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

(slide < O'Reilly)
The 6 Cortical Layers

(slide < O'Reilly)
Functions of Layers

Input
layer 4
from sensation or other areas

Hidden
layers 2 & 3

Output
layers 5 & 6
to motor systems or other areas

(fig. < O’Reilly, Comp. Cog. Neurosci.)
Connection Directions

Feedforward
• from Hidden in lower to Input in higher

Feedback
• from Hidden & Output in higher to Hidden & Output in lower

Lateral
• from Hidden and Output to all three layers in same area

Bidirectionality
• pervasive
Bidirectional Symmetry

(slide < O'Reilly)
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

COSC 494/594 CCN
(fig. < O’Reilly, Comp. Cog. Neurosci.)
Hierarchical Categorical Representations

- Successive layers of neural detectors
- Progressively more abstract

(fig. < O’Reilly, *Comp. Cog. Neurosci.*)
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?
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)
The central problem of AI is: What are \( a \) and \( i \)?
— Douglas Hofstadter
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?

(slide based on O'Reilly)
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?)
Graded Responses

(slides < O'Reilly)
Cell Responses in V4

(fig. < Clark, *Being There*, 1997)
Orientation Columns

(fig. < Nicholls & al., *Neur. to Brain*)
Orientation Columns

(fig. < Nicholls & al., *Neur. to Brain*)
Topographic Organization
Topographic Maps: Bat Auditory Cortex

(figs. from Suga, 1985)
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

(slide < O’Reilly)
Coarse Coding

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

COSC 494/594 CCN
(fig. < O’Reilly, *Comp. Cog. Neurosci.*)
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

Occlusion = 50%

Cosine to Unoccluded

Cycle of Settling

Wyatte, Herd, Mingus, O'Reilly (2012)

(slide < O'Reilly)
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

(fig. < O’Reilly, Comp. Cog. Neurosci.)
Competitive Network
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:
  \[ g_i(t) = g_i[\text{FF}(t) + \text{FB}(t)] \]
- FF term is excess average input over set point:
  \[ \text{FF}(t) = \text{ff}[\langle \eta \rangle - \text{ff0}]^+ \]
  where \( \langle \eta \rangle = n^{-1} \sum_{i=1}^{n} \eta_i \) is average input
- FB term varies with average activity:
  \[ \dot{\text{FB}}(t) = \text{dt}[\text{fb}\langle y \rangle - \text{FB}(t)] \]
  where \( \langle y \rangle = n^{-1} \sum_{i=1}^{n} y_i \) is average activity
- Will stabilize with \( \text{FB}(t) = \text{fb}\langle y \rangle \)
emergent demonstration: Inhibition