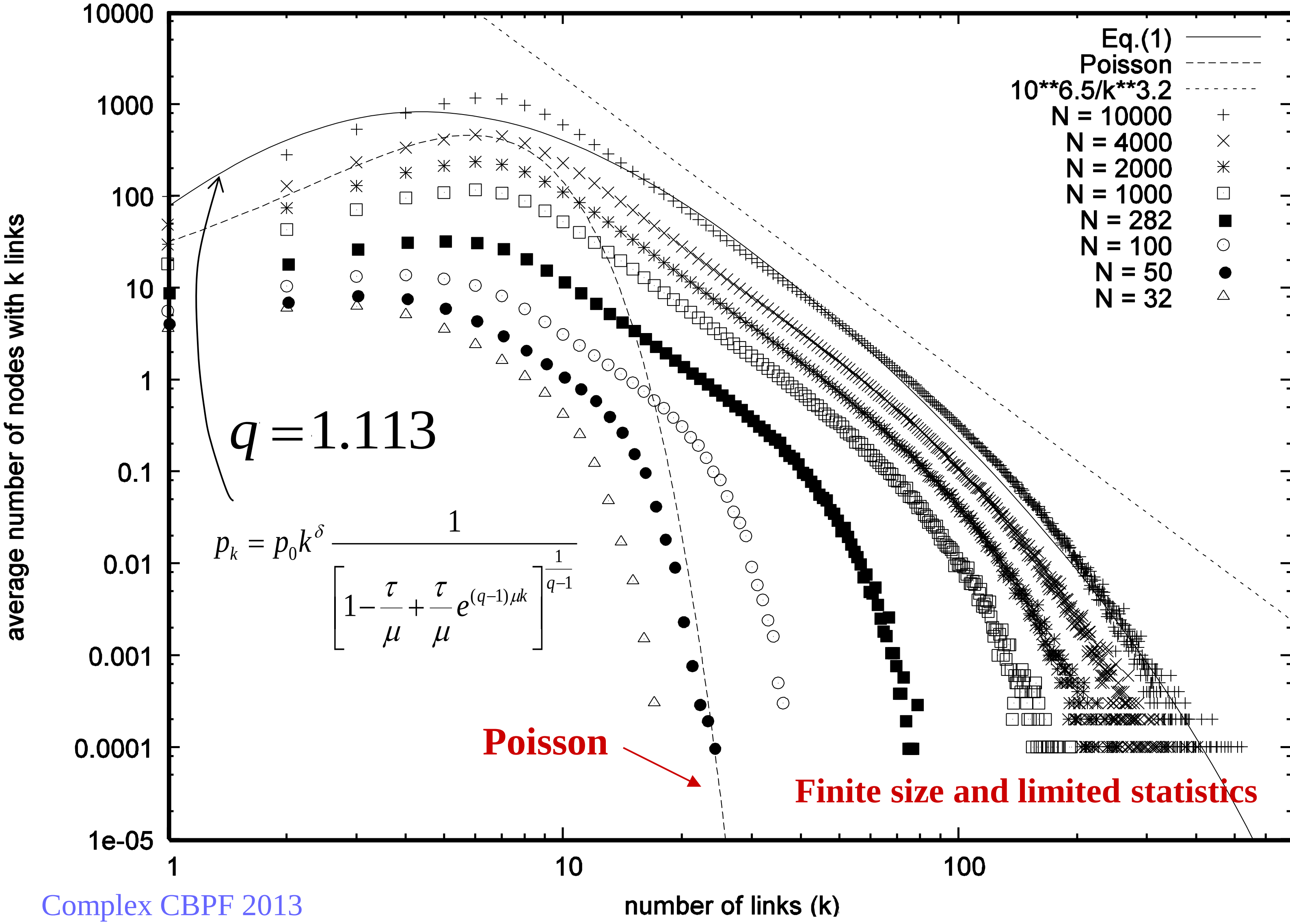
- Deviation from Poisson for large k may be attributed to competitive, associative biological mechanisms, which introduce structure (correlations).

- Average clustering coefficient for $N = 282$ and 0.24.
  For *C.Elegans* worm = 0.28, would be = 0.049 if random graph

- Network stored number of memory patterns initially. Only some remained after working-through.
  Network adapts, freeing itself from neurotic states.

- For $\zeta < 0.5$, network has learning difficulties.
  Difficulty in associating unconscious memory traces
  with symbolic memory.

Node Degree Distributions



Memory Access Mechanism : retrieves memory patterns stored by neural network topology

Originally simulated by Boltzmann Machine (**BM**) with complete graph: Pattern retrieval on net is achieved by a simulated annealing (**SA**) process, where temperature T is gradually lowered by an annealing schedule α.

N nodes with states S_i in {-1, 1}, synaptic weights $\omega_{ij} = \omega_{ji}$

Energy:

$$E(\{S_i\}) = -\frac{1}{2}\sum_{ij} \omega_{ij} S_i S_j$$

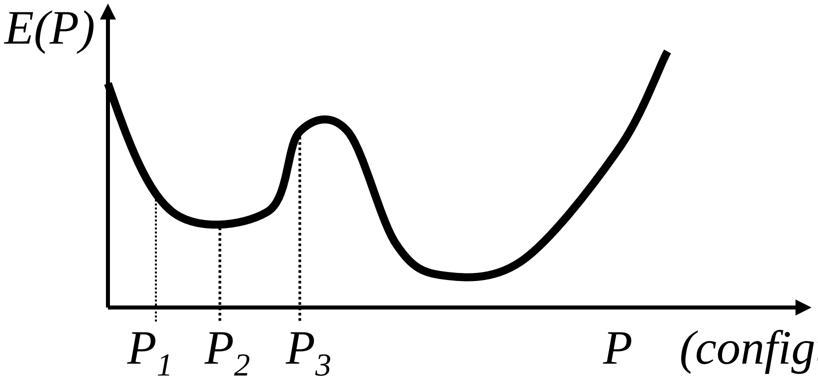
Network state distribution function is Boltzmann-Gibbs (BG):

$$P_{BG}(\{S_i\}) = \frac{\exp\left[\frac{-E(\{S_i\})}{T}\right]}{\sum_{\{S_i\}} \exp\left[\frac{-E(\{S_i\})}{T}\right]}$$

T is network temperature

Boltzmann Machine - Simulated Annealing
Global / Local Minima in the Energy Function

Global / Local minima



Hopfield: gradient descent finds closest local minimum

Boltzmann: Start at initial pattern P_1

Choose a neighbor P'

If $E(P') < E(P_1)$, accept P' as new solution

Else accept P' with probability

$$P_{BG}(P_1 \rightarrow P') = \exp\left(\frac{-(E(P') - E(P_1))}{T}\right)$$

Memory: Boltzmann Machine - Simulated Annealing Global / Local Minima in the Energy Function

For a given memory pattern P presented to the net, do

$y = P$;

$T = T_0$;

While $T \geq T_f$

For $i = 1$ to No. iterations

For $j = 1$ to N do

generate P';

if ($E(P') < E(y)$) then $y := P'$;

else Assign P' to y with probability

$$P_{BG}(y \to P') = \exp\left(\frac{(E(y) - E(P'))}{T}\right)$$

$T := r(\alpha, T)$; (Annealing)

- At zero temperature, doing gradient descent on the energy surface, is an approximation to following the dynamics of the integrate and fire neurons. Following the dynamics necessarily takes you downhill on the energy surface. For binary neurons, dynamics is given by

$$S_i(t+\Delta t) = \mathrm{sgn}\left(\sum_j w_{ij} S_j(t)\right) \quad .$$

- At non-zero temperatures, following the dynamics with some fluctuations (noise) should be equivalent to the Boltzmann machine, if the BG distribution reflects the probability of the system to occupy possible states in phase space.

- The **BG** distribution function favors changes of states with small increases in energy, so that the **BM** will strongly prefer visiting state space in a nearby energy neighborhood from the starting point.
- No theoretical indication of the exact relation between network topology and memory dynamics.
- Power-law behavior for node degree distribution and long-range synapses suggest using a Generalized Statistical Mechanics in the annealing process[3]. This changes the thread of associations of the model.

[1]Hertz, J.A., Krogh, A., Palmer, R.G. (ed.), *Introduction to the Theory of Neural Computation*, Lecture Notes, Vol. I, Santa Fe Institute, Perseus Books, 1991.

[2]V. C. Barbosa, *Massively Parallel Models of Computation*, Ellis Horwood, England, 1993.

[3]Tsallis, C., *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*, Springer, 2009.

Generalized Simulated Annealing (GSA)

Change from P_{BG} to

$$P_q(y \rightarrow P') = \frac{1}{\left[1 + (q-1)\dfrac{(E(y) - E(P'))}{T}\right]^{1/(q-1)}}$$

C. Tsallis and D. A. Stariolo, Physica A, 233, 395—406, (1996).

Memory Associativity Experiment

Perform up to 10000 minimization procedures, starting each one from a different random network configuration, which is presented to both the **BM** and **GSA**.

When a new minimum energy pattern is found, it is stored and the procedure is repeated from other random starting configurations, otherwise the search stops.

Temperature, T and q associated with noise, allowing associativity among memory traces.

- Quantities of neurotransmitters and neuromodulators in brain.
- Regulates synaptic inhibition ⟷ freely talking in analytic sessions and stimulation from analyst lower resistances and allow greater associativity.

