Recurrent Competitive Fields

Lecture 9

Instructor: Anatoli Gorchetchnikov <anatoli@bu.edu>
Shunting Competitive Networks

Properties of feedforward case

- saturation avoidance
- noise suppression (uniform inputs)
- factorization of pattern and energy
- automatic gain control
- tendency toward total activity conservation
  (normalization)
- shift property in log coordinates
- Weber’s law
Recurrent Competitive Fields

Recurrent simply refers to the inclusion of feedback pathways that allow a cell’s output to project back to it’s input either directly or indirectly.

Competitive refers to the fact that there are inhibitory interactions between cells in the network.
  – On-center, off-surround describes the nature of these inhibitory interactions.

Field refers simply to a group of interconnected neurons.
  – The term network could also be used.
Recurrent Competitive Fields

Furthermore, we will primarily be interested in shunting RCFs

Alternative names for the networks we’ll look at include:
- recurrent on-center, off-surround shunting competitive network
- shunting competitive network with feedback
- etc
Deriving a Shunting RCF

Start with feedforward network:

\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - x_i \sum_{j \neq i} I_j
\]

What happens if we allow for inputs from other cells (i.e., recurrent connections)?

Furthermore, what if the on-center, off-surround nature of the network was built into these feedback connections rather than the inputs to the cells?
Deriving a Shunting RCF

\[
\frac{dx_i}{dt} = -Ax_i + \left(B - x_i\right) I_i - x_i \sum_{j \neq i} I_j
\]

Replace inhibitory external inputs with recurrent inhibition

Add recurrent excitation

\[
\frac{dx_i}{dt} = -Ax_i + \left(B - x_i\right) \left(I_i + f(x_i)\right) - x_i \sum_{j \neq i} f(x_j)
\]

All recurrent inputs go through a signal function \(f(x)\)

In this case it can also be called a feedback function
Recurrent Competitive Field

\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i)(I_i + f(x_i)) - x_i \sum_{j \neq i} f(x_j)
\]

Recurrency can allow the activity to reverberate
The system can be used as a persistent short term memory
But only if it maintains the initially imprinted pattern without distortions

– Will this pattern be maintained?
– Will it distort?
– If so how will it distort?

It all depends on the shape of the signal function \( f(x) \)
Method

\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i)(I_i + f(x_i)) - x_i \sum_{j \neq i} f(x_j)
\]

Assume an initial pattern has been impressed on the field and the inputs have been turned off.

Then the equation reduces to

\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i)f(x_i) - x_i \sum_{j \neq i} f(x_j)
\]

We want to see what happens over time to the pattern stored in memory and to the total activity in the network.

We will try to get an intuitive feel for what is happening rather than doing hard-core mathematical analyses.
Possibilities

Let’s say a pattern is provided to the network. After the input is turned off, what would be the activity the network will settle to?

- Perfect storage
- Degradation
- Only the strongest persists
- Strong signals persist
- All signals saturate at some value

\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i)f(x_i) - x_i \sum_{j \neq i} f(x_j)
\]
Look at $f(x)$ and $g(x) = f(x)/x$

<table>
<thead>
<tr>
<th>Linear</th>
<th>$f(x_i) = Cx_i$</th>
<th>$g(x_i) = C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slower than linear</td>
<td>$f(x_i) = \frac{Cx_i}{D + x_i}$</td>
<td>$g(x_i) = \frac{C}{D + x_i}$</td>
</tr>
<tr>
<td>Faster than linear</td>
<td>$f(x_i) = Cx_i^2$</td>
<td>$g(x_i) = Cx_i$</td>
</tr>
<tr>
<td>Sigmoidal</td>
<td>$f(x_i) = \frac{Cx_i^2}{D + x_i^2}$</td>
<td>$g(x_i) = \frac{Cx_i}{D + x_i^2}$</td>
</tr>
</tbody>
</table>

CN 510 Lecture 9
Total Energy Equation

\[ \dot{x}_i = -Ax_i + (B - x_i) f(x_i) - x_i \sum_{j \neq i} F(x_j) \]

