Memory

Declarative vs. non-declarative memory

- **Declarative**
  - “Winter break ended on January 15”
  - “Apples are edible, chairs are not edible”

- **Non-declarative**
  - Throwing a baseball
  - Pattern completion (seeing the dog behind the fence)

Types of memory

- **Declarative**
  - Episodic
  - Semantic

- **Non-declarative**
  - Procedural
  - Conditioning
  - Perceptual
  - Reflex

Hippocampus (MTL)
Cerebral cortex
Basal ganglia
Motor cortex
Cerebellum
Amygdala
Cerebral cortex

Short-term vs. long-term memory

- **Short-term memory** – aka “working” memory
  - Hold facts in memory for 1-200 seconds
  - Sometimes prolonged version of perception
  - Associated with prefrontal cortex (PFC)

- **Long-term memory**
  - Stores facts over years
  - Associated with hippocampus (also, amygdala)
Modeling limits of working memory

• How much can we hold in working memory?
  – 7±2 things
  – Things can be simple: A Q R L G
  – Things can be complex

• Why is our working memory limited?
  – Binding hypothesis: distributed code with synchronous spiking – errors with spurious synchronization

Binding hypothesis

Neurons firing at “same time” represent same thing

<table>
<thead>
<tr>
<th>Pink</th>
<th>1 pink cat</th>
<th>1 pink dog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>1 blue dog</td>
<td></td>
</tr>
<tr>
<td>Cat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cow</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Objects too close!

If spikes occurring within 1 ms of each other are considered synchronous, hard to incorporate increasing number of spikes in fixed time

More features not increase risk of spurious synchronization

Small cat confused with big sad cow
Note adding more features (with more neurons!) to a concept/object does **not** cause a problem – no risk of extra overlap in time with more features.

**Neural dynamics in “cortical sheet”**

- Cortical sheet: group of neurons on same level of hierarchy interacting with lateral connections
- Balance between local cooperation and local inhibition

- \( r^{\text{out}} \) determined from:
  \[
  h = \left( \sum_{j} w_{ij} r^{\text{feedfwd}} \right) + \left( \sum_{k} w_{kr} r^{\text{lateral}} \right) + \left( \sum_{m} w_{mr} r^{\text{feedback}} \right)
  \]

In V1, get feedfwd input from “eyes” (actually thalamus)
Get input from other V1 neurons (lat); get input from V2 (fbkd)

**Neural memory in dlPFC for delayed-action task**

- a: stimulus display onset
- b: stimulus display offset
- c: performance of action

Funahashi et al. 1989

**Neural dynamics in action**

- V1/IT
- Color code:
  - Dark red: 1
  - Light red: 0.5
  - Dark blue: -0.4
  - Light blue: -0.1

- Neurons fire with \( r^{\text{out}} = h \) linear
- Side neurons fire at \( r=0.5 \)
- Center neuron fires at \( r=1 \)

At t=1: bananas start pushing n1,n2,n3 to fire
At t=2: banana push n1,n2,n3 up; n2 pull down n1&n3
Neural dynamics: equations and numbers

\[ r_A^{t+1} = w_{in,A}^{t+1} + w_{B,A}^{t+1} \]
\[ r_B^{t+1} = w_{in,B}^{t+1} + w_{A,B}^{t+1} + w_{C,B}^{t+1} \]
\[ r_C^{t+1} = w_{in,C}^{t+1} + w_{B,C}^{t+1} \]

\[ w_{B,A} = -0.4 \quad w_{B,C} = -0.4 \quad w_{A,B} = -0.1 \quad w_{C,B} = -0.1 \]
\[ w_{in,A} = 0.5 \quad w_{in,B} = 1 \quad w_{in,C} = 0.5 \]

<table>
<thead>
<tr>
<th></th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>t=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>??</td>
<td>??</td>
<td>??</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>??</td>
<td>??</td>
<td>??</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>??</td>
<td>??</td>
<td>??</td>
</tr>
</tbody>
</table>

(Feedforward)

Neural dynamics in action

V1/IT

Alternate area

<table>
<thead>
<tr>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>t=4</th>
<th>t=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

in 0 0 0 0

Neural dynamics, alternate area: equations and numbers

\[ r_B^{t+1} = w_{in,B}^{t+1} + w_{A,B}^{t+1} + w_{C,B}^{t+1} \]
\[ r_C^{t+1} = w_{in,C}^{t+1} + w_{B,C}^{t+1} \]

\[ w_B,A = 0.5 \quad w_B,C = 0.5 \quad w_A,B = 0.1 \quad w_C,B = 0.1 \]
\[ w_{in,A} = 1 \quad w_{in,B} = 1 \quad w_{in,C} = 1 \]