- **Catecholamines (dopamine, norepinephrine, etc.) alter signal-to-noise ratio at neuronal level, regulating the way the brain performs associations. Temperature and noise in BM and GSA ⟷ associativity of thought.**

↑dopamine (↓T) → enhances signal → higher signal-to-noise ration → weaker associativity → inflexible thought

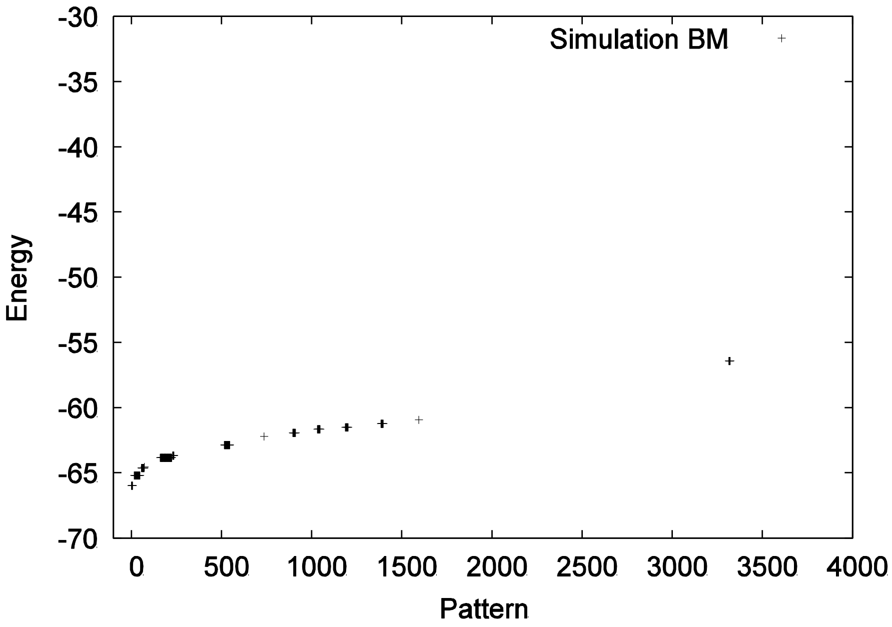
↓dopamine(↑T) → increases noise → more associativity → distant ideas, disorganized thought…

disorganized thought | **creative thought** | **fixations to ideas and delusions**

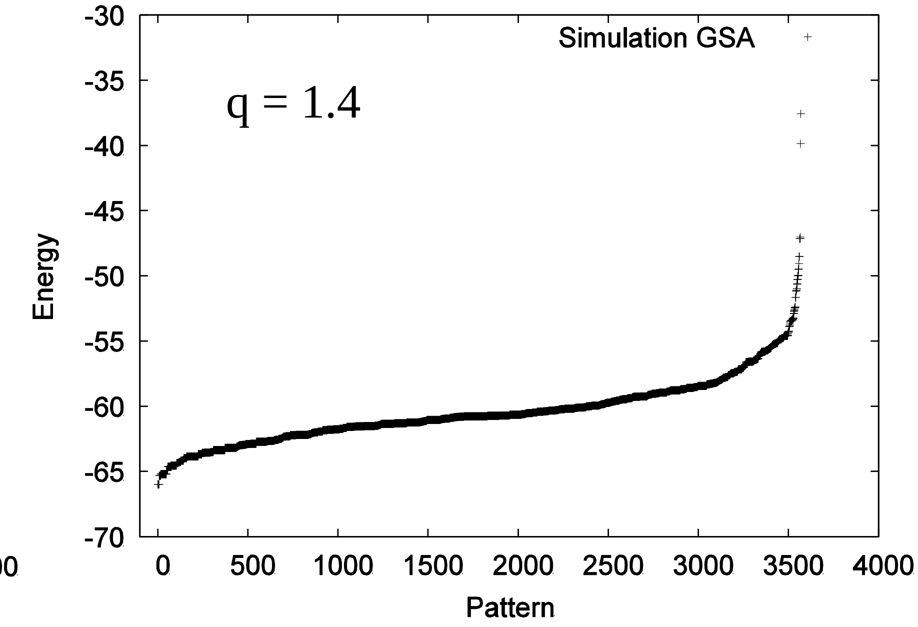
→

increase dopamine

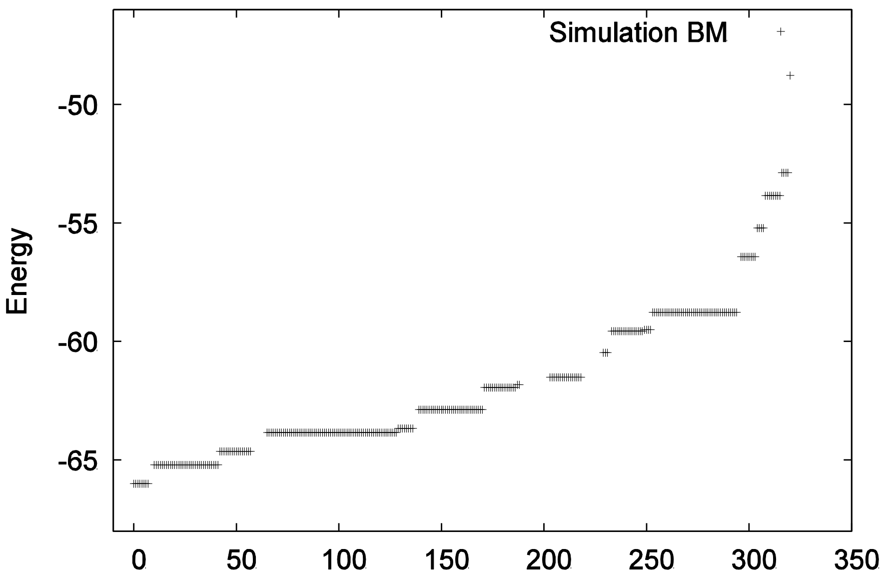
T = 0.2



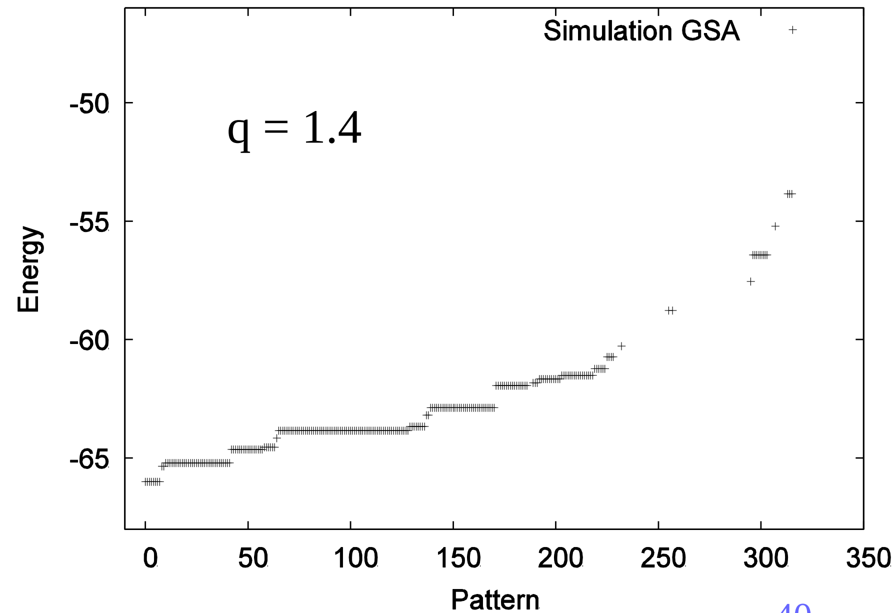
Simulation BM
Energy
Pattern
-30
-35
-40
-45
-50
-55
-60
-65
-70
0
500
1000
1500
2000
2500
3000
3500
4000



Simulation GSA
q = 1.4
Energy
Pattern
-30
-35
-40
-45
-50
-55
-60
-65
-70
0
500
1000
1500
2000
2500
3000
3500
4000



Simulation BM
Energy
Pattern
-50
-55
-60
-65
0
50
100
150
200
250
300
350



Simulation GSA
q = 1.4
Energy
Pattern
-50
-55
-60
-65
0
50
100
150
200
250
300
350

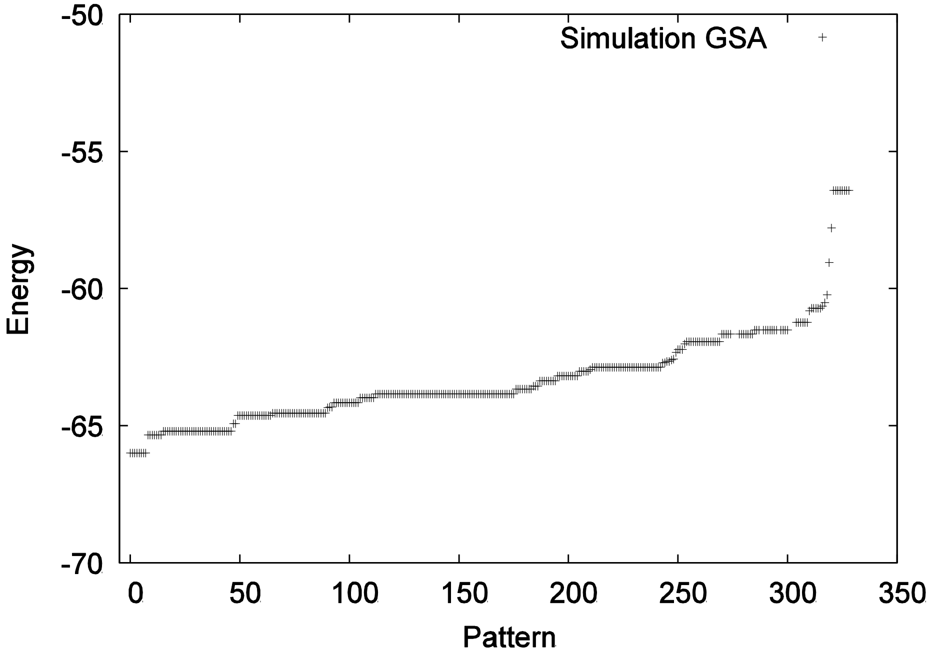
T = 0.1

- **GSA** appears to visit state space more loosely at higher temperatures, while the **BM** visits state space more uniformly at lower temperatures.