Sum up for all \( x_i \)

\[ \sum_{i=1}^{n} \dot{x}_i = -A \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} (B - x_i) f(x_i) - \sum_{i=1}^{n} x_i \sum_{j \neq i} F(x_j) \]

Recombine the terms

\[ \dot{x} = -Ax + B \sum_{i=1}^{n} f(x_i) - \sum_{i=1}^{n} x_i f(x_i) - \sum_{i=1}^{n} x_i \sum_{j \neq i} f(x_j) \]

Join terms with the sum of \( f(x) \)

\[ \dot{x} = -Ax + B \sum_{i=1}^{n} f(x_i) - \sum_{i=1}^{n} x_i \sum_{j=1}^{n} f(x_j) \]

Pull the sum out

\[ \dot{x} = -Ax + (B - x) \sum_{i=1}^{n} f(x_i) \]

CN 510 Lecture 9
Total Energy Equation

For each signal function we would like to know what would happen to the total energy in the system.

Given non-negative signal functions the equation

\[
\dot{x} = -Ax + (B - x) \sum_{i=1}^{n} f(x_i)
\]

ensures that the total energy of the system will stay bounded between 0 and \( B \).

In order to analyze what value it will converge to it is useful to rewrite the equation.
Total Energy Equation

\[ \dot{x} = -Ax + (B - x) \sum_{i=1}^{n} f(x_i) \]

Pulling \( x(B-x) \) out

\[ \dot{x} = x(B - x) \left( \sum_{i=1}^{n} \frac{f(x_i)}{x} - \frac{A}{B - x} \right) \]

And substituting \( f(x_i) = g(x_i)x_i \)

\[ \dot{x} = x(B - x) \left( \sum_{i=1}^{n} \frac{x_i}{x} g(x_i) - \frac{A}{B - x} \right) \]

CN 510 Lecture 9
Total Energy Equation

\[ \dot{x} = x(B - x) \left( \sum_{i=1}^{n} \frac{x_i}{x} g(x_i) - \frac{A}{B - x} \right) \]

Now using our usual pattern variable \( X_i = \frac{x_i}{x} = \frac{x_i}{\sum_{k=1}^{n} x_k} \)

\[ \dot{x} = x(B - x) \left( \sum_{i=1}^{n} X_i g(x_i) - \frac{A}{B - x} \right) \]

And finally assigning

\[ G = \sum_{i=1}^{n} X_i g(x_i) \quad \text{we end up with} \quad \dot{x} = x(B - x) \left( G - \frac{A}{B - x} \right) \]

CN 510 Lecture 9
\[ \dot{x} = x(B - x) \left( G - \frac{A}{B - x} \right) \]

has the following critical points:

\( x=0 \), and a set of points \( G = \frac{A}{B - x} \) or \( x = B - \frac{A}{G} \)

The case \( x=B \) turns the denominator of \( \frac{A}{B - x} \) to 0, so it is not a valid critical point.

Furthermore, when \( 0 < x < B \), the sign of the derivative is the same as the sign of

\[ G - \frac{A}{B - x} \]
We will be looking at the relationship between \( \frac{A}{B - x} \) and 

\[
G = \sum_{i=1}^{n} X_i g(x_i)
\]

Since the range of \( G \) is the same as the range of \( g(x_i) \)

Simplifying approximation:
Assume that the shape of \( G \) is the same as the shape of \( g(x) \)

Basically replacing a sum of functions with a function of a sum

\[
G = \frac{1}{x} \sum_{i=1}^{n} f(x_i) \rightarrow \frac{1}{x} f \left( \sum_{i=1}^{n} x_i \right) = \frac{f(x)}{x} = g(x)
\]
This approximation is not valid for quantitative analysis
For one strong $x_i$

$$G = \sum_{i=1}^{n} X_i g(x_i) = g(x_i) = g(x)$$

For uniform $x_i$’s

$$G = \sum_{i=1}^{n} X_i g(x_i) = g(x_i) \neq g(x)$$

However, in both cases the shape of $G$ is the same as the shape of $g()$
Pattern Variables Equation

\[ X_i = \frac{x_i}{\sum_{k=1}^{n} x_k} = \frac{x_i}{x} \]

\[
\dot{x} = -Ax + \left( B - x \right) \sum_{i=1}^{n} f(x_i)
\]

\[
\dot{x}_i = -Ax_i + \left( B - x_i \right) f(x_i) - x_i \sum_{j \neq i} f(x_j)
\]

