Contents

1 "Spontaneous Activity in the Brain: and what it means for neural computation" 2
1.1 The Question 2
1.2 The Problem 2
1.3 Three (essential) stories 3
1.4 Implications 8

2 Lecture 7 9
2.1 Recap 9
2.2 Hopfield networks 11
2.3 Cf: the Ising model: all edges are +1 11
2.4 A discrete-time system: Hopfield net (cont) 11
2.5 Goal of dynamical systems: pattern completion 11
2.6 How do you train a Hopfield net so that it pattern completes? 11
2.7 Do Hopfield nets work? 12
2.8 Boltzmann Machines 12
2.9 Back to modeling brain networks: the feedforward + recurrent (FFR) network 12
2.10 What is the solution? 13
2.11 Finally, suppose $\lambda_1 = 1$ 13
2.12 Nonlinear FFR networks 13
2.13 Modeling complex cells 14
2.14 Another FFR network: excitation-inhibition balance 14
2.15 E - I balance 14
2.16 How fruit flies remember smells 15
2.17 Flys algorithm (RP&C) outperforms LSH (2017) 15

1 "Spontaneous Activity in the Brain: and what it means for neural computation"

Speaker: Kenny Kay (kaykenneth@gmail.com)

1.1 The Question

- How is X reducible to neural activity?
  - Human intelligence
  - A general problem solver
  - Cognition
    - ↑ now described as following:
    - Perception
    - Memory
    - Decision-making
    - Imagination
    - Language
    - Etc.

1.2 The Problem

- We may have (unknowingly) adopted a **dominant paradigm** that is limiting us
  - Akin to What if the Sun doesnt revolve around the earth?
- Is there a dominant paradigm?
  - Lets look at the conventional framework
  - Levels
    - Function
    - Algorithm
    - Implementation
  - Approaches
    - Formal [down-arrow] (Cartesian) - top-down
    - Cognitive science
    - Comp neuro
    - Empirical [up-arrow] (Baconian) - bottom-up
    - Neurophysiology

        - Neuroscience
    - Apparently not, as many researchers take both approaches
    - Lets see where the empirical one takes us

1.3 Three (essential) stories

- ↑ about spontaneous activity in the brain
    - The dark horse
        - Scientists battle noise in sensory neurons
    - The elephant in the room
        - Scientists encounter strange giant neural activity
    - The rabbit hole
        - Scientists jump off the deep end
- The dark horse: background
    - 1926 neural spiking discovered at the sensory periphery
        - Adrian and Zotterman (1926)
    - Lord Adrian (isolated single frog muscle nerves) discovered two things
        - Spikes
        - Stimulus-controlled response
    - 1959 - now
        - In sensory cortex, evocative stimulus features are found to drive responses
        (a) Record further downstream (inside brain matter)
            - Visual context (V1)
            - Cat experiment
        (b) Stimulate
            - Hubel & Weisel 1959
    - 1980s - now
        - Upon release from anesthesia and restraint, neurons in sensory cortex are found to fire with **extreme variability** (noise).
        - Remove anesthesia from cat
            - Livingstone & Hubel (1981)
            - Awake
            - ↑ Neurons specific to direction of stimulus (bar moving)
            - Asleep

        - ↑ Spiking present in between stimulus!
        - ↑ Noise
        - ↑ Stimulus that had large effect when awake does not have much effect when asleep
    - Remove restraint
        - Erisken et al. (2014)
        - Self-generated movement
        - Allow animal (rodent) to run (circular treadmill)
        - Ramping spking when mouse is moving spontaneously
    - Review: McGinley & McCormick (2015)
  - This ongoing work is starting to clarify **a different type of independent variable** for explaining activity in the brain
  - Dark horse - unexpected movement

