Bifurcating Recursive Nodes Networks and Multi-assemblies Structures as Tools for Intentional Dynamics Modeling

Emilio Del Moral Hernandez
Department of Electronic Systems Engineering

University of São Paulo - Polytechnic School (Brazil)
Outline of the presentation

- **RPEs networks as dynamic systems**
  (RPE - Recursive Processing Element)
- **Associative assemblies with RPE nodes**
- **Multiple assemblies and connectivity aspects**
- **Diversity of features in different RPE nodes**
- **Degradation of associative performance with synaptic connections scale**
- **Relation to incremental learning and number of stored patterns**
Generic RPE nodes … (or bifurcating maps)  
RPEs - Recursive Processing Elements

- Local functionality of nodes (“neurons”) is given by a first order recursion with a numeric parameter “\( p \)”
- The state variable \( x \) assumes continuous values and evolves with the discrete time \( n \) according to:

\[
x_{n+1} = R_p(x_n)
\]

- \( R_p \) is the parametric function that relates \( x_n \) and \( x_{n+1} \)
- \( p \) is a numeric (“bifurcation”) parameter
- \( x_n \) and \( x_{n+1} \) are consecutive values of the state variable \( x \)

- **One specific example**: the Logistic recursion

\[
R_p(x_n) = p \cdot x_n \cdot (1 - x_n)
\]
One-dimensional bifurcating recursive maps

\[ x_{n+1} = R_p(x_n) \]

... A family of recursive maps:
Networks of oscillators / attractor networks / associative memory of coupled RPEs

Value $X_{10}$

Value $X_{20}$

Value $X_{30}$

Value $X_{40}$

$L1$  Sequence $X_{1n}$

$L2$  Sequence $X_{2n}$

$L3$  Sequence $X_{3n}$

$L4$  Sequence $X_{4n}$

$p_1$ to $p_4$
Example of a period-2 attractor: the evolution in time of a node state variable $x$.

From analog $x$ to Digital:
The phase of oscillation between A and B codes a binary digit.
The logistic map, an example of Recursive Processing Element (RPE) Bifurcating Node

\[ x_{n+1} = R_p(x_n) \quad \leftrightarrow \quad x_{n+1} = px_n(1-x_n) \]

Sweeping the bifurcation parameter \( p \) / Sweeping the family of logistic recursions

EPUSP – University of Sao Paulo: wwwlsi.usp.br/~emilio Emilio Del-Moral-Hernandez
Hopfield nets deal with fixed-point attractors
RPEs nets deal with more complex attractors

\[
W = \left( \sum_{\mu=1}^{M} \xi_{\mu} \xi_{\mu}^T \right) - MID
\]

Prototypical applications of associative networks on binary strings: recovery of binary patterns from distorted or partial versions of the stored memories

12% noise          0% noise

EPUSP – University of Sao Paulo:  www.lsi.usp.br/~emilio  Emilio Del-Moral-Hernandez
Experiments with Associative Architectures - measuring pattern recovery power (100 nodes)
Contrasting the RPEs Associative Memory with the Hopfield Neural Networks

AVERAGE ERROR PLOTS FOR RPEs and HOPFILED ARCHITECTURES

HAMMING ERROR IN RECOVERY (%)

NOISE IN PROMPTING (% OF FLIPPED BITS)

HOPFIELD NN

RPEs NN
Search for stored patterns ... each $p_i$ can evolve within the limits $p_{min}$ and $p_{max}$ ...

- $p_{min}$ chosen for period-2 attractors &
- $p_{max}$ chosen for chaotic search
Search for stored patterns / synchronization of all the RPEs, through $p_i$ driven coupling

- range $[p_{\text{min}}, p_{\text{max}}]$ limits the excursions of $p_i$

  $\Delta p_i = \Delta p_{i+1} = p_i + \Delta p_i$ ...

  where ...