Using derivative formula

\[
\frac{dX_i}{dt} \frac{d}{dt} \left( \frac{x_i}{x} \right) = \frac{1}{x^2} \left( x \dot{x}_i - x_i \dot{x} \right) = \frac{1}{x} \left( \dot{x}_i - \frac{x_i}{x} \dot{x} \right)
\]

And substituting

\[
\dot{X}_i = \frac{1}{x} \left( -Ax_i + Bf(x_i) - x_i \sum_{i=1}^{n} f(x_i) + Ax_i - \frac{Bx_i}{x} \sum_{i=1}^{n} f(x_i) + x_i \sum_{i=1}^{n} f(x_i) \right)
\]
Pattern Variables Equation

\[ \dot{X}_i = \frac{B}{x} \left( f(x_i) - \frac{x_i}{x} \sum_{k=1}^{n} f(x_k) \right) \]

\[ \dot{X}_i = \frac{Bx_i}{x} \left( g(x_i) - \sum_{k=1}^{n} \frac{x_k}{x} g(x_k) \right) \]

Since the pattern variables are normalized

\[ \dot{X}_i = BX_i \left( g(x_i) \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} X_k g(x_k) \right) \]

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( g(x_i) - g(x_k) \right) \right) \]
Linear Signal Function

\[ f(x_i) = C x_i \]

Firing rate directly proportional to activation

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( g(x_i) - g(x_k) \right) \right) \]

Here the derivative is 0

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( C - C \right) \right) = 0 \]

Thus the pattern does not change
Linear Signal Function

Let’s compute the value of $G$

\[ G = \sum_{i=1}^{n} X_i g(x_i) = \frac{1}{x} \sum_{i=1}^{n} f(x_i) = \]

\[ = \frac{1}{x} \sum_{i=1}^{n} C x_i = \frac{C}{x} \sum_{i=1}^{n} x_i = C \]

And compare \( \frac{A}{B-x} \) and \( G = \sum_{i=1}^{n} X_i g(x_i) = C \)
Linear Signal Function

Case 1:
Single intersection if \( G = C > A/B \)

On the left \( G \frac{A}{B-x} \) is positive, so \( x \) increases
On the right \( G \frac{A}{B-x} \) is negative, so \( x \) decreases

The resulting critical point \( x = B - \frac{A}{C} \) is stable.

\[ \dot{x} = x(B-x) \left( G - \frac{A}{B-x} \right) \]
Linear Signal Function

Thus no matter what the initial (non-zero) energy is it will settle to a fixed value:

– total activity normalization
– amplification of the weak patterns
– reduction of energy for the strong patterns

Thus this case is a perfect pattern preservation with a stable energy level – ideal working memory

Unfortunately, linear signal function is not exactly realistic
Linear Signal Function

Case 2:

If $C \leq A/B$ there is no intersection

The derivative is always negative
The system converges to 0:
pattern degradation

$$\dot{x} = x(B - x) \left(G - \frac{A}{B - x}\right)$$
Linear Signal Function

Summary:

\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i) f(x_i) - x_i \sum_{j \neq i} f(x_j)
\]

\[
f(x_i) = Cx_i
\]

Pattern storage depends on parameter choices:
If \( C \leq A/B \) all \( x_i \) go to 0; no storage
If \( C > A/B \) – perfect storage of the pattern with total energy converging to

\[
x = B - \frac{A}{C}
\]
Slower-than-Linear Signal Function

\[ f(x_i) = \frac{C x_i}{D + x_i} \]

Firing rate tend to increase slower than activation, Type I excitability

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( g(x_i) - g(x_k) \right) \right) \]

Here

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( \frac{C}{D + x_i} - \frac{C}{D + x_k} \right) \right) \]

CN 510 Lecture 9
Slower-than-Linear Signal Function

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( \frac{CD + Cx_k - CD - Cx_i}{(D + x_i)(D + x_k)} \right) \right) \]

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( \frac{C(x_k - x_i)}{(D + x_i)(D + x_k)} \right) \right) \]

Here if \( x_k > x_i \) the sum component is positive, since for the \( X_i \) that are small the derivative has most of \( x_k - x_i \) positive and \( X_i \) will grow.