- The elephant in the room
  - 1929 Highlight 1 (alpha rhythm)
    - Hans Berger
    - 1929 - advent of EEG
    - Discovery of the alpha rhythm
    - Experiment on son:
        - Awake, eyes closed
        - But sin-like wave from EEG!!
        - Brain doing this on its own!
    - Papers:
        - Engel & Singer (2001)
        - Buzsaki (2006)
        - Yuste (2015)
  - 1989-2012 Highlight 2 (sharp wave - ripple)
    - The sharp-wave ripple (SWR)
    - Eeg on rats, no stimuli (Buzsaki (1989))
        - Spontaneous bursts of activity (100 ms)
        - 1. Occur in stillness & sleep (no overt stimulus)
        - 2. Originate in deep region (hippocampus)
        - 3. 5-10x increase in firing in hippocampus
        - SWRs drive activity across the brain
    - Oeltermann et al. 2012

	"Outside-in" (Stimulus transformation)	"Inside-out" (Autonomous generation)
Paradigm ("classic") cases	V1 simple cells / classic deep NNs	Alpha rhythm place- or grid-cells
Independent variable	External stimulus	Spontaneous activity (behavioral or neural)
The ideal behavior	Reliable	Natural
Behavioral constraint of the animal?	Essential	Removed

        * Buzsaki 2015
    - So what is this stuff??
    - Nothing obviously comes from the outside, yet spontaneous patterns in the brain are found
    - But from the dark horse story, variance

- The rabbit hole
    - At this point, we can revisit the matter of **dominant paradigm**
    - Our previous 2 stories implied this view
        * Stimulus $\rightarrow$ cognition $\rightarrow$ action
        * **The Outside-In paradigm**
        * The whole neural organism...is, physiologically considered, but a machine for converting stimuli into reactions. - William James (1890)
    - Two paradigms for brain science
        * Outside In (Stimulus transformation)
            · Paradigm (classic) cases
            · ↑ V1 simple cells,
            · ↑ Classic deep NNs
            · Independent variable
            · ↑ External stimulus
            · The ideal behavior
            · ↑ Reliable
            · Behavioral constraint of the animal?
            · ↑ Essential
        * **Inside out** (autonomous generation) - alternative view
            · Paradigm (classic) cases
            · ↑ Alpha rhythm
            · ↑ Place- or grid- cells
            · Independent variable
            · ↑ Spontaneous activity (behavioral or neural)
            · The ideal behavior
            · ↑ natural

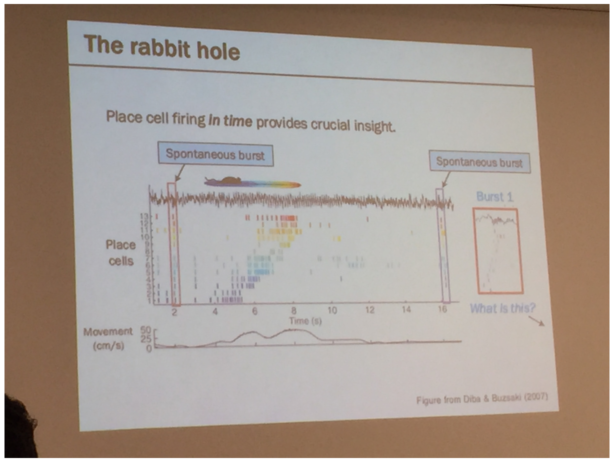
- Behavioral constraint of the animal?
- ↑ Removed ("jump off the deep end)
- ↑ See strange things and not understand whats going on
- Minority of researchers take this paradigm
- ↑ Also researchers that take some middle ground

– Brain scientists **abolish behavioral constraints** (jump off deep end)
  * Three (radical) bases of neural activity identified
  (a) (1969) self-generated movement
    * .. is identified as a basis for neural activity
    * Spontaneous behavior of rat
    * But eeg of neuron reveals that neuron is firing rather regularly
  (b) (1971) an abstract concept (space)
  (c) (2000s) **prediction/simulation**
  * (Why do these sound bizarre or invalid?)
  * Some scientists even reject these observations
  * Possible to synthesize both paradigms
    - E.g., visual cortex stimulus quieting the noise

