Computational Models of Neural Systems

David S. Touretzky

Lecture 2:
Brains and Computation

What Is a Function?

- Abstract definition: a function maps inputs to outputs.

  domain: the set of possible inputs ("input space")

  range: the set of possible outputs ("output space")

- Theinput and output spaces can be vector spaces:

  [2, -7.4, 0.003] ➜ [1.5, 2.8]

- Input from the optic nerve at time *t*: a vector of dimension one million.

- What about stochastic behavior?

  f: *number* ➜ {spike, no spike}

  f(15) = spike *(38% of the time)*
  no spike *(72% of the time)*

  This function does not map its input to a single output.

- One solution: let the output be a probability distribution.

  f: *number* ➜ (g: {spike, no spike} ➜ [0,1])

What Is Computation?

- Not precisely defined yet. Many things might qualify.
- Turing machines, cellular automata, RAMs (random access memory) machines. *Brains?*
- The Church-Turing hypothesis: any computable function can be computed by a Turing machine.
- When is a physical system a computational system?

  1. Its physical states correspond to the representations of an abstract computational system.
  2. Transitions between its states can be explained in terms of operations on those representations.

- There must exist an appropriate (*revealing*) mapping between physical states and elements of the domain and range of the function being computed.

Structured vs. Unstructured Spaces

Let $\boldsymbol{S}$ be the set consisting of the following 24 items:

A1P	A1Q	A1R	A1S
B1P	B1Q	B1R	B1S
A2P	A2Q	A2R	A2S
B2P	B2Q	B2R	B2S
A3P	A3Q	A3R	A3S
B3P	B3Q	B3R	B3S

Let **f:** $\boldsymbol{S} \rightarrow [0,1]$ be a function implementing an arbitrary mappng from a set to an interval. Can't say much about the domain; it's just a set of discrete points ("atoms", "symbols").

Let $\boldsymbol{S}_1 = \{A, B\}$ $\quad \boldsymbol{S}_2 = \{1, 2, 3\}$ $\quad \boldsymbol{S}_3 = \{P, Q, R, S\}$

Let **g:** $\boldsymbol{S}_1 \times \boldsymbol{S}_2 \times \boldsymbol{S}_3 \rightarrow [0,1]$ be a function from a Cartesian product space to an interval.

The domain of **g** has a *concise description* as a 3D space. The input points have *internal structure* which can be exploited in the definition of **g**.

<A,1,P> is more similar to <A,1,Q> than <B,3,S>.

Physical Realization of Computation

Functions are just mathematical abstractions.

How can werelate them to actions in the physical world?

1. Pick a function, and define its input and output spaces.
2. Decide how these input and output spaces should map onto states of the physical world.
3. A physical process must map the representation of each input into the representation of the corresponding output.
4. Whenever this process occurs, "computation" has taken place.

Physical Computation: The Slide Rule

Function being computed: product of two numbers.
Input: a pair of numbers (2-dimensional vector)
Output: a number.

Physical Realization:
The first input: a point on the surface of the (fixed) D scale.
The second input: a point on the surface of the (sliding) C scale.
The output: a point on the surface of the (fixed) D scale.

Computation process:

1. Move the sliding C scale so that the digit “1” is above the first input, which is a point on the D scale.
2. Move the cursor so that the index is over the second input, a point on the C scale.
3. Read the result point under the index on the D scale.

When Does Computation Happen?

If a slide rule falls in the forest and nobody sees it, has computation taken place?

Yes. If the mapping is defined for us, then from our point of view, physical action = computation. No observer is required.

The Hole in Our Definition

How to Play Expert Chess with an Anvil

Input space: a board configuration (64-dimensional vector.)
Output space: a legal move (one item in a set of a few hundred.)

Physical representation of the input: a chess board with plastic pieces.

Input mapping: pieces on the board should be physically laid out so as to agree with the input vector.

Computational process: position a 2 ton anvil directly above the chess board, at a height of 50 feet. Let it drop.

Physical representation of the output: a pattern of smooshed plastic shapes embedded in very thin cardboard.

What's wrong with this computation?

Interpreting the Debris

Output mapping:

1. If the board looks like any of:
   result is "P-K4".

2. If the board looks like any of:
   result is "P-Q3".

... Etc.

What's wrong?

To fully define this output mapping would require trillions of smooshed board images.

The physical process (dropping the anvil) isn't exploiting any regularities in the domain.

All the work is being done by the output mapping.
  Way too much work.

Expert-level chess can be described far more concisely if you use a digital computer instead of an anvil for your computations.

Systematic Interpretability

What makes for a good example of a computation?

Physical representations must be *systematically interpretable* as input and output values. The mapping is *concise.*

Arbitrary mappings don't provide a satisfying analogy between the abstract function and the physical system..