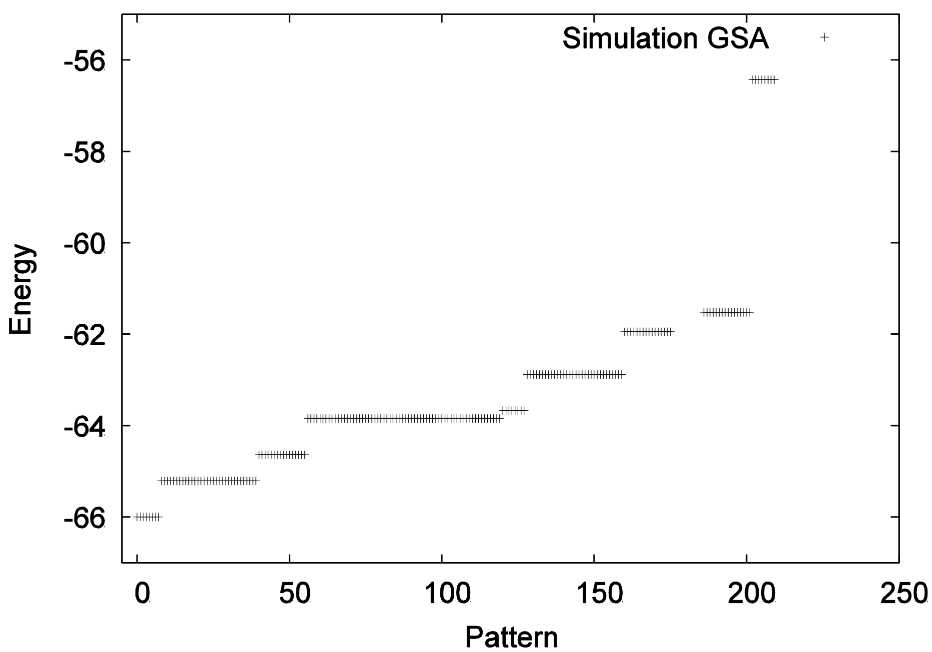
- For lower temperatures, the **BM** functions more like a gradient descent method, and randomly generated patterns will stabilize at the closest local minima.