  $\Delta p_i = d_i \cdot c - d$

- driving of $p_i$: $p_{i+1} = p_i + \Delta p_i$

- disagreement measure (between $x_i$ and $X_{\text{network}}$):

  $d_i = - (x_i - k1). (\sum_j w_{ij} \cdot (x_j - k1)) + k2$

EPUSP – University of Sao Paulo:  www.lsi.usp.br/~emilio  Emilio Del-Moral-Hernandez
Associative (RPEs) networks: appearance of collective (coordinated) period-2 attractors

LIMIT SET DIAGRAM AND Pi CONFIGURATION

VALUES OF Xi

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

NEURON NUMBER

0 2 4 6 8 10 12 14 16 18 20

ones
zeros

A

B

？

$\pi_i/10$

EPUSP – University of Sao Paulo: www.lsi.usp.br/~emilio Emilio Del-Moral-Hernandez
Complex functionalities require cooperation of more than one associative assembly.

"conversation" between assemblies
“Connecting the pieces” ... Inter-assemblies connections and Time-dependent inputs

Connections Coming from Another Assembly of RPEs or From Time-Dependent Inputs

From Other RPEs Or Inputs Paths

Output Pattern

Connections From $p_1$ to $p_4$

Initial $X_{10}$

Initial $X_{20}$

Initial $X_{30}$

Initial $X_{40}$

Sequence $X_{1n}$

Sequence $X_{2n}$

Sequence $X_{3n}$

Sequence $X_{4n}$

EPUSP – University of Sao Paulo: wwwlsi.usp.br/~emilio Emilio Del-Moral-Hernandez
Inter- assemblies connections ... use of the same mechanisms used for RPEs coupling

\[ di = -(x_i - k1) \sum_j w_{ij} \cdot (x_j - k1) + k2 \]

\[ \sum_j w_{ij} \cdot (x_j - k1) + \sum_k SW_{ik} \cdot (S_k - k1) \]

\[ di = -(x_i - k1) \left[ \sum_j w_{ij} \cdot (x_j - k1) + \sum_k SW_{ik} \cdot (S_k - k1) \right] + k2 \]

**note:** \( SW_{ik} = \sum_{\mu=1:M} (\xi_{i,\mu} \cdot \xi_{S_k,\mu}) \)
Illustrative time-dependent input, assuming the input signal $S_k$ is generated by a logistic node.
A possible coding of analog input patterns

- The phase of cycling of the period-2 attractors codes the sign ($\pm$) of the analog amount.
- The amplitude of cycling $|A-B|$ in the sensory nodes encodes the magnitude … $|A-B| = (1 - 2/p - 3/p^2)^{1/2}$
... In this way, we allow for the “conversation” between assemblies (& sensing of environment)
Beyond Logistic RPEs: other bifurcating recursions can also be used …

- Each recursion has its own properties and features
  - *different computational advantages*
  - *different modeling needs*
- Different types of recursive nodes can be used in different assemblies of a modular architecture
- Some of the differences that appear from recursion to recursion:
  - The range of values visited by the $x$ state variable
  - The amplitude and average value of the oscillations of attractors in each RPE
  - The conditions of limit cycle stability
Beyond Logistic RPEs: other bifurcating recursions can also be used ...

\[ x_{n+1} = p x_n (1 - x_n) \]

\[ x_{n+1} = p x_n^3 - p x_n + x_n \]
Adapting equations to deal with diverse Recursive Elements Nodes

- To deal with different amplitudes of oscillation \( |A_j - B_j| \), in different \( j \) neighboring neurons, a multiplying factor in the \( w_{ij} \) connection:
  \[
  w_{ij} \propto \frac{1}{|A_j - B_j|}
  \]

- Dealing with the average of each particular neighboring variable \( x_j \) in the coupling equations ...
  - In self connections: \( k1_j = (A_j + B_j)/2 \)
  - In inter-assemblies connections: \( k1_k = (A_k + B_k)/2 \)
Defining $x_n$ cycles in period-2 attractors in Logistic nodes ...

- $x_n = R_p(R_p(x_n))$, where $R_p(x) = p.x.(1-x)$
- Values of $x_n$ for a given value of parameter $p$ ...
  