- Avoid input/output mappings of spaces having a simple structure that require lookup tables with trillions of entries.
- The physical process that takes input states to output states must also have a concise description; it can't have trillions of distinct steps.

Interesting functions are ones that cleverly exploit non-obvious regularities of the domain.

Brains as Computing Devices

What are some properties of brains as computing devices that distinguish them from digital computers?

For comparison, Sejnowski and Churchland list 13 properties of brains as computing devices.

The following pages are based on their discussion in chapter 2 of *The Computational Brain.*

1. Specialization of Function

- The brain is not homogeneous. Different areas subserve different functions.
- Subcortical areas: thalamus, basal ganglia, brainstem, etc.
- Cortex: six-layered structure. Is there a canonical circuit?
- Primary cortical areas: V1 (vision), S1 (somatosensory), A1 (auditory), M1 (motor).
- Specialization is not strict.

  Some cells in V1 respond to nonvisual signals, such as eye movement.
- At the cellular level: variety of cell types with different characteristics.
- At the sub-cellular level: variety of channel types, receptor types.

2. Numbers

	Neurons	Synapses
Humans	10^{12}	10^{15}
Rats	10^{10}	10^{13}
1 mm^3 of cortex	10^5	10^9

A cortical neuron averages 4.12×10^3 synapses (cat or monkey.)

[Note: Cherniak cites estimates as high as 39,000 synapses per neuron in some parts of cortex.]

In primary visual cortex, because cells are packed so densely (100,000 per mm^3), the number of synapses is lower. Estimated at 1.17×10^3.

Can these Numbers be Trusted?

Quantitative anatomical estimates are still in need of refinement.

> *Of the 10^{10} neurons in the human brain, 10^{11} are in the cerebellum.*

Cherniak (1990) observes an order of magnitude difference in estimates of the surface area of the cortical sheet in mm^2:

Rakic (1981)	80,000
Shepherd (1979)	~120,000
Hubel (1979)	139,000
Blinkov & Glezer (1968)	160,000
Colonnier (1981)	200,000
Carpenter and Sutin (1983)	200,000
Hoffman (1983)	307,500
Mountcastle (1978)	400,000
Anderson & Hinton (1981)	1,000,000

Physical Constraints

- In his paper, *The Bounded Brain: Toward Quantitative Neuroanatomy*, Cherniak focuses on volume and area.
- Volume of the human brain: about 1.4 liters.
- Bounds on the connectivity of neurons can be derived if we can measure the volume of a connection.
- Models of the brain that require unreasonably large numbers of units or connections are "impossibility engines."
- They might still compute interesting things, but they cannot tell us how real brains work.

Demystifying the Brain

More observations from Cherniak:

- The brain is very complex, but not infinitely so.
- The cerebellum, concerned with posture and movement, contains four times as many neurons as cortex, the seat of language and conscious reasoning.
- A few spoonfuls of yogurt contain 10^{11} lactobaccillus bacteria: tens times the number of cortical neurons in a human brain.
- There are roughly 10^{13} synapses in cortex. Assume each stores one bit of information: that's 1.25 terabytes.
- The Library of Congress (80 million volumes, average 300 typed pages each) contains about 48 terabytes of data.
- Even the Lenin Library (25 million volumes) contains more data than can fit in one cortex.

3. Connectivity

- Across species, cortical neurons have roughly constant fan-in/ fan-out.

- Cortical neurons are sparsely connected.

- A single neuron contacts about 3% of the neurons underlying the surrounding square millimeter of cortex (Stevens 1989).

- Most connections are between, not within cell classes.

- Example: reciprocal connections between excitatory pyramidal cells and various types of inhibitory cells.

- Forward projections from area A to area B are generally matched by recurrent projections from B back to A.

- Curious fact: there appear to be 10 times as many feedback connections from V1 to LGN as feedforward connections from LGN to V1.

4. Analog Inputs/Discrete Outputs

- Neuron inputs are analog (continuous) values: epsp's and ipsp's.
- Neuron outputs are discrete: spike, or no spike.
- *Is this fair?*
- *What about spike rate, or phase?*
- Gap junctions? Ephatic coupling?

5. Timing: General

- Nervous systems must operate in real time.

- In perception, the time scale of computation must be matched to the time scale of relevant events in the world.

- In motor control, the time scale of motor commands must be matched to the response time of the body.

- Multiple inputs to a neuron can interact.

  So arrival times matter.

  Propagation delays matter.

6. Timing: Specifics

- An action potential (spike) lasts about 1 msec.
- Synaptic transmission, including electrotonic conduction in dendrites, takes about 5 msec.
- Synaptic potentials can last for 1 msec to many minutes.
- Transmission velocity in myelinated axons: 10-100 m/sec.
- In un-myelinated axons: less than 1 m/sec.
- Peak firing rate of a cortical neuron is typically about 100 Hz.
- No neuron fires faster than 1 kHz.
- How long can the chain of neuron firings be between an input and a response? Cf. data from tachistoscope experiments.
- Monosynaptic vs. multisynaptic paths: measure the time difference between input and output spikes.

