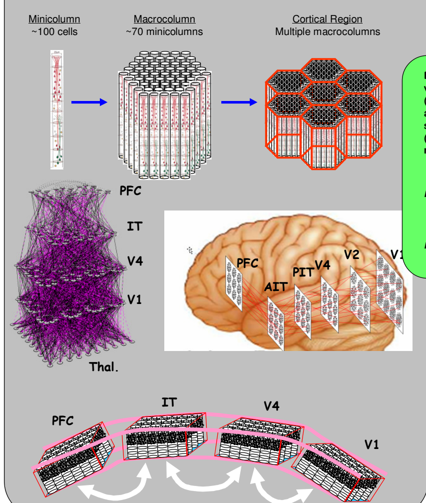
Gerard J. Rinkus,
Visiting Scientist, Lisman Laboratory
Volen Center for Complex Systems,
Brandeis University, Waltham, MA
(grinkus@brandeis.edu)

Population Coding using Familiarity-Contingent Noise

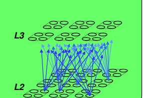
Main Idea

- The principal neurons in a competitive cluster should compete on the basis of their specific inputs (via synapses carrying informational content, reflecting prior learning).
- How does a sparse set of cells in a patch **of cortex** (e.g., macrocolumn) become bound together as a *population code* representing a particular concept?
- One way is to compute a measure of the familiarity (inverse novelty) that depends on the pattern of inputs to all cells in the patch, i.e., a global familiarity measure.
- We propose a specific algorithm for this.
- The algorithm implies that the process of choosing which cells become active in a region in any small Δt, e.g., a gamma cycle, has two stages.
  - Stage 1: principal cells integrate their specific inputs and compete with neighbors. One emerges as victor rapdily and transiently and its input is gated out of the area as a *local* (to the competitive cluster) measure of familiarity.
  - Stage 2: these local familiarty measures are all gated to a centralized place that computes their average, resulting in a global (to the macrocolumn) familiarity measure, *G*. *G*, or rather, a function of *G*, is then fed back to to the **macrocolumn**, where it influences a second round of competition (in the second half of the gamma cycle). In particular:
    - If *G* ≈ 0 (global unfamiliarity), a large amount of noise is added into the second round of competition. Thus, the specific informational inputs are drowned out. This results in a choice of winners having only chance-level intersection with any previous population code previously assigned in this **macrocolumn**.
    - If *G* ≈ 1 (global familiarity), then no noise is added into the second round, allowing specific informational inputs, reflecting prior learning, to determine which cells win.
    - Rather than picking (all the cells comprising a sparse) code in *one* decision event, our algorithm picks the code with Q independent decision events (one per cluster). This does not deterministically guarantee that the code as a whole is selected or not (as in other, localist neural decision theories), but depending on the parameters of the distributions in each cluster, it can make the likelihood of the code as a whole being reactivated be arbitrarily close to one.

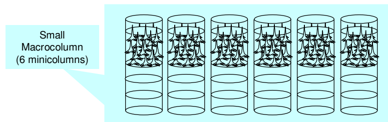
Overarching Hypothesis



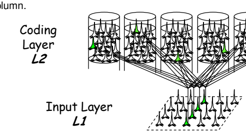
BU input summations, v, can be normalized (Step 2) because activity level in the subjacent level (afferent region) is normalized (to P)



- Minicolumn functions as a winner-take-all (WTA) Module
  - One principal cell becomes active (wins) in each discrete processing cycle
- Processing occurs simultaneously, and in phase, in all of the macrocolumn's minicolumns, e.g., in one gamma cycle (~30 ms)
  - cf. Fries et al. (2007) …WTA in a gamma cycle…
- In each cycle, the ***set*** of winners in the macrocolumn constitutes a sparse ***population code*** within that macrocolumn.



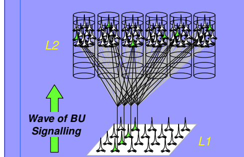
- ***But***, how do ***multiple*** minicolumns function as a unit?
- What could ***bind*** together the simultaneous winners across multiple minicolumns into a permanent ***population code***?
- ***Answer:***
  - Coactivity and the coordinated learning that occurs during the coactivity.
  - A ***global*** (e.g., macrocolumn-level) measure of the familiarity of the input.
  - A process by which that measure influences which cells win in each minicolumn.



The Need for a Global Familiarity Measure

Learning

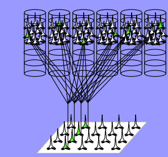
- Six winners, code C_1, are chosen randomly
- P_1 is associated to C_1



Input pattern P_1 presented

Test: Familiar

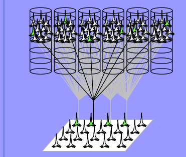
- C_1 cells have high bottom-up (BU), *v*, summations
- The desired behavior is that all C_1 cells should be reactivated



P_1 presented again

Test: Novel

- C_1 cells have low, *but still maximal*, BU summations, *v*
- The Local WTA, alone, would reactivate the entire C_1 code
- The desired behavior is that C_2 has small overlap with C_1



Novel input P_2 presented

Unfamiliar (Novel) ($G \approx 0$)

