3. Networks

Networks

1. Biology of Neocortex (“cortex”)
2. Categorization and Distributed Reps
3. Bidirectional Excitation and Attractors
4. Inhibitory Competition and Activity Regulation

Neurons: Excitatory and Inhibitory

Excitatory = main info processing, long-range connections
Inhibitory = local, activity regulation and competition
The 6 Cortical Layers

Functions of Layers

Connection Directions
Bidirectional Symmetry

Biology ⇒ Function

• Feedforward excitation = categorization of inputs
  — larger patterns, more invariant w.r.t instances & space
• Feedback excitation = attractor dynamics
  — ambiguity resolution & constraint satisfaction
• Lateral inhibition = competition, activity regulation
  — sharpens response

Ambiguity Resolution
Hierarchical Categorical Representations

- Successive layers of neural detectors
- Progressively more abstract

The Chair Category
How would you define “chair”?
Socrates asks, “what is that common quality, which is the same in all these cases, and which is called courage?” (Laches 191e)

Getting the right ones is key
- Two men are dead in a cabin in the woods.
- The cabin itself is not burned, but the forest all around is burned to cinders.
- How did the men die?
Categories: A Philosophical Problem

- A long-standing problem:
  - Socrates (d. 399 BCE) says, “that which we know we must surely be able to tell.” (*Laches* 190c)
  - Must knowledge be encoded in language-like structures?
- What makes a mental categorization accurate? Is there something “real” about a “chair”?
- Stereotypes are mental categories...
- Can you encode multiple categories at the same time?

Distributed Representations

- Let a 1,000 categories bloom… You’ve got the room in your head (billions of neurons)
- Each neuron can respond to multiple things (graded similarity)
- And each thing activates many neurons (who knows what is going to be relevant this time?)
Cell Responses in V4

Orientation Columns

Orientation Columns
Topographic Organization

Topographic Maps: Bat Auditory Cortex

Sparse Distributed Representation

- Localist representation
  - "grandmother cells"
  - unlikely in brain
- K-out-of-N detectors
  - typically 15–25% of neurons active
- Approximate orthogonality

(monkey IT cortex)
Sparse Distributed Representations

Not Just Monkeys

Coarse Coding

• Broadly-tuned receptive fields
• Population-coding of precise values
• Common throughout sensory and motor areas
Localist Representations?

emergent demonstration: Face Categorization I

Bidirectional Excitation

- Functions
  - recognition
  - top-down imagery
  - ambiguity resolution
  - pattern completion
- Attractor dynamics
  - convergence on good representation
  - energy vs. harmony

(fig. < O'Reilly, Comp. Cog. Neurosci. )
What Are These?

A Big Network Model...

Bidirectional Dynamics

Wyatte, Herd, Mingus, O'Reilly (2012)
emergent demonstration: Face Categorization II

emergent demonstration: Cats and Dogs

emergent demonstration: Necker Cube
Inhibitory Competition and Activity Regulation

- Activity regulation
- Selective attention
- Competition
- Sparse distributed representation

Activity Regulation

- Feedback: reactive, reflects actual level of activity, robust, responsive, may be unstable
- Feedforward: anticipatory, limits feedback oscillation, slow, brittle
- Work well together

Competitive Network

(fig. < O'Reilly, Comp. Cog. Neurosci.)
Competitive Learning

- Competitive learning network
  - two layers, randomly initialized weights
  - second is self-reinforcing, mutually inhibitory
  - "winner takes all" dynamics
- Learning
  - winner moves toward last
  - weight vectors move to centers of clusters

FFFB Inhibition Approximation

- Approximates total effect of all inhibition in a layer
- Inhibition determined by feedforward and feedback terms:
  \[ g_i(t) = g_i(FF(t) + FB(t)) \]
  \[ g_i(t) = g_i(FF(t) + FB(t)) \]
  - FF term is excess average input over set point:
    \[ FF(t) = \frac{1}{n} \sum_{i=1}^{n} \eta_i - \eta_0 \]
    where \( \eta(t) = n^{-1} \sum_{i=1}^{n} \eta_i \) is average input
  - FB term varies with average activity:
    \[ FB(t) = dt [fb(y) - \frac{FB(t)}{n}] \]
    where \( y(t) = n^{-1} \sum_{i=1}^{n} y_i \) is average activity
  - Will stabilize with \( FB(t) = fb(y) \)

emergent demonstration: Inhibition