SAND2016-1944C

A Combinatorial Model of Dentate Gyrus Sparse Coding and Pattern Separation

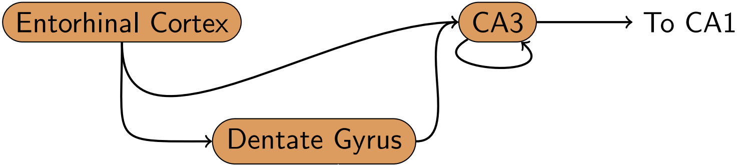
William Severa, Ojas Parekh, Conrad James, James B. Aimone

Sandia National Laboratories, Albuquerque, New Mexico, USA

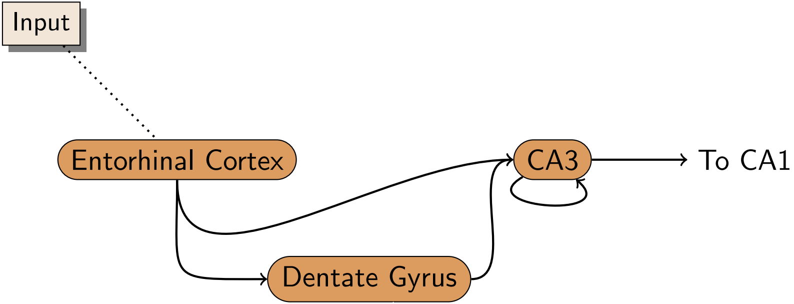
March 9, 2016

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

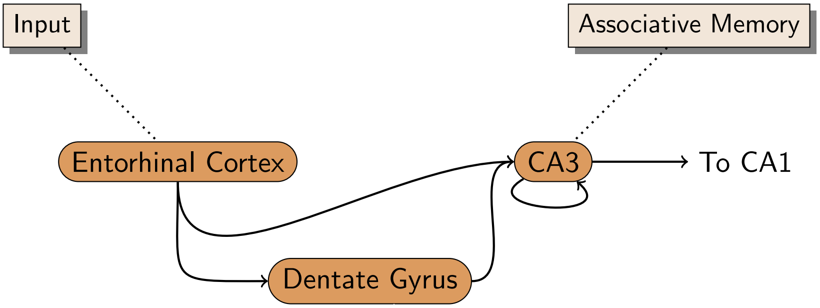
Dentate Sparse Coding Basics



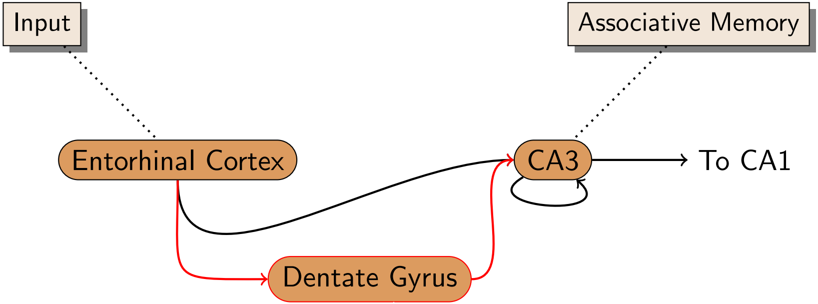
Dentate Sparse Coding Basics



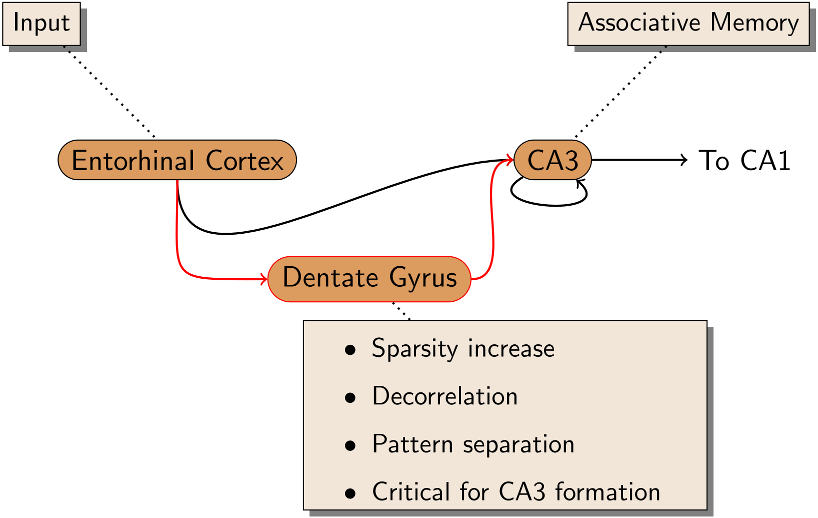
Dentate Sparse Coding Basics



Dentate Sparse Coding Basics



Dentate Sparse Coding Basics



Goals

- $F = fA : \{0,1\}^n \to \{0,1\}^k$, where A is a weight matrix and f is a threshold (indicator) function