7. Cell-to-Cell Effects

- The effect of an individual input to a neuron is weak.
- Typical single psp's are 1% to 5% of firing threshold.
- There are some exceptions: basket cells and chandelier cells are strongly inhibitory.

8. Firing Patterns

- Different types of neurons have different firing patterns:
  regular spiking
  fast-spiking
  repetitive-bursting

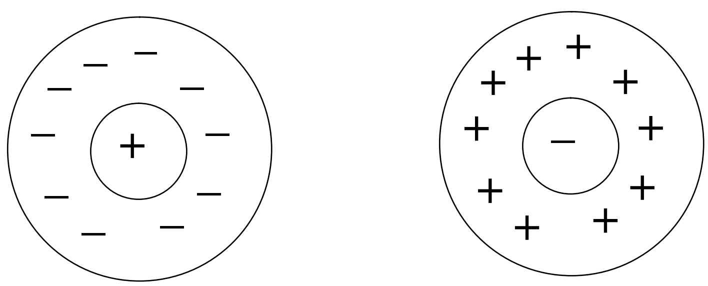
- Cells have "internal state" and can switch firing patterns.

- Some cells are natural oscillators. Determined by characteristics of the various types of ion channels.

- Most neurons exhibit random, spontaneous activity of a few spikes/sec. up to as much as 50 spikes/sec.

9. Receptive Fields

- The concept of “receptive field” comes from sensory systems.
- The region in which a particular cell responds to inputs.
- In the somatosensory system receptive fields are defined over the surface of the body. Fields for the fingertips are much smaller (and more densely packed) than fields for the arm.
- In the visual system, RF’s are defined on the retina.
- Retinal ganglion cells come in two types:
  on-center (off-surround)
  off-center (on-surround)



- Center-surround fields are also found in somatosensory neurons in thalamus and cortex.

10. Nonclassical Receptive Fields

- Cells’ responses can be modulated by events outside of their classical receptive fields.
- In area 17 of cats (V1), cells that repond to vertical edges had an enhanced response when distant cells, co-oriented and co-axial to them, were electrically stimulated (Nelson and Frost 1985.)
- Wavelength-sensitive cells in V4 are influenced by the color balance of the surrounding area outside their classical receptive field.
- Receptive fields can dynamically change shape. Modulated by attention?
- Receptive fields can shift in response to damage, e.g. RF’s of cells in V1 shift when their preferred area of retina is lesioned.

11. Specific and Nonspecific Systems

- There are five sources of widely-projecting neurons, each associated with a specific transmitter.

What	***Where***	***Transmitter***
locus coeruleus	brain stem	norepinephrine
raphe nucleus	midbrain	serotonin
substantia nigra	midbrain	dopamine
nucleus basalis	basal forebrain	acetylcholine
mammilary cells	hypothalamus	GABA

- These play a modulatory role, regulating the sleep-wake cycle, current level of awareness, and attention.

12. Action at a Distance

- Some neurotransmitters are dumped into the extracellular space, where they can affect nonsynaptic sites some distance from the point of release.
- Hormones produced by the endocrine system, such as estradiol, are released into the bloodstream.
- These can alter neural activity over wide regions of the brain.

13. Parallel Architecture

- Processing in the brain appears to be highly parallel, not just at the unit level, but at the systems level.
- In the monkey visual system:

  two parallel streams from the retina

  each originates in a different set of ganglion cells

  each projects to distinct sets of layers of the LGN:
  the *magnocellular* and *parvocellular* systems

  these in turn project to distinct sublaminae of layer 4 of V1

- The two visual streams are not perfectly separate.
- There are probably interactions between them at all levels.

Is the Brain Hierarchical?

- Many theories of the brain involve a hierarchy of processing levels.
- The nerves running to the sensory and motor organs are at the base of the hierarchy.
- “Reasoning” or “consciousness” is at the top.
- There are certainly hierarchical aspects to the brain’s organization, but...
- Not strictly bottom up: there are extensive feedback paths.
- There are many lateral connections between processing streams.
- There is no single apex to the hierarchy.
- There is not much hierarchy visible in higher association areas, such as frontal cortex.
- Motor control appears to be distributed; there is no one “command center.”

Types of Computation

1. Simple table lookup.
   Hard-wired circuitry.
2. Associative memory.
   Hippocampus? Piriform cortex? Cerebellum?
3. Competitive learning.
   Piriform cortex?
4. Parallel constraint satisfaction.
   Higher level association areas?
5. Gradient descent learning.
   Neocortex?