T = 0.2

q = 1.3

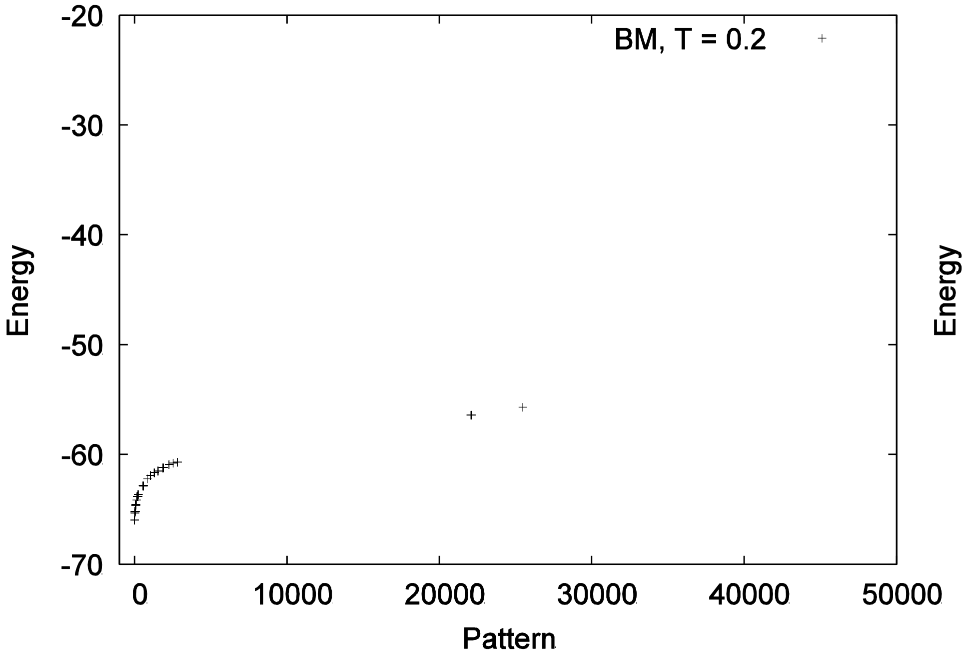


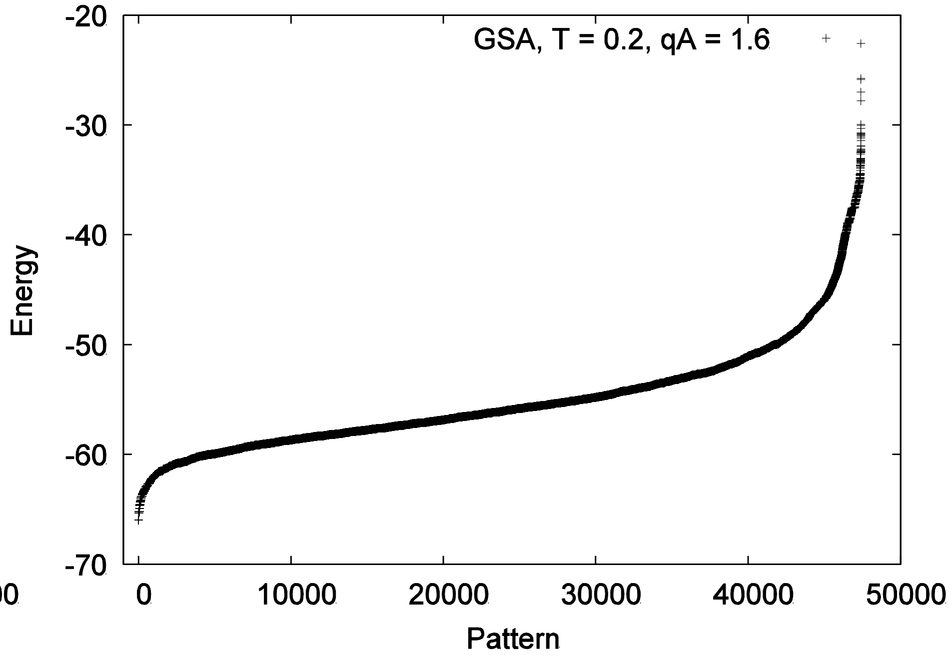
q = 0.7

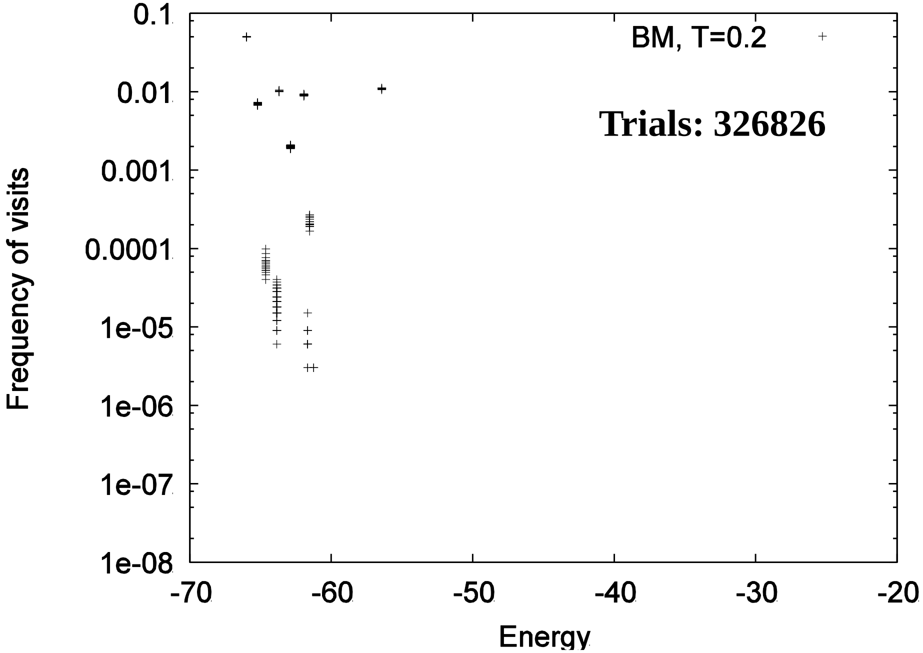


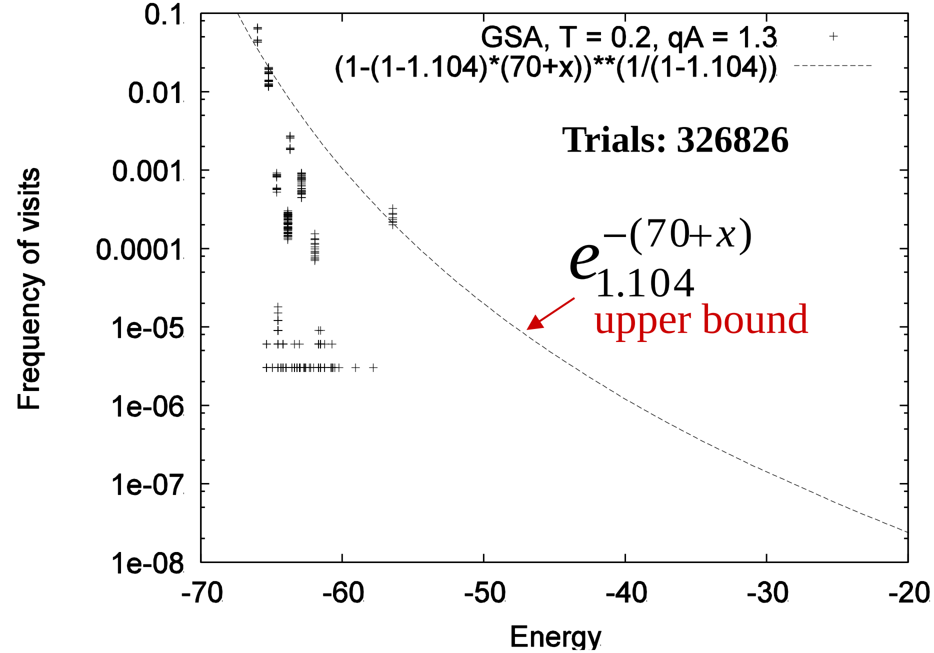
When we lower the value of q, **GSA** becomes less associative and finds less patterns.

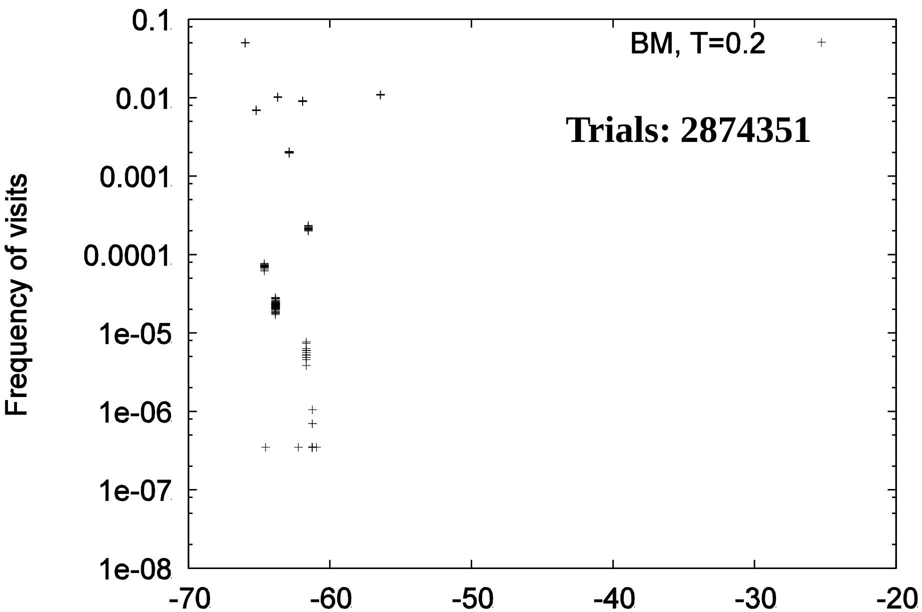
$q = 1.6$

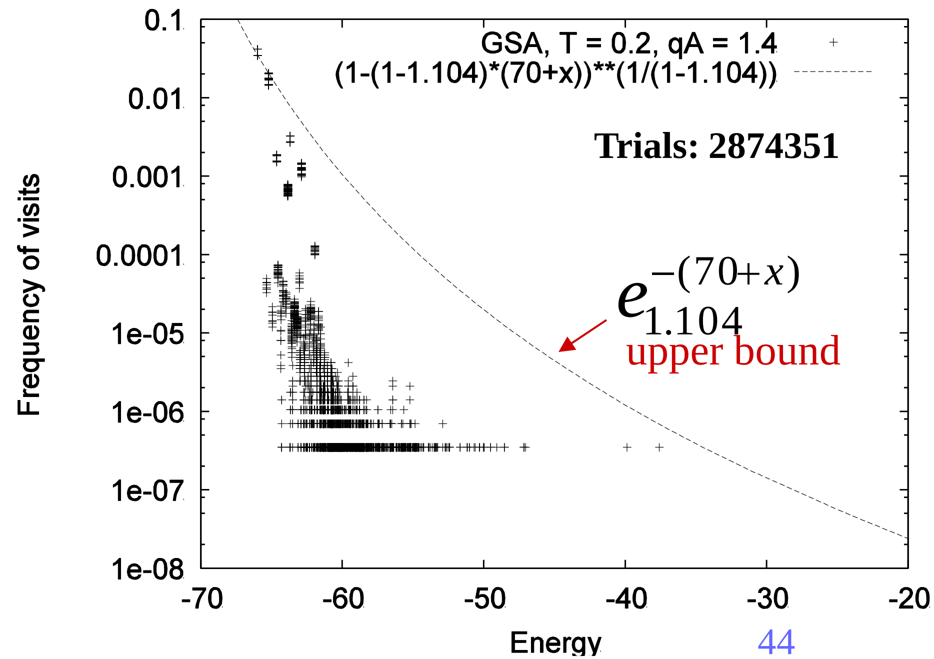

BM, T = 0.2
Energy
Pattern
-20
-30
-40
-50
-60
-70
0
10000
20000
30000
40000
50000


GSA, T = 0.2, qA = 1.6
Energy
Pattern
-20
-30
-40
-50
-60
-70
0
10000
20000
30000
40000
50000

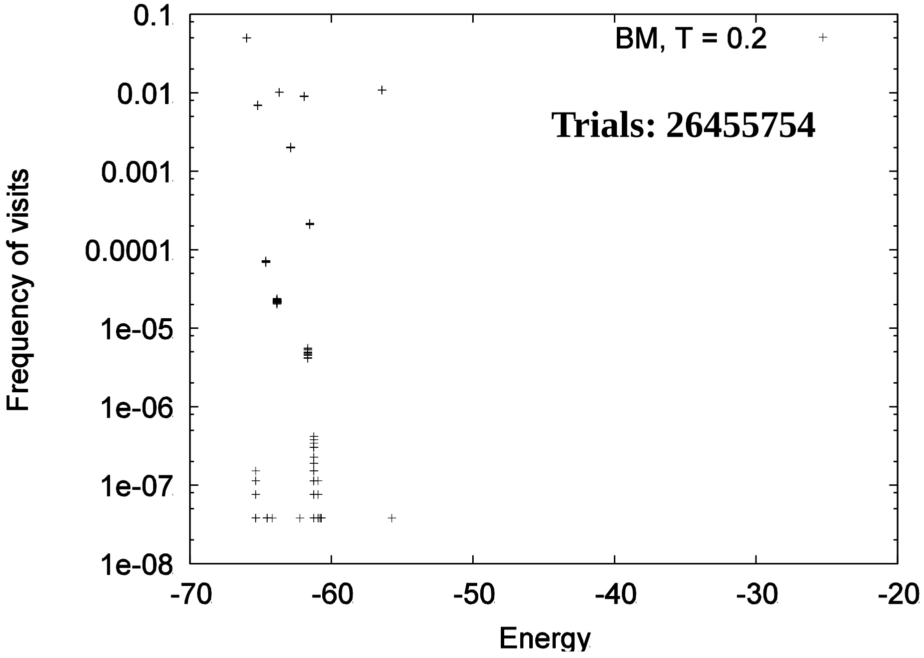

BM, T=0.2
Trials: 326826
Frequency of visits
Energy

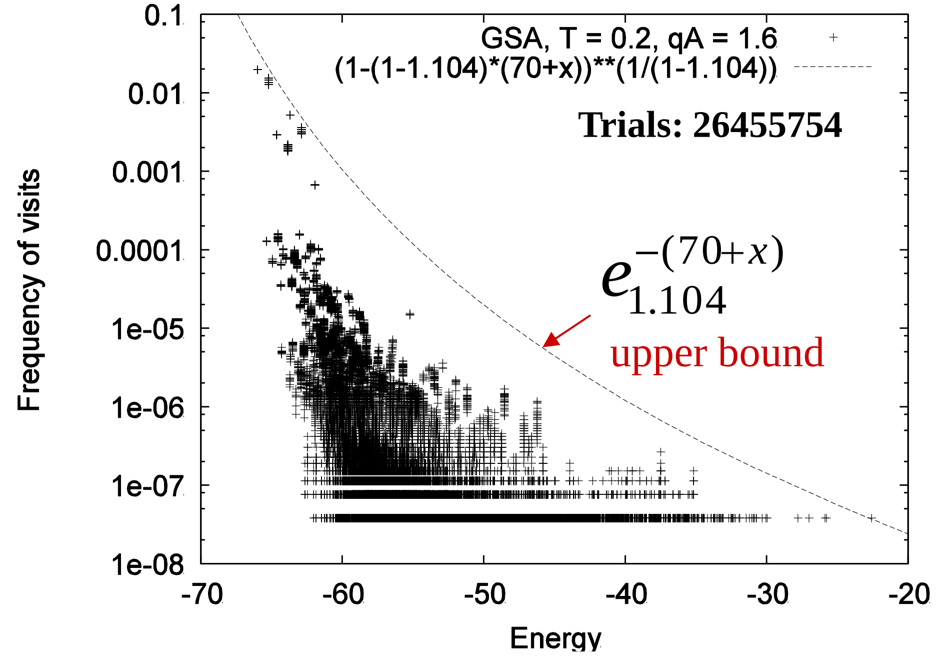

GSA, T = 0.2, qA = 1.3
(1-(1-1.104)*(70+x))**(1/(1-1.104))
Trials: 326826
$e_{1.104}^{-(70+x)}$
upper bound
Frequency of visits
Energy