  $A \& B = (1+1/p \pm (1-2/p -3/p^2)^{1/2})/2$

- Condition of stability:
  
  $| R_p'(A) \cdot R_p'(B) | < 1 \Leftrightarrow | p^2 - 2p - 4 | < 1$
  
  (… valid $p$ in the range 3 to 3.45)

- Separation … $|A-B| = (1- 2/p -3/p^2)^{1/2}$

  for other RPEs, changes are needed …
Revisiting $x_n$ cycles period-2 attractors features:
Limit cycle $A$ & $B$, and stability conditions

- $x_n = R_p (R_p (x_n))$, where $R_p (x) = p.x.(1-x)$

- Values of $x_n$ for a given value of parameter $p$ ...
  
  $A$ & $B = \frac{(1+1/p \pm (1-2/p -3/p^2)^{1/2})}{2}$

- Condition of stability:
  
  $| R_p'(A) \cdot R_p'(B) | < 1 \iff | p^2 - 2p - 4 | < 1$
  
  (... valid $p$ in the range 3 to 3.45)

- Separation ...
  $|A-B| = \frac{(1-2/p -3/p^2)^{1/2}}$
Emergence of GENERIC recursive maps with Integrate and Fire relaxation oscillators

Any 1st order recursive map is achievable!!

Theory / Method in IEEE Transactions on C.A.S., Dec 2003 (in the list of references)
Emergence of GENERIC recursive maps with Integrate and Fire relaxation oscillators

\[ V_q = \sum_{p=1}^{2} \text{comparator input} \]

\[ x_n, x_{n+1}, x_{n+2} \]

“2\pi = 1”
The sine-circle map, another example of a Recursive Processing Element (RPE)

\[ x_{n+1} = \left( x_n + \Omega - \frac{K}{2\pi} \sin(2\pi x_n) \right) \mod 1 \]
Some issues that emerge when dealing with several assemblies with diverse features

- The coupling between outputs of an assembly and the input of another assembly
- Different RPE nodes in different assemblies
- Different magnitudes of the $W$ connection matrix, for example due to different number of stored patterns in each assembly
Important also in contexts of continuous learning ... the # of stored patterns ($M$) grows!!
Update of the matrix $W$, for the incremental storage of the new patterns ...

$$W_{new} = W_{old} + (\xi_1 \xi_1^T - \text{ID})$$

$$W = \left( \sum_{\mu=1}^{M} \xi_\mu \xi_\mu^T \right) - M\text{ID}$$

$$= (\xi_1 \xi_1^T - \text{ID}) + (\xi_2 \xi_2^T - \text{ID}) + \cdots + (\xi_M \xi_M^T - \text{ID})$$

$$= W_{old} + (\xi_1 \xi_1^T - \text{ID})$$
Relating the number of stored patterns $M$ with the magnitudes of $w_{ij}$...

- For $w_{ij} = \sum_{\mu=1:M} (\xi_i, \mu \cdot \xi_j, \mu)$, $i \neq j$
- For statistically independent $\xi_i, \mu$:
  
  $$<|w_{ij}|> \propto M^{1/2}$$

- Therefore, two processes contribute for the degradation of performance with larger $M$:
  - The load factor $L=M/N$ of the associative matrix $W$, (this relates to the Capacity)
  - The increase of the average $w_{ij}$ magnitude

The second effect can be avoided!!

EPUSP – University of Sao Paulo: www.lsi.usp.br/~emilio Emilio Del-Moral-Hernandez
Departing from the optimal scale for the matrix $W$ ... Scaling it up (x 3)
The reason for the degradation of performance with the scaling of $W$ matrix ...