P : Set of active cells in the afferent pattern to this level

$w(k,i)$: Weight (0,1) of cell k to cell i

$a(k)$: Activity level (0,1) of cell k

1 Potential

$$v(i) = \sum_{k \in P} a(k) \times w(k,i)$$

2 Normalized Potential

$$V(i) = v(i) / |P|$$

3 Max Normalized Potential in j^{th} minicolumn

$$\hat{V}_j = \max_{i \in M_j} (V(i))$$

4 Average Max Normalized Potential over all minicolumns

$$G \equiv \sum_{k=1}^{Q} \hat{V}_k / Q \approx 0.1$$

5 Determine amount of noise to add into winner selection process

6 V-to-ρ Transform

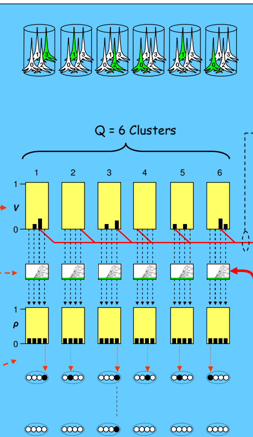
$$a = \left(\left[\frac{G - V_\alpha}{1 - V_\alpha} \right]^+ \right)^d \times Q \times K$$

$$\xi(i) = \frac{a}{1 + e^{-b(V(i) - c)}} + 1$$

$$\rho(i) = \frac{\xi(i)}{\sum_{k \in \text{cluster}} \xi(k)}$$

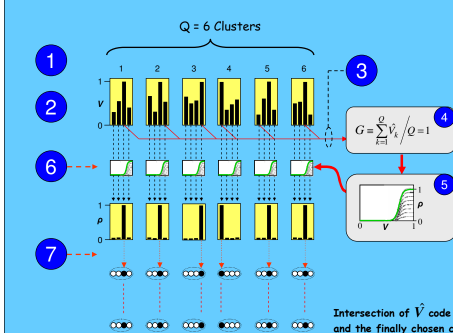
7 Select Final Winners

Choose final winner in each cluster according to ρ-distribution



Intersection of $\hat{V}$ code and the finally chosen code

Familiar ($G = 1$)



$$G \equiv \sum_{k=1}^{Q} \hat{V}_k / Q = 1$$

Intersection of $\hat{V}$ code and the finally chosen code

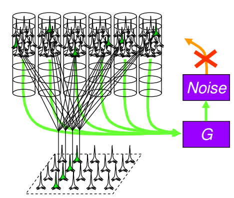
What to do with G?

G ≈ 1: Highly familiar input

Should reactivate the same code that was learned initially.

Global information does ***NOT*** need to be fed back to the local minicolumns.

- Simply, allow max-V cells to win in their respective minicolumns.

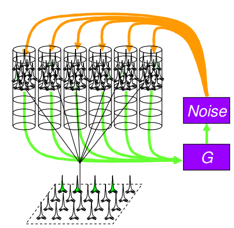


G ≈ 0: Highly novel

Should choose winners pretty much at random.

Global information ***DOES*** need to be fed back.

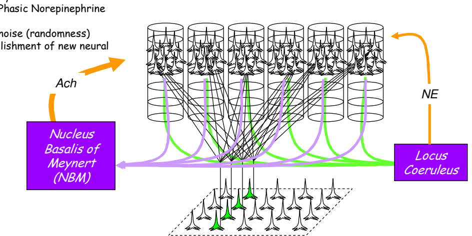
- Add a lot of noise into process of cells computing their individual potentials



Important Computational Implication

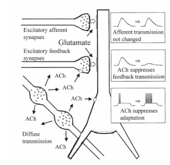
- Principal cells undergo ***two*** rounds of integration and competition within the basic computational cycle.
- The 1st round results in the activation of a ***preliminary*** code which drives the computation of *G*
- The 2nd round includes a variable amount of noise (depends inversely on *G*), resulting in the ***final*** code for the cycle.

- Novelty
- High Phasic Norepinephrine (NE)
- High noise (randomness)
- Establishment of new neural codes



Ach Data

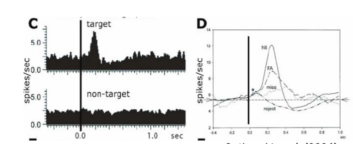
- Acetylcholine, not NE, is the main regulator of the level of spontaneous activity of cortical neurons: Isakova & Mednikova (2007)
- Kimura, Fukada, Tsumoto (1999): Ach causes:
  - synaptic facilitation,
  - synaptic suppression,
  - direct hyperpolarization,
  - direct depolarization
- ACh increases depolarization, excitability, and reduces spike frequency adaptation:
  - Tateno et al. (2005)
  - Hasselmo
- Increased Ach leads to learning of finer categories (more detail)
  - Olfactory - Linster et al. (2001)
  - Auditory – Weinberger et al. (2006)



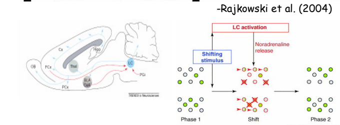
- Hasselmo & McGaughy (2004)

NE Data

- LC activated by novelty: Vankov et al. (1995)
- Phasic NE: latency (~100-200 ms), short duration (~100-200 ms): Clayton et al. (2004)
  - Signals "unexpected uncertainty", i.e., novelty. Dayan & Yu (2006)
  - Increase signal/noise (cf. Hasselmo et al. 1997)
  - "provoke or facilitate dynamic reorganization of target neural networks, permitting rapid behavioral adaptation to changing environmental imperatives" - Bouret & Sara (2005)
  - NE burst causes rapid state shift in hippocampal network: Brown et al (2005)
- "The overall profile suggests lower levels of norepinephrine may facilitate pattern completion or memory retrieval while higher levels would recruit global remapping and promote long-term episodic memory." - Harley & Helen (2007) (…in hippocampus…)
- PFC sends fibers back down to LC (Arnsten & Goldman-Rakic, 1984; Sara & Herve-Minvielle, 1995; Jodoj et al., 1998).



-Rajkowski et al. (2004)



- Bouret & Sara (2005)

• NIH Training Grant, 5 T32 NS07292, with John Lisman, Brandeis University Redwood Neuroscience Institute (Jeff Hawkins) Scientific Systems Co Inc