BM, T=0.2
Trials: 2874351
Frequency of visits
Energy


GSA, T = 0.2, qA = 1.4
(1-(1-1.104)*(70+x))**(1/(1-1.104))
Trials: 2874351
$e_{1.104}^{-(70+x)}$
upper bound
Frequency of visits
Energy

$$e_q^x \equiv \begin{cases} [1+(1-q)x]^{\frac{1}{1-q}}, & \text{if } 1+(1-q)x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$



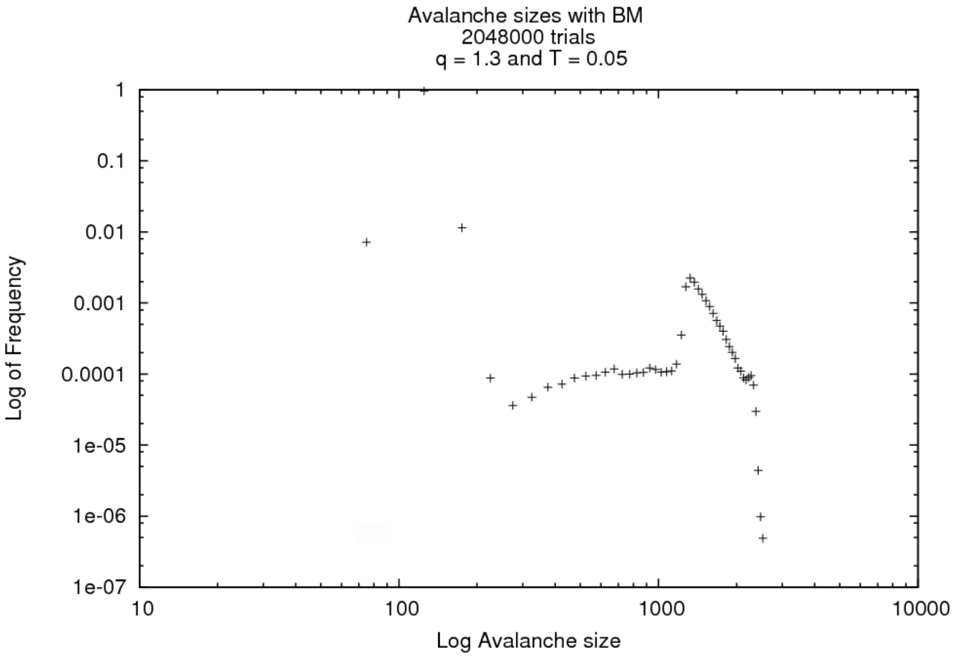


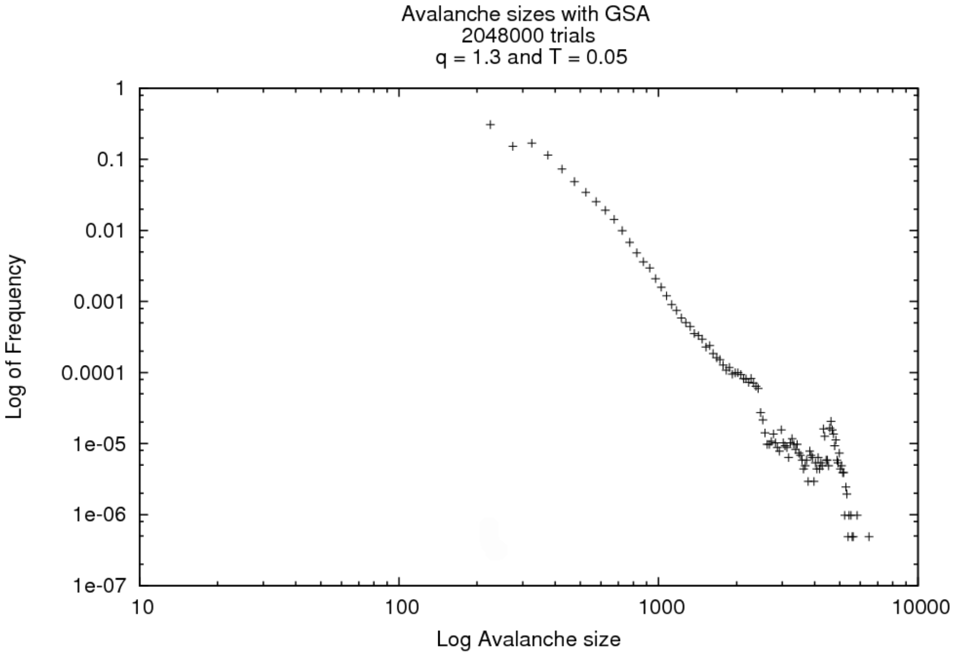
With **GSA**, the frequency with which the hardest to detect patterns are found is much larger than the corresponding ones in the **BM**. In particular, many patterns that are not found by the **BM** are detected employing **GSA**.

There is a q-exponential upper bound for the frequency of visits, as a function of energy for **GSA**.

Avalanche Size: Number of steps with state changes to reach minimum energy $\leftrightarrow$ time for propagation of signal to access information in memory.