- large magnitudes of $w_{ij}$ make the network more unstable

- disagreement measure (between $x_i$ and $X_{network}$):
  $$ d_i = - (x_i-k1) \cdot (\sum_j w_{ij} \cdot (x_j-k1)) + k2 $$

- driving of $p_i$:
  $$ p_{i+n+1} = p_{i+n} + \Delta p_i \ldots \text{where} \ldots $$
  $$ \Delta p_i = d_i \cdot c - d $$
# Hamming error in recovered pattern for different scaling factors for the $W$ matrix

<table>
<thead>
<tr>
<th>Scaling for $W$ matrix →</th>
<th>$\frac{1}{10}$</th>
<th>$\frac{1}{3}$</th>
<th>REFERENCE (factor is 1)</th>
<th>$x \ 3$</th>
<th>$x \ 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hamming Error for Prompting Noise 10%</strong></td>
<td>0.13 %</td>
<td>0.00 %</td>
<td>0.00 %</td>
<td>0.00 %</td>
<td>0.04 %</td>
</tr>
<tr>
<td><strong>Hamming Error for Prompting Noise 20%</strong></td>
<td>1.32 %</td>
<td>0.28 %</td>
<td>0.11 %</td>
<td>0.26 %</td>
<td>2.23 %</td>
</tr>
</tbody>
</table>

*(average amounts for 500 experiments)*

**EPUSP – University of Sao Paulo:**  www.lsi.usp.br/~emilio  Emilio Del-Moral-Hernandez
Degradation of the basin of attraction
(from the extended paper in the NNs Special Issue)

<table>
<thead>
<tr>
<th>Scaling for W matrix →</th>
<th>$\frac{1}{10}$</th>
<th>$\frac{1}{3}$</th>
<th>REFERENCE (factor 1)</th>
<th>3</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Radius of Attraction (bits)</td>
<td>13</td>
<td>20</td>
<td>24</td>
<td>23</td>
<td>19</td>
</tr>
<tr>
<td>Radius for Guaranteed Recovery (bits)</td>
<td>2</td>
<td>5</td>
<td>12</td>
<td>10</td>
<td>6</td>
</tr>
</tbody>
</table>

(average for 500 experiments - network size is 100 nodes)
<table>
<thead>
<tr>
<th># of Patterns</th>
<th>M=3 (reference)</th>
<th>M=4</th>
<th>M=5</th>
<th>M=6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompting Noise 5%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.38%</td>
<td>1.60%</td>
</tr>
<tr>
<td>Prompting Noise 10%</td>
<td>0.00%</td>
<td>0.05%</td>
<td>1.58%</td>
<td>2.22%</td>
</tr>
<tr>
<td>Prompting Noise 15%</td>
<td>0.04%</td>
<td>0.55%</td>
<td>2.02%</td>
<td>3.05%</td>
</tr>
</tbody>
</table>
New results, with normalization of $W$ scale i.e., $<|w_{ij}|> \propto 1/M^{1/2}$ (second value in cell, in black)

<table>
<thead>
<tr>
<th># of Patterns</th>
<th>M=3 (reference)</th>
<th>M=4</th>
<th>M=5</th>
<th>M=6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompting Noise 5%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.38%</td>
<td>1.60%</td>
</tr>
<tr>
<td>Prompting Noise 10%</td>
<td>0.00%</td>
<td>0.05%</td>
<td>0.30%</td>
<td>1.08%</td>
</tr>
<tr>
<td>Prompting Noise 15%</td>
<td>0.04%</td>
<td>0.55%</td>
<td>1.58%</td>
<td>2.22%</td>
</tr>
</tbody>
</table>
Final Comments / Summary of ingredients of the RPEs architectures / Potential

- The attractors which emerge at the output are spatiotemporal patterns, not static attractors
  \emph{(production of multidimensional & spatiotemporal patterns)}

- Modularity, heteroassociation, and heterogeneous multi-assemblies architectures
  \emph{(modeling / implementation of complex structures / functions)}

- Time-dependent inputs
  \emph{(sensing of changing environment)}

- Arbitrary recursive nodes easily explored
  \emph{(modeling different dynamical phenomena)}

- Many other periodic attractors are still there!!
  \emph{(... period – 4, period – 8 and etc ... )}
My coordinates … & Acknowledgments

Emilio Del-Moral-Hernandez
Polytechnic School - University of São Paulo (EPUSP)
Department of Electronic Systems Engineering
emilio_del_moral@ieee.org
wwwlsi.usp.br/~emilio
Some sources for additional information