Goals

- $F = fA : \{0,1\}^n \to \{0,1\}^k$, where A is a weight matrix and f is a threshold (indicator) function
- Theoretically rigorous, combinatorially simple

Goals

- $F = fA : \{0,1\}^n \rightarrow \{0,1\}^k$, where A is a weight matrix and f is a threshold (indicator) function
- Theoretically rigorous, combinatorially simple
- Formal Dentate Gyrus properties
    - Decorrelation of input (measured by normalized dot product)
    - Error-correcting information and redundancy
    - Pattern separation
    - The fidelity of F is a controlled parameter

Goals

- $F = fA : \{0,1\}^n \to \{0,1\}^k$, where A is a weight matrix and f is a threshold (indicator) function
- Theoretically rigorous, combinatorially simple
- Formal Dentate Gyrus properties
  - Decorrelation of input (measured by normalized dot product)
  - Error-correcting information and redundancy
  - Pattern separation
  - The fidelity of F is a controlled parameter
- Refinements matching biological constraints

A General Approach

- Suppose n is determined at the outset and let

$$\Delta = \{\eta_i\} \subset \mathcal{P}(\{1, \ldots, n\}).$$

A General Approach

- Suppose n is determined at the outset and let

$$\Delta = \{\eta_i\} \subset \mathcal{P}(\{1, \ldots, n\}).$$

- Δ will control the fidelity of our coding.

A General Approach

- Suppose n is determined at the outset and let

$$\Delta = \{\eta_i\} \subset \mathcal{P}(\{1, \ldots, n\}).$$

- Δ will control the fidelity of our coding.
- Define $A_\Delta = [a_{i,j}]$ where $a_{i,j} = 1/|\eta_i|$ if $j \in \eta_i$ and 0 otherwise.

A General Approach

- Suppose n is determined at the outset and let

$$\Delta = \{\eta_i\} \subset \mathcal{P}(\{1, \ldots, n\}).$$

- Δ will control the fidelity of our coding.
- Define $A_\Delta = [a_{i,j}]$ where $a_{i,j} = 1/|\eta_i|$ if $j \in \eta_i$ and 0 otherwise.
- With $f(\alpha) = 0$ for $\alpha < 1$, $f(\alpha) = 1$ for $\alpha \geq 1$, the map

$$F = fA_\Delta$$

is a sparse coding of $\{0,1\}^n$ into $\{0,1\}^k$, where $k = |\Delta|$.

Example

Let $n = 4$, $\Delta = \{\{1,2\},\{2,3\},\{2,3,4\}\}$. Then,

$$A_\Delta = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Example

Let $n = 4$, $\Delta = \{\{1,2\},\{2,3\},\{2,3,4\}\}$. Then,

$$A_\Delta = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Effectively $F(x)_i = 1$ if and only if $x_j = 1$ for all $j \in \eta_i$, and $F(x)_i = 0$ otherwise.

Example

Let $n = 4$, $\Delta = \{\{1,2\},\{2,3\},\{2,3,4\}\}$. Then,

$$A_\Delta = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Effectively $F(x)_i = 1$ if and only if $x_j = 1$ for all $j \in \eta_i$, and $F(x)_i = 0$ otherwise.

First simplification: Assume Δ is composed of all p-sized subsets for some $1 < p \leq n$.

Some Basic Results

Proposition

If $||x_i||_1 = q_i$*, then* $||y_i||_1 = \binom{q_i}{p}$*, and that* F *maps* $E = \{||x||_1 \geq p\}$ *injectively into a subset* $F(E)$ *of* $\{0,1\}^k$ *where* $k = \binom{n}{p}$*.*

Some Basic Results

Proposition

If $||x_i||_1 = q_i$*, then* $||y_i||_1 = \binom{q_i}{p}$*, and that* F *maps* $E = \{||x||_1 \geq p\}$ *injectively into a subset* $F(E)$ *of* $\{0,1\}^k$ *where* $k = \binom{n}{p}$*.*

Theorem

The action of F *on* E *non-trivially decreases normalized dot produdcts.*

Some Basic Results

Proposition

If $||x_i||_1 = q_i$, *then* $||y_i||_1 = \binom{q_i}{p}$, *and that* F *maps* $E = \{||x||_1 \geq p\}$ *injectively into a subset* $F(E)$ *of* $\{0,1\}^k$ *where* $k = \binom{n}{p}$.