– 1969 self-generated movement
  * ... is identified as a basis for neural activity
  * Spontaneous behavior of rat
  * But EEG of neuron reveals that neuron is firing rather regularly

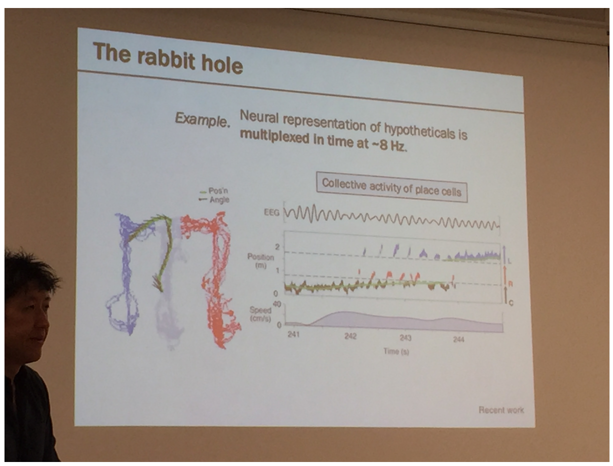
– John OKeefe
  (a) Free behavior in rats
  (b) Single cells
  * → Place cells (1971)
  * Grid cells (2005) → neurons can represent location in space → an abstract concept (space) is identified as a basis for neural activity
  * Over decades the next insight was made by studying these cells **in time**

– Place cell firing **in time** provides crucial insight
  * Diba & Buzsaki (2007) figure
  * Hippocampal replay
  * Two big spontaneous bursts of activity when the animal wasn't moving or receiving stimuli
  * ↑ bursts classically called pre-replay



  - *

- Place cells represent hypothetical paths



  - *
  - Place cell maps → infer → spatial path
  - Question: what does this path represent?
    - Seems spatial paths more concentrated in areas of interest (food, play)
    - BUT, paths/fields present in uninteresting places too
  - Papers:
    - Pfeiffer & Foster (2013)
    - Carr, Jadhav, & Frank (2012)
    - Foster (2017)
- 2000s **prediction/simulation** is clarified as a basis for neural activity
  - Related concepts

        - Psychological
        - Imagination
        - Deliberation
        - Planning
        - (day)dreaming
        - Mind-wandering
        - Prospection
      - Representational
        - Hypotheticals
        - Counterfactuals
        - (???) thought
  - Neural representation of hypotheticals is **multiplexed in time at ~8 Hz** (Kenneth Kay)
    - Rat in maze
    - Place cell 1 - left path in maze
    - Place cell 2 - right path in maze
    - EEG: place cells seem to avoid each other in firing time (i.e., one fires but not at the same time as the other)
  - Collective activity of place cells
    - Segmented (jumps between left arm and right arm)
    - Reflects the traveling wave in hippocampus

1.4 Implications

- The Question - How is X reducible to neural activity?
  - The true implications (of these stories) remain unclear, whether via biological or artificial neural computation
- The problem - a default paradigm
- Implications
  - We need researchers with open and critical minds
- Reconsider neural computation via two alternative paradigms
  - Outside-in vs inside-out
  - O-I
    - Independent var.

            - External input
        - Fundamental repr.
            - The present
        - Essential property of the NN
            - feedforward
    - I-O
        - Independent var.
            - Activity generated within the NN
        - Fundamental repr.
            - Hypotheticals
        - Essential property of the NN
            - Recurrence, re-entrance
    - Rabbit hole properties of neural representation:
        - Depends on internal dynamics (synchronization)
        - Rhythmic
        - Multiplexed