For the \( X_i \) that are large the derivative has most of \( x_k - x_i \) negative and \( X_i \) will decrease.

This will lead to uniformization of the pattern.
Slower-than-Linear Signal Function

Let’s compute the value of $G$

\[ G = \sum_{i=1}^{n} X_i g(x_i) = \frac{1}{x} \sum_{i=1}^{n} f(x_i) = \]

\[ = \frac{1}{x} \sum_{i=1}^{n} \frac{C x_i}{D + x_i} = \frac{C}{x} \sum_{i=1}^{n} \frac{x_i}{D + x_i} \]

And compare \( \frac{A}{B - x} \) and \( G \)

Note that if \( x=0 \) then \( G = \frac{C}{D} \)
Slower-than-Linear Signal Function

Case 1:
Single intersection if $G_0 = C/D > A/B$

On the left $G - \frac{A}{B - x}$ is positive, so $x$ increases
On the right $G - \frac{A}{B - x}$ is negative, so $x$ decreases

The resulting critical point is stable
Slower-than-Linear Signal Function

Case 2:

If \( C/D \leq A/B \) there is no intersection

The derivative is always negative

The system converges to 0:

pattern degradation
Slower-than-Linear Signal Function

Summary:
\[
\frac{dx_i}{dt} = -Ax_i + (B - x_i) f(x_i) - x_i \sum_{j \neq i} f(x_j)
\]

\[
f(x_i) = \frac{Cx_i}{D + x_i}
\]

Pattern storage depends on parameter choices:
If \( C/D \leq A/B \) all \( x_i \) go to 0; no storage
If \( C/D > A/B \) – then total energy is converging to a stable point, but the pattern is uniformized in the process

Note that although it is the case for Type I excitability to have slower than linear signal function, here \( f(x) \) lumps together synaptic integration and signal function

CN 510 Lecture 9
Faster-than-Linear Signal Function

\[ f(x_i) = Cx_i^2 \]

\[ g(x_i) = Cx_i \]

Firing rate tend to increase faster than activation, not too realistic, but again synaptic integration can kick in...

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( g(x_i) - g(x_k) \right) \right) \]

Here

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k C(x_i - x_k) \right) \]
Faster-than-Linear Signal Function

\[ \dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k C(x_i - x_k) \right) \]

Here if \( x_k > x_i \) the sum component is negative, since for the \( X_i \) that are small the derivative has most of \( x_i - x_k \) negative and \( X_i \) will decrease.

For the \( X_i \) that are large the derivative has most of \( x_i - x_k \) positive and \( X_i \) will grow.

This will lead to contrast enhancement of the pattern and eventually to a winner-take-all situation.
Faster-than-Linear Signal Function

Let’s compute the value of $G$

\[ G = \sum_{i=1}^{n} X_i g(x_i) = \frac{1}{x} \sum_{i=1}^{n} f(x_i) = \]

\[ = \frac{1}{x} \sum_{i=1}^{n} C x_i^2 = \frac{C}{x} \sum_{i=1}^{n} x_i^2 \]

And compare $\frac{A}{B-x}$ and $G$

Note that if $x=0$ then $G = 0$
Faster-than-Linear Signal Function

\[ \dot{x} = x(B - x) \left( G - \frac{A}{B - x} \right) \]

Case 1:
Two intersections if \( G \) is high enough

On the left and right \( G - \frac{A}{B - x} \) is negative, so \( x \) decreases
In the middle \( G - \frac{A}{B - x} \) is positive, so \( x \) increases

The resulting critical point on the right is stable, on the left is unstable
Faster-than-Linear Signal Function

Case 2:
One intersection if $G$ is just touching $\frac{A}{B - x}$

On the left and right $G - \frac{A}{B - x}$ is negative, so $x$ decreases

The resulting critical point is a saddle, depending on the initial condition the system can converge to this point or to 0
Faster-than-Linear Signal Function