<table>
<thead>
<tr>
<th></th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>t=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

in 1 1 0 0
Neural dynamics, alternate area: equations and numbers

\[ w_{B,A} = 0.5 \quad w_{B,C} = 0.5 \quad w_{A,B} = 0.1 \quad w_{C,B} = 0.1 \]
\[ w_{in,A} = 1 \quad w_{in,B} = 1 \quad w_{in,C} = 1 \]

By changing our weights, control speed of reset-to-zero (or prevent reset to zero)

<table>
<thead>
<tr>
<th></th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>t=4</th>
<th>t=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>0.6</td>
<td>0.15</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1</td>
<td>1.2</td>
<td>0.3</td>
<td>0.12</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>0.6</td>
<td>0.15</td>
</tr>
<tr>
<td>in</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Neural dynamics, alternate area: equations and numbers

\[ w_{B,A} = 2 \quad w_{B,C} = 2 \quad w_{A,B} = 1 \quad w_{C,B} = 1 \]
\[ w_{in,A} = 1 \quad w_{in,B} = 1 \quad w_{in,C} = 1 \]

Over-weighting > epilepsy

<table>
<thead>
<tr>
<th></th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>t=4</th>
<th>t=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>in</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Neural system dynamics

- In an interconnected cortical sheet, neural activity can continue after feedforward input is gone
Neural dynamics in action

**Feedback input** sending message: “keep in short-term memory”

Frontal cortex – executive decision making – provides feedback input

Neuron activated
Neuron exciting
Neuron inhibiting

t=1  t=2  t=3  t=4  t=5

Additional color code:
Dark green: .3

Frontal cortex – executive decision making – provides feedback input

**Neural dynamics + memory**

\[
\begin{align*}
& w_{B, A} = .5 & w_{B, C} = .5 & w_{A, B} = .5 & w_{C, B} = .5 \\
& w_{in, A} = 1 & w_{in, B} = 1 & w_{in, C} = 1 \\
& w_{mem, A} = .3 & w_{mem, B} = .3 & w_{mem, C} = .3
\end{align*}
\]

<table>
<thead>
<tr>
<th>t</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>in</th>
<th>mem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1.3</td>
<td>1.3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1.3</td>
<td>2.6</td>
<td>1.9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.3</td>
<td>1.95</td>
<td>1.9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1.95</td>
<td>1.6</td>
<td>.95</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1.43</td>
<td>1.43</td>
<td>.95</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1.43</td>
<td>1.43</td>
<td>.95</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Remove feedforward: A, B, C rates decrease a bit
Remove top-down mem: A, B, C rates decrease faster

**Neural dynamics + memory**

\[
\begin{align*}
& w_{B, A} = .5 & w_{B, C} = .5 & w_{A, B} = .5 & w_{C, B} = .5 \\
& w_{in, A} = 0 & w_{in, B} = 1 & w_{in, C} = 0 \\
& w_{mem, A} = .5 & w_{mem, B} = 0 & w_{mem, C} = .5
\end{align*}
\]

<table>
<thead>
<tr>
<th>t</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>in</th>
<th>mem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Remove feedforward: A, B, C rates decrease a bit
Remove top-down mem: A, B, C rates decrease faster
Neural system dynamics

Trappenberg 7.3.2

- **Decaying activity:** mutual inhibition suppresses continued neural activity after feedforward input is gone – V1
- **Growing activity:** mutual excitation produces global, non-stop activity over time – epilepsy
- **Memory activity:** balance of mutual excitation (and mutual inhibition) produces maintained activity (sparse) distributed coding during “working memory” time period – PFC
Anatomy of long term memory

Hippocampus ("sea horse")
In medial temporal lobe (MTL)
• Input: Entorhinal cortex – EC
• Dentate gyrus – DG
• Cornus ammonis – CA1, CA3
• Perforant pathway: EC -> CA3

Recurrent networks

• Extensive collateral connections in CA3 enhance associative memory

Recurrent networks

• Extensive collateral connections in CA3 enhance associative memory
• Broader loop:
  EC -> CA3 -> CA1 -> EC

\[ \Delta w_{ij} = r_i r_j - r_i w_{ij} \]
Cells that fire together, wire together
Loop repeatedly increases weight — increasingly encourage simultaneous firing

Learning/remembering

• Learning: neurogenesis in DG
• Retrieval: pattern completion in CA3
• Alternate between learning and retrieval phases
  – DG granule cells enable learning
  – Perforant pathway probes memory
Learning locations

- Rats learn neural representations of locations within a maze
- Hippocampal place cells in CA1, CA3

Further hippocampal representations

Grid cells
- In dorsocaudal medial EC
- Represent multiple locations

Samsonowich, J Neurosci 1997
Neurons organized in 2D based on similarity of tuning curves

(C) Some rights reserved, Torkel Hafting