2 Lecture 7

2.1 Recap

- Response to Bengio et al. (Towards Biologically plausible deep learning)
    - Backprop thought of as evolution
    - Feedforward circuit and/or the weights are a phenotype that depends on many genes
    - Minibatch: the collective experience of a generation
    - Selection changes the allele statistics of the next generation through a variant of gradient descent (if allele didn't perform well, fewer percent of next generation will have that allele)
    - **Can learning happen this way?**
- How to measure information in spike trains (a bit array for each Δt, then subtract noise)
- Result: frog auditory neurons respond with much higher entropy to sounds that resemble frog calls than to white noise – despite the fact that the latter has higher entropy

- Summary of dynamical systems:
  - 1D: no cycles, just fixpoints or growth)
    - But have bifurcations: change in parameters qualitatively changes behavior
  - 2D: cycles happen (oscillators)
    - Poincare-Bendixson
    - Because of planarity and logic, flow lines cannot cross (or else we'll have cycles)
      - Control and discipline each other
  - 3D: chaos (Lorenz)
    - $x^{t+1} = ax^t(1 - x^t)$: already chaos
- Against chaos: properties you want your dynamical system to have
  - Linear: can sole closed form: e^{At}
  - 2D: (P-B to the rescue)
  - Conservative systems: conserve energy
  - Reversible systems: they can be run backwards
  - Systems that have a Lyapunov function (progress toward convergence)
- Fundamental theorem of dynamical systems (conley 1984)
  - Redefine cycle in discrete time systems
  - Infinitesimal jumps between steps are allowed (think round-off error).
  - Then the (compact) domain is decomposed into the chain connected components and the transient part; in the latter a Lyapunov function is at work
    - **If you squint a little, chaos goes away**
- The feedforward + recurrent network
  - Two populations of neurons
  - Feedforward and recurrent synaptic connections
  - U, v: vectors of firing rates
  - W, R: matrices of synaptic weights
  - $\tau \cdot dv/dt = -v + F(W\dot{u} + R \cdot v)$
  - F is the response function of the red neurons (recurrent)
  - What if F is linear?
  - W: matrix of synaptic weights

2.2 Hopfield networks

- A discrete-time dynamical system: the hopfield network
- Nodes have two values: $+1$, -1
- Nodes are happy when $\sum_j v_i v_j w_{ij} \geq 0 \geq \Theta_i$
- When there is an unhappy node, flip it

2.3 Cf: the Ising model: all edges are +1

- All are happy
- Algorithm:
  - Repeat
    * Pick a node at random
    * If unhappy, flip
    * If k-happy, flip with probability $e^{-k/T}$
  - Until ferromagnetic (all are same)

2.4 A discrete-time system: Hopfield net (cont)

- Theorem [Hopfield 1982]: dynamical system converges
- at each step total happiness increases by 2 $\times$ the unhappiness of flipped node

2.5 Goal of dynamical systems: pattern completion

- Regions of attraction
- Equilibria

2.6 How do you train a Hopfield net so that it pattern completes?

- Given a set of desired memories $M_1, \ldots, M_m \in \{-1, +1\}$
- Set the weight of edge $a - b$ to $\sum_k M_{ka} \times M_{kb}$
- For every memory k, we bias the weight in the direction "memory k wants to be"
- With n nodes, $0.138n$ random patterns can be stored with probability of erroneous retrieval $< 0.4\%$

2.7 Do Hopfield nets work?

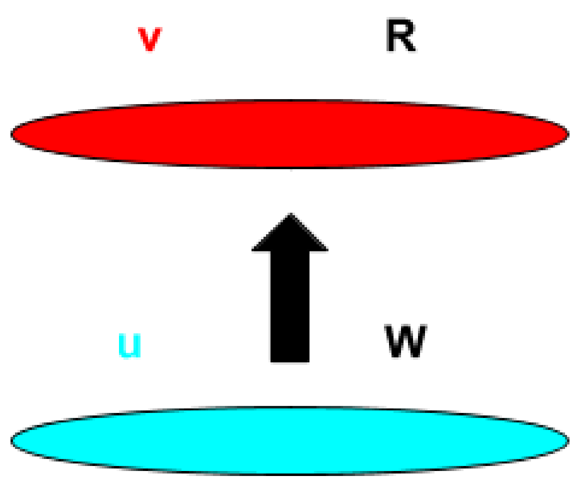
- Not really, because of spurious memories
- Spurious memories: if M is stored, $-M$ is also retrieved
- Also, if M, M, M are stored, so are $\pm M$, $\pm M$, $\pm M$
- Finally, if you store M K times, it will be retrieved L ¿¿ K times more often than other memories