Case 3:
No intersections if $G$ is under $\frac{A}{B - x}$

Here $G - \frac{A}{B - x}$ is always negative, so $x$ decreases

Energy in this system will decay to 0 no matter the initial conditions

\[ \dot{x} = x(B - x) \left( G - \frac{A}{B - x} \right) \]
Faster-than-Linear Signal Function

Summary:

\[ \frac{dx_i}{dt} = -Ax_i + (B - x_i) f(x_i) - x_i \sum_{j \neq i} f(x_j) \]

\[ f(x_i) = Cx_i^2 \]

Pattern storage depends on parameter choices:

If \( C \) is not large enough, then all \( x_i \) go to 0; no storage

If \( C \) is large enough – then small initial patterns will decay to 0, while for large energy patterns the total energy will converge to a stable point, but the pattern will be transformed to a winner-take-all activation

This is a useful behavior for categorization tasks, but not very realistic of a signal function
Grossberg suggested that this function combines all three previous cases:

- Contrast-enhancing small signals
- Stores intermediate signals with little distortion
- Uniformizes large signals

Unfortunately, this requires a true linear segment in the middle.

CN 510 Lecture 9
Quenching Threshold

Contrast-enhancement of small signal basically selects the strongest of small input components, amplifies them and suppresses the weaker components. As the stronger components overgrow the faster-than-linear segment they survive, while the weaker components get quenched to 0.

The quenching threshold is a measure of how large the input component has to be in order to survive the faster-than-linear part. The exact value of the threshold depends on the parameters of a sigmoid as well as the input distribution. In the real sigmoid it is the inflexion point that the signal has to cross.
Quenching Threshold

Since the closer the signals to the linear part or inflexion point, the better chance of survival they have, one way to ensure they persist is to add a non-specific activation (arousal or modulation) that will bump up the values of all inputs.

Trough the manipulation of modulation level one can control the noise-suppression properties of a sigmoid.
Sigmoidal Signal Function

$$f(x_i) = \frac{Cx_i^2}{D + x_i^2}$$

$$g(x_i) = \frac{Cx_i}{D + x_i^2}$$

Firing rate tend to increase faster than activation for low activations, saturating for high activations

$$\dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k \left( g(x_i) - g(x_k) \right) \right)$$

Here

$$\dot{X}_i = BX_i \left( \sum_{k=1}^{n} X_k C \left( \frac{x_i}{D + x_i^2} - \frac{x_k}{D + x_k^2} \right) \right)$$
Sigmoidal Signal Function

Case 1:
Two intersections if $G$ is high enough

In the middle $G - \frac{A}{B - x}$ is positive, so $x$ increases
On the left and right right $G - \frac{A}{B - x}$ is negative, so $x$ decreases

The resulting critical point on the right is stable, the one on the left is unstable
Sigmoidal Signal Function

Cases 2 (single point intersection) and 3 (no intersection) are also similar to respective faster-than-linear cases

Summary:

\[
\frac{dx_i}{dt} = -Ax_i + \left( B - x_i \right) f(x_i) - x_i \sum_{j \neq i} f(x_j)
\]

\[
f(x_i) = \frac{C x_i^2}{D + x_i^2}
\]

Pattern storage depends on parameter choices:
For certain parameter choices all \( x_i \) go to 0; no storage
For better choices small initial patterns will decay to 0, while for large energy patterns the total energy will converge to a stable point, but the pattern will be deformed with large activities saturated and small activities going to 0
General RCF Summary
More on RCF

The most interesting cases can arise when recurrent inhibition is distance dependent.

In real spiking neurons the signal function is more complex than any of the four studied here.

If you want to use RCF properties with spiking neurons you should study the signal function of your neurons.

Another aspect to consider is temporal dynamics:

- True sigmoid will eventually uniformize part the pattern and kill the rest.
- In the spiking case the stronger inputs have an advantage of spiking faster, thus preventing other components from spiking at all, few winners take all situation…
Next Time

Synaptic dynamics as a biophysical basis of long term memory changes