Theorem

The action of F *on* E *non-trivially decreases normalized dot produdcts.*

Theorem

Denote $n - ||x - x'||_1 = r$, $||x||_1 = q$, $||x'||_1 = q'$, $\delta_1 = q - r$, $\delta_2 = q' - r$. *We have*

$$d(F(x), F(x')) = \delta_1 \binom{q-1}{p-1} + \delta_2 \binom{q'-1}{p-1}.$$

Biology Inspires Constraints

- General framework has been established

Biology Inspires Constraints

- General framework has been established
- Look to biology for plausible extensions

Biology Inspires Constraints

- General framework has been established
- Look to biology for plausible extensions
  - More realistic behavior
  - Generalize from assumption that Δ comprises p-sized subsets
  - Model's reaction to these constraints measure suitability

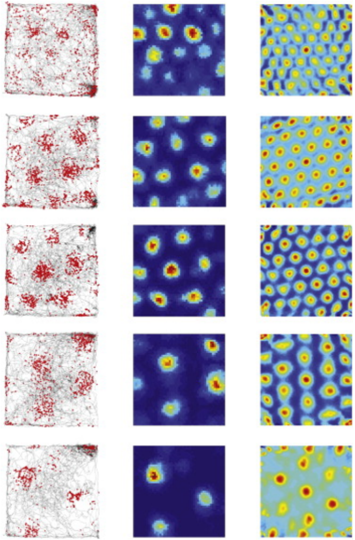
Biology Inspires Constraints

- General framework has been established
- Look to biology for plausible extensions
  - More realistic behavior
  - Generalize from assumption that Δ comprises p-sized subsets
  - Model's reaction to these constraints measure suitability
- Two types of generalization:

Concept	Method	Biology
Mimic Input Structure	Prune Δ	Grid Cells
Adaptive Fidelity	Mixed Coding	Adult Neurogenesis

Pruning Δ

If the distribution of the input data is known, we can remove entries from Δ and still maintain good fidelity.

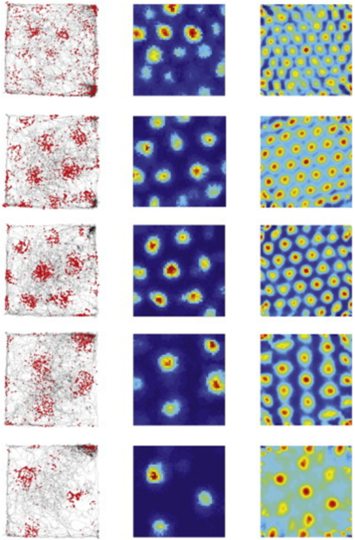


Giocomo, Moser and Moser

Pruning Δ

If the distribution of the input data is known, we can remove entries from Δ and still maintain good fidelity.

If we remove η from Δ and η is never observed, previous results still hold.



Giocomo, Moser and Moser

Pruning Δ

If the distribution of the input data is known, we can remove entries from Δ and still maintain good fidelity.

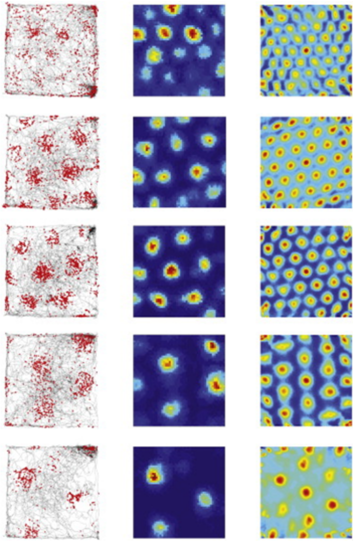
If we remove η from Δ and η is never observed, previous results still hold.

Example

Grid cells in EC encode spatial information using modular code. For $\lambda_0, \ldots, \lambda_T$ relatively prime,

$$x \mapsto (x \bmod \lambda_0, \ldots, x \bmod \lambda_T)$$

Chinese Remainder Theorem gives uniquely represented position.