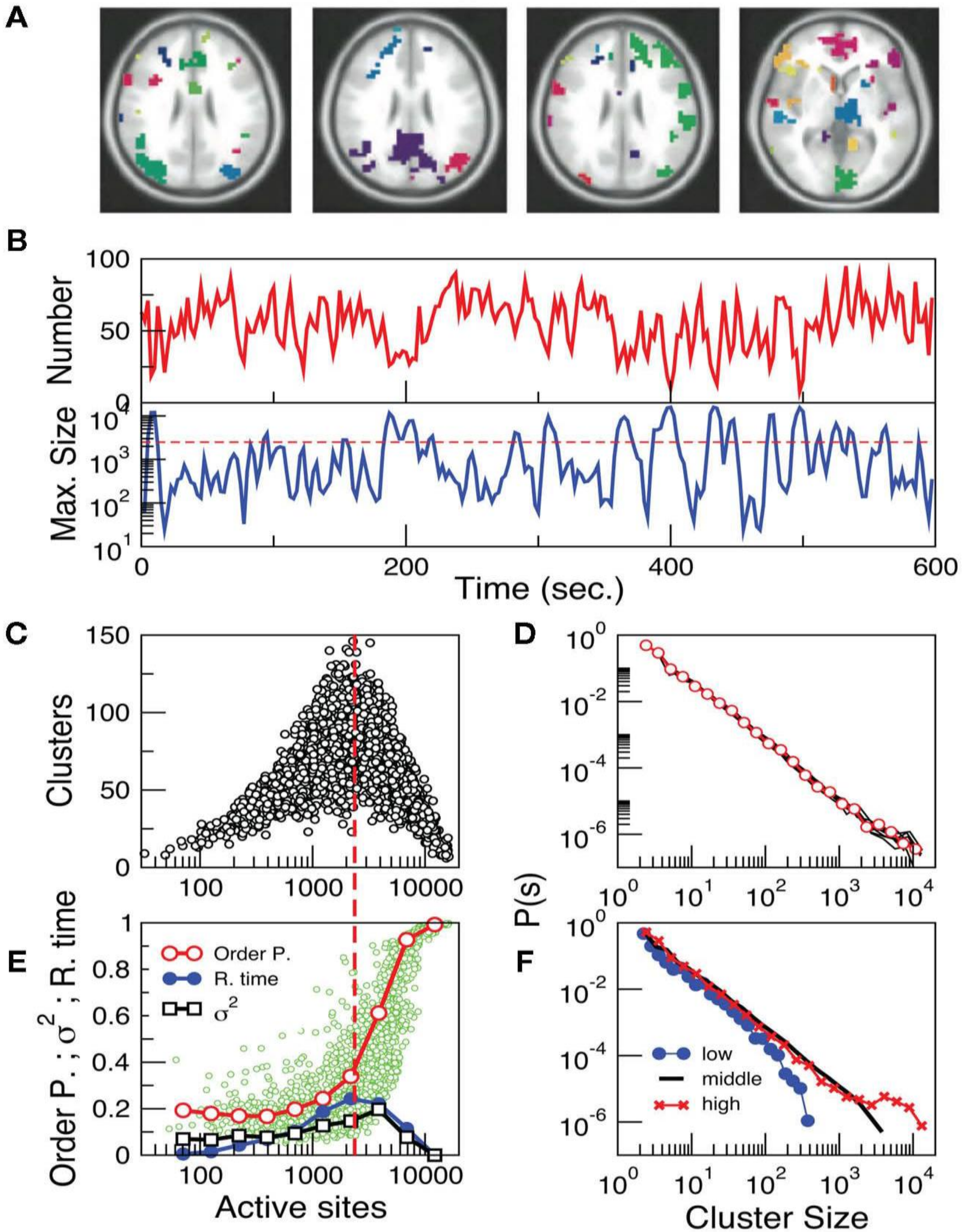
Distribution of avalanche sizes



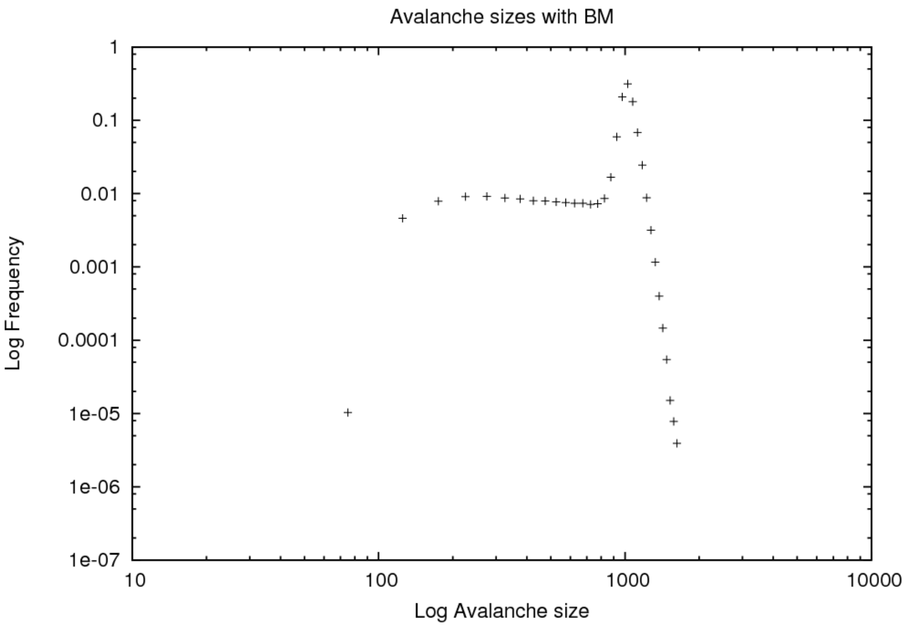


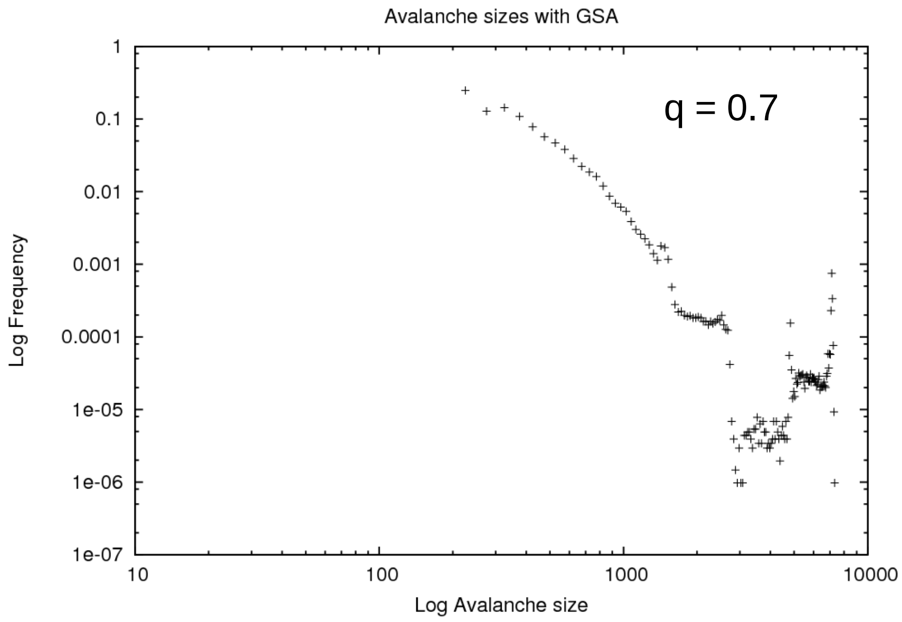
fMRI experiments reported in:

EnzoTagliazucchi, Pablo Balenzuela, Daniel Fraiman and Dante R. Chialvo, *Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis*, **Frontiers in Physiology** | Fractal Physiology **3**, Article 15, 2012



T = 0.2


Avalanche sizes with BM
Log Frequency
Log Avalanche size
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
10
100
1000
10000


Avalanche sizes with GSA
q = 0.7
Log Frequency
Log Avalanche size
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
10
100
1000
10000

- If memory access functions as GSA, with long-range connections, the neuronal system should be capable of making more distant associations among stored traces ⟶ more metaphors and creativity - *'homo metaphoricus'*[1].

- Values of T and q regulate these capabilities, reflecting network properties: ex. availability of neuromodulators and neurotransmitters.

- More associative networks enhance capabilities for self-reconfiguration and learning process (change in network connectivity by working-through) ⟶ new outcomes in the subject´s life history.

[1]Tsallis, Tsallis, Magalhães, Tamarit, *Human and Computer Learning: An Experimental Study*, Complexus 2003;1:181–189.

- We are currently investigating the relation of our neuroses model to cognitive models for attention and consciousness.

- Freud's observations regarding the unconscious may give us important insight regarding basic neuronal mechanisms underlying consciousness.

Dehaene, S., Kerszberg, M., Changeux, J.P.: A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks. Proc. Natl. Acad. Sci. **95**, 114529–14534 (1998)

Changeux, J.P.: The Molecular Biology of Consciousness Investigated with GeneticallyModified Mice. Phil. Trans. R. Soc. B **361**, 2239–2259 (2006)

Taylor, J.G.: A Neural Model of the Loss of Self in Schizophrenia. Schizophrenia Bull; **0**: sbq033V1–sbq033 (2010)

Freud, S.: The Unconscious. Standard Edition, Vol. XIV. The Hogarth Press, London,(1957). First German edition (1915)

Thank you!

Special thanks to

- Angela Bernardes, Romildo Rego de Barros and Paulo Vidal from Escola Brasileira de Psicanálise (EBP).
- Sumiyoshi Abe, Evaldo Curado and Constantino Tsallis.
- Brazilian research financing agencies: CNPq, CAPES and FAPERJ.
- T2 Hepgrid Brazil – UERJ Laboratory for high performance computing (cluster for distributed computing).
- Complexity and Networks Group at Imperial College London.

Conclusions

Common Features in Models of Consciousness

- If a stimulus gains access to a specific neuronal subset, it becomes conscious. In CODAM, Working Memory (WM) and in DKC-GW it is the Global Workspace (GW). Stimuli that intensely mobilize access to WM or GW are available for operations such as thinking, reasoning, verbal report and symbolization, evaluation, memorization and action guidance, associated with the subjective feeling of conscious effort.
- Mechanisms for selecting (setting goals and gating) a subset of input stimuli that will gain access to consciousness, functioning as an attentional filter, which amplifies the attended stimulus and inhibits distracters.
- Mechanisms for sustaining the attended stimulus for a certain period of time in WM or GW are also included in both models.
- Stimuli and memory traces which have inhibited access to WM and GW constitute the unconscious.
- Error correction mechanisms for guaranteeing that the attended stimulus corresponds to specified goals.

What can Psychopathologies tell us about Consciousness?

- Our model illustrates how Freud's ideas regarding the unconscious strongly indicate that symbolic processing, language and meaning are essential for the emergence of consciousness.
- In neurotic functioning, the subject is strongly attracted (as in an attentional mechanism) to repressed memories, of which he has no consciousness and which determine his acts.
- As in Cleeremans et al., we hypothesize that a second level network which is capable of producing metarepresentions of traces stored in a lower level network is necessary for the emergence of conscious experience. → Higher Order Thought perspective.
- Functions such as integration, differentiation and selection of information, along with synchronization mechanisms of different brain areas, are present also in unconscious mental activities, such as the production of neurotic acts and dreamwork. This suggests that, although these mechanisms should be important for the production of consciousness, they are not sufficient.

Cleeremans et al.

"Metarepresentations inform the agent about its own internal states, making it possible for it to develop an understanding of its own workings. And this, we argue, forms the basis for the contents of conscious experience, provided of course – which cannot be the case in any contemporary artificial system - that the system has learned about its representations by itself, over its development, and provided that it cares about what happens to it, that is, provided its behaviour is rooted in emotion-laden motivation (to survive, to mate, to find food, etc.)."

Cleeremans, A., Timmermans, B., Pasquali, A.: Consciousness and metarepresentation: A computational sketch. Neural Networks 20, 1032–1039 (2007)

- A model of consciousness should not only guarantee that a stimulus be available for report through an attentional mechanism to become conscious, but also that, in order to be conscious of an attended stimulus, a subject should be capable of reporting, by having the capability of creating and associating metarepresentations or abstract symbols to sensory information.

- This attachment of symbolic processing to sensorial input illustrates a phenomenological view of mind, where consciousness is deeply rooted in the subject's sensorial experience within the environment and his body. Symbols are associated to meaning.

The model

- is biologically plausible and in agreement with aspects of psychoanalytic clinical experience.
- is illustrative, very schematic and far from explaining the complexities of brain / mental processes. → Develop better relation to research in neurophysiology: Models by Edelman, Baars and Changeux.
- is in lack of experimental data for verification. Plan experiments. Future work
- seems to be a good metaphorical view of some basic concepts of neurotic behavior and the unconscious, as described by Freud.
- suggests directions of search for a neuronal substratum.
- Generalized simulated annealing for simulation of memory dynamics changes thread of associations in thought. Change q and T → change associativity, creativity, pathology.

