In Transition

• from Part I: Basic Mechanisms.
• to Part II: Perception, Attention, Memory, Language, Higher Level Cognition

Summary of Part I: Basic Mechanisms

1. Biological realism
2. Distributed Representations
3. Inhibitory Competition
4. Bidirectional Activation Propagation
5. Error-driven Learning
6. Hebbian Learning

Micro and Macro-Neurocomputomics

Micro = basic mechanisms common across brain areas.
Macro = organization, differentiation, interactions of brain areas.
Need to consider general principles for macro organization before we can think about larger cognitive functions.

Macro Structural Principles

• Hierarchical sequence of transformations.
  – Emphasize some distinctions, ignore others
  – For object recognition you want to ignore differences in location, lighting, size, rotation
  – When reaching for objects, you want to emphasize location

• Specialized pathways.
  – Location-invariant object recognition vs. recognizing orientation & location for actions (seeing for identifying and seeing for action)
  – Patients with ventral stream damage have “blindsight” (e.g., Milner & Goodale 1995): they can reach and grasp objects at different locations/orientations but cannot perceive them!

• Inter-pathway interactions.
  – Visual attention is an emergent property of interactions between different brain areas.
  – Visual attention is an emergent property of interactions between different brain areas.
  – Visual attention is an emergent property of interactions between different brain areas.

Level Cognition

• To Plan (Perception, Attention, Memory, Language, Higher Functions, Executive Functions)
• To Plan (Perception, Attention, Memory, Language, Higher Functions, Executive Functions)
• To Plan (Perception, Attention, Memory, Language, Higher Functions, Executive Functions)
Macro Structural Principles

- Higher-level association areas
  - Integration of e.g., visual and auditory information
  - At extreme, thought to underlie synesthesia

- Large-scale distributed representations
  - Knowledge is distributed across multiple brain areas
  - Multiple areas participate in representing a given thing (e.g., apple)
  - Each area represents multiple things
  - Similar idea as distributed representation among units, but now across multiple areas/modalities, etc.

Macro Dynamic Principles

- Processing as multiple constraint satisfaction
- Attractors, settling dynamics, amplification
- Inhibitory competition: attention

Where do constraints come from?
- Perceptual inputs ("bottom-up" constraints)
- We also have the ability to maintain firing of neurons in the absence of bottom-up stimulation
- Make use of bidirectional excitatory connections
- Therefore, constitute an active memory
- Which do constraints come from?
- Where do constraints come from?
  - Perceptual inputs ("bottom-up" constraints)
  - Which do constraints come from?

General Functions of the Cortical Lobes

- Frontal lobe: Motor control, cognitive control (planning)
- Parietal lobe: Representing body, external spaces
- Occipital lobe: Representing visual perception, object recognition
- Temporal lobe: Visual processing

Other Areas

- Hippocampus (rapid episodic encoding)
- Amygdala (emotion, affective associations)
- Basal ganglia (motor control, sequencing, reward learning, gating of prefrontal cortex)
- Cerebellum (motor learning, cognitive roles via timing?)
- Midbrain neuromodulators (VTA - dopamine, raphe - serotonin)
Tripartite Functional Organization

Defined by set of functional trade-offs.

Multiple systems in decision making

Computational Trade-offs in Learning & Memory

Trade-offs: Computational objectives that are mutually incompatible and thus cannot be achieved by a single brain system.

→ Begin to address psychological distinctions between different learning & memory processes, informed by mechanisms required.

- Learning statistical structure vs. memorizing specific events
- Isolated maintenance (holding in mind multiple items of info) vs. inference (spreading activation: smoke → fire)
- Robust maintenance vs. rapid updating

Learning must also be able to learn quickly.

Learning must be slow to capture (statistical) structure (averaging).

But you also have to be able to learn rapidly.

Tradeoffs solved by 2 systems: cortex learns slowly, hippo rapidly.

3rd system: Active memory (prefrontal cortex) ≈ fastest (immediately accessible) but learning to develop pfc reps in first place is slow, allows abstraction.

Slow vs. Fast Learning

lrate = .005
lrate = .1
lrate = 1

Learning must be slow to capture (statistical) structure (averaging).

But you also have to be able to learn rapidly.

Tradeoffs solved by 2 systems: cortex learns slowly, hippo rapidly.

3rd system: Active memory (prefrontal cortex) ≈ fastest (immediately accessible) but learning to develop pfc reps in first place is slow, allows abstraction.
1b. Slow vs Fast [Reinforcement] Learning

- Slow learning is necessary to capture best actions that work on average.
- But you also have to be sensitive to rapid changes in value (e.g., stock market).

Tradeoffs solved by two systems: BG learns slowly, PFC flexibly updates new states and can override habitual choices.
What do these findings tell us?

- The same action (lever-pressing) can arise from two psychologically & neurally dissociable pathways:
  - Moderately trained behavior is goal-directed: dependent on outcome representation of what might happen.
  - Overtrained behavior is habitual: apparently not dependent on outcome, like S-R learning.

- S-R habits really exist (in humans too), they just don’t describe all of behavior.

- Lesions suggest two parallel systems, in that the intact one can apparently support the behavior at any stage. (see also BG vs Hippo in S-R vs cognitive map)

...
Summary: Model-based vs Model-free RL

Instrumental conditioning reveals that rats indeed have S-R habits (and humans, Tricomi et al, 2009). But even humble rats are cognitive: must distinguish habits from goal-directed behaviors.

Understand this distinction algorithmically in terms of different RL strategies for decision making, and mechanistically in terms of functional properties of biological systems involved (BG and PFC).

Note: The same overt behavior can be the product of different neural (computational) systems (controllers).

For computational models of these and related phenomena, including how the brain might arbitrate between the two systems, see Daw, Niv & Dayan (2005) and Frank & Claus (2006).
Reinforcement learning: Dopamine can reinforce rewarding actions so that they are more likely to be executed in the future. This allows an agent to exploit the best possible actions in a situation that are most likely to lead to reward. But what if other possible actions are even better? How would you ever know?

Norepinephrine (NE) modulates the noise in cortical representations, allowing the agent to sometimes randomly select some other action. This NE system is itself controlled by overall long-term utility, so that when previous actions no longer lead to reward over an extended period of time, the NE system responds by increasing noise and exploration of new actions.

Also: "Directed" exploration, e.g. toward uncertain options (optimistic).
Recursion and Subroutine-like processing

- In middle of processing, need to perform same processing (recursion) or different processing (subroutine).
- Easy in standard serial computer (store current state, call subroutine with appropriate arguments).
- Harder when data and processing not separated!

HCMP, PFC

Nobody's perfect...

The mouse the cat the dog bit chased squeaked.

Generalization

- How to recognize new inputs given dedicated, specialized reps?
- Distributed representations: combinations of existing features.
- Abstraction: learn that all dogs might bite, not just that Spike bit me.

Important Distinctions

- Controlled vs Automatic Processing.
- Declarative/Procedural vs Explicit/Implicit Processing.

Consciousness = influence (on Constraint Satisfaction):

- Intensity: Higher = more influence
- Duration: Longer = more influence
- Centrality: More influence on other areas

HCMP, PFC

Transfer is not good at all.

How to recognize new inputs given dedicated, specialized reps?

A Cognitive Architecture