Giocomo, Moser and Moser

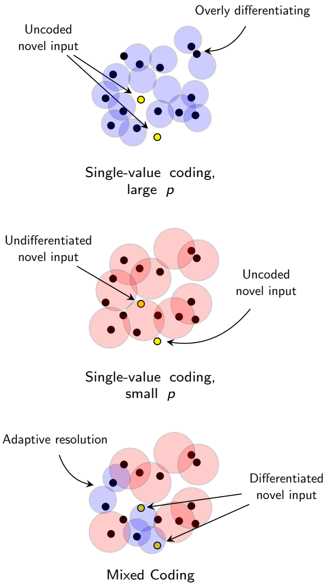
Pruning Δ

- Take input space group isomorphic to grid cell input
- Remove inadmissible η_i from Δ
- Dramatically reduces target dimension k
- Size and activity within literature expectations

Mixed Coding

Example

- The brain experiences adult neurogenesis—the development of new neurons throughout life.
- Young neurons = broadly tuned
- Old neurons = tightly tuned
- Mixed coding increases information capacity.

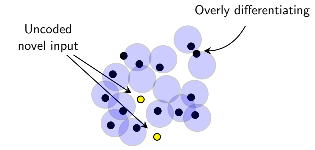


Mixed Coding

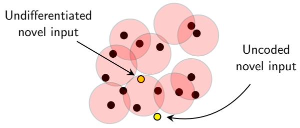
Example

- The brain experiences adult neurogenesis—the development of new neurons throughout life.
- Young neurons = broadly tuned
- Old neurons = tightly tuned
- Mixed coding increases information capacity.

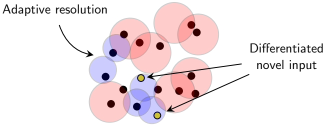
Choose $\eta_i \in \Delta$, $p' < p$.



Single-value coding, large p



Single-value coding, small p



Mixed Coding

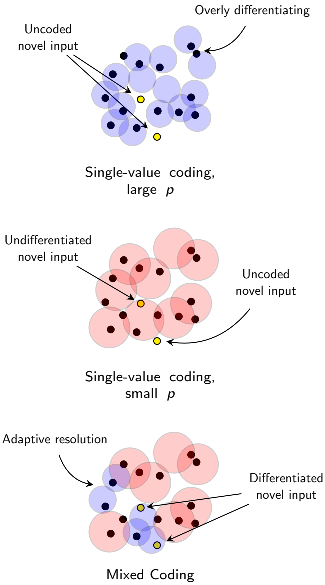
Mixed Coding

Example

- The brain experiences adult neurogenesis—the development of new neurons throughout life.
- Young neurons = broadly tuned
- Old neurons = tightly tuned
- Mixed coding increases information capacity.

Choose $\eta_i \in \Delta$, $p' < p$.
Expand Δ to Δ' by

$$\Delta' = (\Delta \setminus \{\eta_i\}) \cup \{p'\text{-sized subsets of } \eta_i\}.$$



Mixed Coding

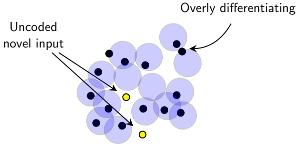
Example

- The brain experiences adult neurogenesis—the development of new neurons throughout life.
- Young neurons = broadly tuned
- Old neurons = tightly tuned
- Mixed coding increases information capacity.

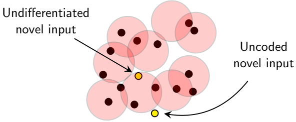
Choose $\eta_i \in \Delta$, $p' < p$.
Expand Δ to Δ' by

$$\Delta' = (\Delta \setminus \{\eta_i\}) \cup \{p'\text{-sized subsets of } \eta_i\}.$$

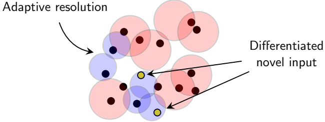
Conditions exist to guarantee sparsity and decorrelation.



Single-value coding, large p



Single-value coding, small p



Mixed Coding

Summary

- F is a simple combinatorial code for Dentate Gyrus sparse coding and pattern separation.

Summary

- F is a simple combinatorial code for Dentate Gyrus sparse coding and pattern separation.
- Theoretical tractability allows for formal properties.

Summary

- F is a simple combinatorial code for Dentate Gyrus sparse coding and pattern separation.
- Theoretical tractability allows for formal properties.
- Biological contexts inform refinement.

Summary

- F is a simple combinatorial code for Dentate Gyrus sparse coding and pattern separation.
- Theoretical tractability allows for formal properties.
- Biological contexts inform refinement.
- Two generalization methods give control over encoding fidelity.

Summary

- F is a simple combinatorial code for Dentate Gyrus sparse coding and pattern separation.
- Theoretical tractability allows for formal properties.
- Biological contexts inform refinement.
- Two generalization methods give control over encoding fidelity.