- System is small. Simulations are slow in current processors.
  → Purpose is to illustrate basic concepts and mechanisms at semantic level.

- Short range microscopic mechanisms are scalable.
  → Mapping to biological substratum.

- Needs more realistic distributions for links related with language.

- We are investigating the possibility of creating a memory model with asymmetric weights among neurons.

- Parallelization of algorithms for larger systems. Future work.

Thank you!

Special thanks to

- Angela Bernardes, Romildo Rego de Barros and Paulo Vidal from Escola Brasileira de Psicanálise (EBP).
- Sumiyoshi Abe, Evaldo Curado and Constantino Tsallis.
- Brazilian research financing agencies: CNPq, CAPES and FAPERJ.
- T2 Hepgrid Brazil – UERJ Laboratory for high performance computing (cluster for distributed computing).
- Complexity and Networks Group at Imperial College London.

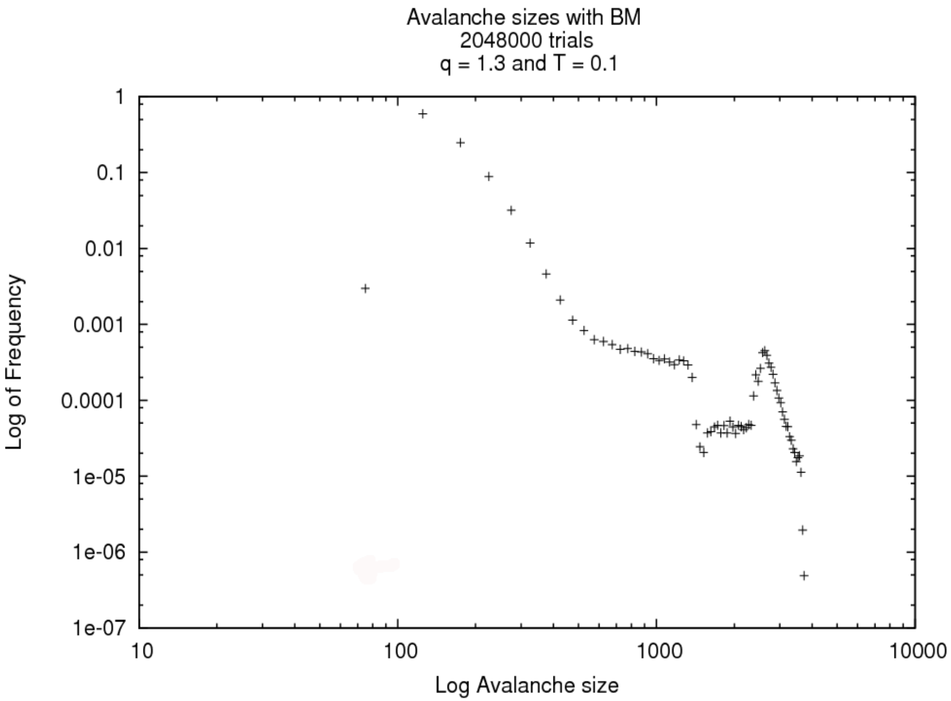
Contact

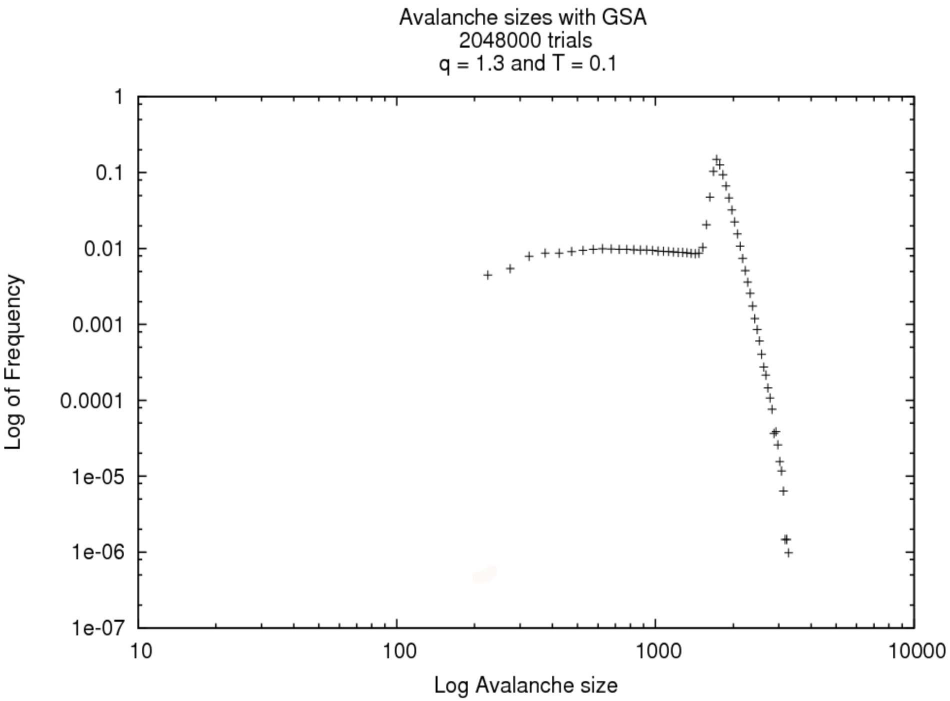
- roseli@ime.uerj.br
- LuisAlfredo@ufrj.br
- donangel@if.ufrj.br

"Once before I ventured to tell you that you nourish a deeply rooted faith in undetermined psychical events and in free will, but that is quite unscientific and must yield to the demand of a determinism whose rule extends over mental life."

Freud, *Introductory Lectures on Psycho-Analysis*

Standard Edition, W. W. Norton and Co, 1966, first German Edition in 1917


Avalanche sizes with BM
2048000 trials
q = 1.3 and T = 0.1
Log of Frequency
Log Avalanche size
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
10
100
1000
10000


Avalanche sizes with GSA
2048000 trials
q = 1.3 and T = 0.1
Log of Frequency
Log Avalanche size
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
10
100
1000
10000

Discussion

Motivations

- In real biological neuronal circuits, interactions among neurons (coupling, weights) are not (necessarily exactly) symmetrical.

$$\omega_{ij} \neq \omega_{ji}$$

- No Hamiltonian?
- What are the necessary and sufficient conditions for having a Hamiltonian?
- Is it possible to define some Lyapunov (Energy Function, Hamiltonian) for these systems, whose optimization corresponds to solving equations of motion for the neuronal system?
- If so, can we relate this to some physical energy of the neuronal system?

- Should these systems require a time-independent Hamiltonian? $\longrightarrow$ Boltzmann statistics ?
- What is a possible Statistical Mechanics treatment for these systems?
- Under what conditions (what physical properties of the system) would we need a generalized statistical mechanics treatment (non-Boltzmann)?
- Advantages of doing stochastic treatment such as simulated annealing instead of just repeating dynamics many times?
- Simulated annealing procedure described earlier is taken from analogy with Ising Model (atoms with binary spins on a lattice).

Definitions

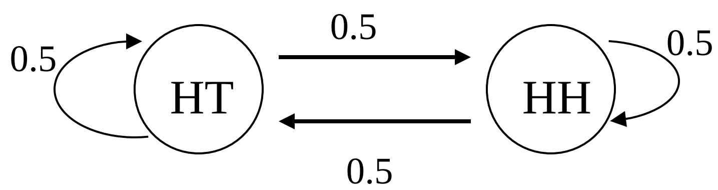
- **Isolated System in Equilibrium:** *"Such an equilibrium situation is characterized by the fact that the probability of finding the system in any one state is independent of time (i.e. the representative ensemble is the same irrespective of time)."*
- **Ergodic System:** *a dynamical system which has the same behavior averaged over time as averaged over the space of all the system's states (phase space).*
- **Ergodic hypothesis:** *says that, over long periods of time, the time spent by a particle in some region of the phase space of microstates with the same energy is proportional to the volume of this region.*

[1]F. Reif, Fundamentals of Statistical and Thermal Physics, McGraw Hill, 1985.

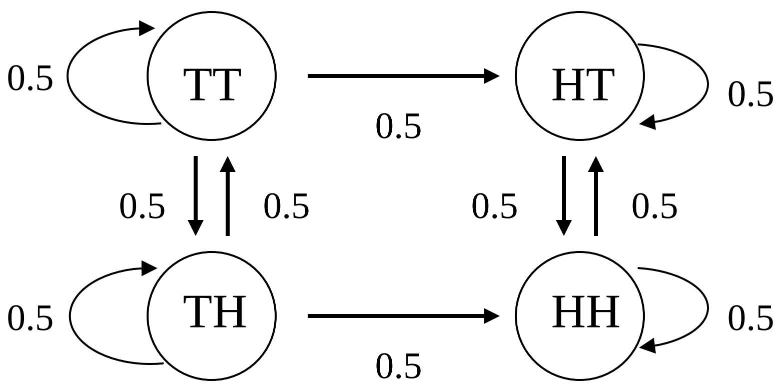
Example

Two coins, tossed one at a time according to the following finite state machines:

one coin is clamped always



ergodic



non-ergodic:

visited states depends on initial state

Assumptions for deriving Boltzmann-Gibbs distribution of occupying states

- Derived for system *A* in contact with a heat reservoir *A'* in [1].

- *A* << *A'*, *A* has many fewer degrees of freedom than *A'*.
  Exs: Bottle of wine imersed in a swimming pool.
  Atom (system) at some lattice site in a solid (the heat reservoir)

- Weak interaction between *A* and *A'* ⟶ Energies are additive.

$$A_{tot} = A + A' \;, \quad E_{tot} = E_A + E_{A'} \;, \quad \text{only} \quad E_{tot} \quad \text{is conserved}$$

- **Under conditions of equilibrium, what is the probability P_r of finding *A* in any one particular state *r* of energy E_r ?**

[1]F. Reif, Fundamentals of Statistical and Thermal Physics, McGraw Hill, 1985.

- If A is in one definite state r, the reservoir must have energy $E_{A'} = E_{tot} - E_r$. The number of states accessible to A_{tot} is just the number of states accessible to A', $\Omega'(E_{tot} - E_r)$, when the energy lies in a range δE near $E_{A'} = E_{tot} - E_r$.

- According to the fundamental statistical postulate, the probability of occurrence in the ensemble of a situation where A is in the state r is proportional to the number of states accessible to A_{tot} under these conditions.

$$P_r = C'\Omega'(E_{tot} - E_r) \quad (1) \qquad \sum_r P_r = 1 \quad (2)$$

Sum is over all r, irrespective of energy.

- $A << A'$, so $E_r << E_{tot}$, so (1) can be approximated, by expanding the slowly varying logarithm of $\Omega'(E_{A'})$ about the value $E_{A'} = E_{tot}$.

$$\ln \Omega'(E_{tot} - E_r) = \ln \Omega'(E_{tot}) - \left[\frac{\partial \ln \Omega'}{\partial E_{A'}}\right]_{E_{tot}} E_r \ldots$$

- Since *A'* acts as a heat reservoir, $E_r <<< E_{tot}$ and higher order terms in the expansion can be neglected. The derivative

$$\left[\frac{\partial \ln \Omega'}{\partial E_{A'}}\right]_{E_{tot}} = \beta$$

is constant and independent of energy E_r of *A*. $\beta = (kT)^{-1}$ characterizes the heat reservoir *A'*. Hence

$$\ln \Omega'(E_{tot} - E_r) = \ln \Omega'(E_{tot}) - \beta E_r$$
$$\Omega'(E_{tot} - E_r) = \Omega'(E_{tot}) e^{-\beta E_r}$$

- Since $\Omega'(E_{tot})$ is a constant (independent of *r*), (1) becomes

$$P_r = Ce^{-\beta E_r} \qquad (3) \quad \text{Canonical Distribution}$$

- C is a constant (independent of *r*). Using normalization condition (2)

$$C = \frac{1}{\sum_r e^{-\beta E_r}}$$

- Thus (3) can be written as

$$P_r = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}} \quad .$$

- What about our neural networks?
- What is the heat bath?

In the Ising model, to apply Boltzmann, one must assume that each atom interacts weakly with other degrees of freedom of the substance.

Example from [1]: Using the analogy from the Ising Model, you have the idea for the Boltzmann machine. Then

A = the network and the brain substance in which the beural network is embedded is the heat bath (*A'*).

However, in our system:

- topology is not a grid structure. There are short and long range in a hierarchical, clustered structure.
- Signal to noise ratio in a synapse may depend on the interaction itself.
- It seems neural networks do not have properties of a Boltzmann-like system.

$$E(\{S_i\}) = -\frac{1}{2}\sum_{ij} \omega_{ij} S_i S_j$$

- If the system does not satisfy Boltzmann's assumptions and total energy is conserved, can we suppose that the probability of it occupying a state depends only on its total energy?
- If so, the system is ergodic (no dissipation of energy).
- Then we would have a system in equilibrium and ergodic, but where the Boltzmann distribution does not apply.

- Is this reasoning ok?

- How could could we derive probability of occupying states?

transference: new editions of old conflicts → **reconfiguration of topology of brain´s neural network, new connections and reinforcement**

→ **new network energy minima** → **new conscious or unconscious acts**

repetition → **network learning** → **long duration of psychoanalytic process**

Hierarquical Clustering Algorithm

1. Generate a uniform distribution of neurons on two 2-dimensional memory sheets.

2. Distribute synapses between nodes n_i and n_j and their respective weights according to gaussian distribution P_G with standard deviation σ. Solution of diffusion equation of neural growth factors, for simulation of synaptic growth.

3. For each neuron n_k corresponding to a center of a region that represents a symbol generated by a previous SOM model, repeat step 4 until clustered.

4. For each n_l adjacent to n_k

$$prob = |\omega_{kl}| \;/\; \Sigma_j |\omega_{kj}|$$

If *random_number* < *prob*, increase $|\omega_{kl}|$ by

$$\Delta = \eta\, prob \qquad\qquad \eta \in (0, 1]$$

and decrease Δ from all neighbors ω_{km} , $m \neq l$ by

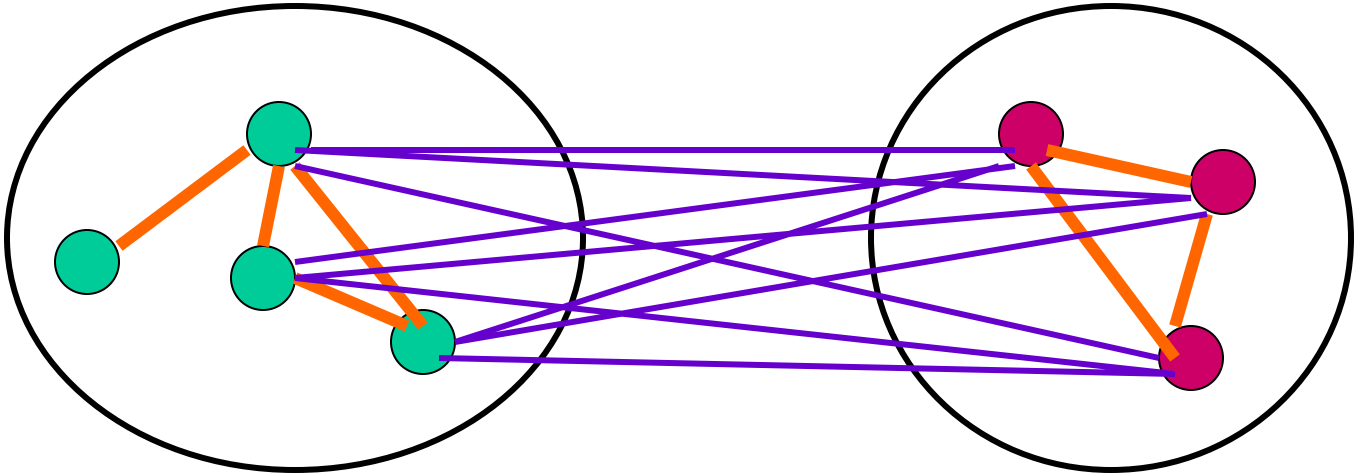
$$\delta_{\mathrm{km}} = (1 - |\omega_{km}| \,/\, (\Sigma_{j \neq l} |\omega_{kj}|)\, \Delta$$

Pattern retrieval on net is achieved by standard simulated annealing (SA) process where T is gradually lowered by a factor α (annealing schedule).

Algorithm Neuroses

1. Initially take random values for ω_{ij}.

2. Divide network into two weakly linked subsets, ω_{ij} between sets multiplied by a number λ in (0, 1].



3. Find initially stored patterns by presenting many random patterns to the Boltzmann machine, with α that allows stabilizing onto local minimum energy states. These are neurotic states since they have weakly linked conscious and unconscious traces.

4. Simulate working-through by stimulating net, changing a randomly chosen n_j in unconscious part of stored pattern P. Present this new pattern P' to network.

5. Response to P' is a pattern P^* with symbolic set state different than symbolic state of P $\longrightarrow$ interpret as new conscious association and enhance all ω_{ij} from n_j to changed nodes in symbolic set.

$$\Delta\omega_{ij} = \beta S_i S_j \omega_{\max}, \qquad \beta \in (0,1] \qquad \textbf{Hebbian}$$

6. Repeat from step 4 for various reinforcement iterations.

7. Repeat steps 4 to 6 for various initial values of T to test associativity with varying noise.

Resulting new set of ω_{ij} will define a new network configuration.

Learning only occurs when stimulus from analyst is not similar to neurotic memory trace.

Simulation Model Illustration
Network Parameters

N	**50**
$N_{Sens} = N_{Symb}$	**25**
initial T for finding global minimum with SA	**1.0**
initial T for seeking local minima with SA	**0.05**
lowest T for SA	**0.001**
α for finding global minimum	**0.95**
α for finding local minima	**0.50**
No. iterations per T in SA	**20**
parameter for initializing inter-cluster synapses λ	**0.5**
size of bi-dimensional memory sheets	**1.9 x 1.9**
standard deviation of Gaussian σ	**0.58**
β	**0.3**
No. of iterations for reinforcement learning	**1000**

Dopaminergic Modulation Hypothesis

- **Neurotransmitters: Act faster through synaptic membrane ion channels. Involved in immediate processing of signals.**

- **Neuromodulators (catecholamines: norepinephrine, epinephrine and dopamine): Produced in brain sites and released in distant and widespread areas. Activate intracelular messengers, promote longer effects, regulate operational characteristics of receptor neurons, i.e. their responses to neurotransmitters.**

Psychoanalytic working-through explores the neurophysiologic potentialities of the subject, including network reconfiguration. When these potentialities are limited by chemical alterations, such as in strong psychotic cases, working-through should be limited or even impossible.