2.8 Boltzmann Machines

- Exactly hopfield nets, but they have hidden nodes, and can be trained
- Visible nodes receive training data by clamping
- After each training data vector is put in, the whole network is left free to run until it reaches "thermal equilibrium"
- Restricted Boltzmann machine: graph is bipartite [Hinton 2005]

2.9 Back to modeling brain networks: the feedforward + recurrent (FFR) network

- $\tau \cdot dv/dt = -v + F(W \cdot u + R \cdot v)$, F is the response of the red neurons



-
- $\tau \cdot dv/dt = h + (M - I) \cdot v$

- Symmetric synaptic matrix M
  - $\rightarrow$ pos. Eigenvalues
  - Symmetry helps (assumed for Boltzmann, assume it here too)

2.10 What is the solution?

- Assume $lambda_k \leq 1$ (exponential growth otherwise)

subsectionSuppose that $\lambda_2 << \lambda_1 \approx 1$

- $(e_1 \cdot h) \cdot e_1$: projection of eigenvector

2.11 Finally, suppose $\lambda_1 = 1$

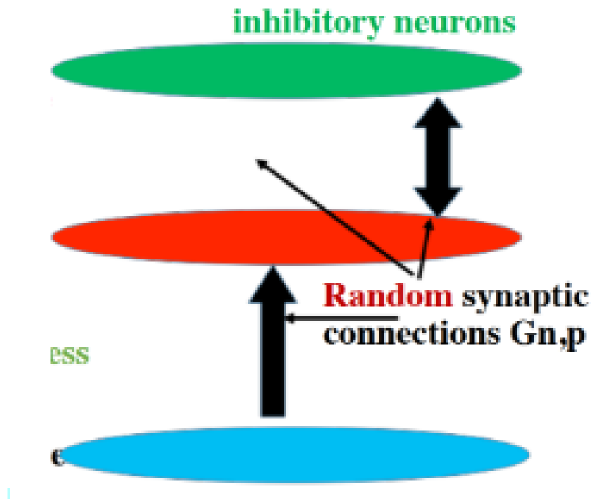
- So the circuit integrates the feedforward inputs projection on e_1
- NB: integration means memory - integration
- A system like this seems to be at work in the brain stem of vertebrates, remembering eye position

2.12 Nonlinear FFR networks

- Cant solve explicitly
- But can simulate
- Can model simple and complex cells

2.13 Modeling complex cells

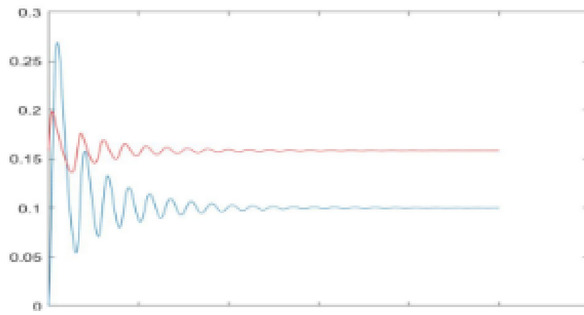
2.14 Another FFR network: excitation-inhibition balance



-
- Gn, p: everything is either present or absent with the same probability
- Blue cells spike, causing red cells to fire, which causes green cells to fire and inhibit red
- then green cells receive less input from red, and they inhibit less
- after 2 or 3 jumps, the excitatory cells will be constant

2.15 E - I balance

- $\tau E\dot{E} = GT(TE, np(E - I), np(1 - p)(E + I)) - E$
- $\tau I\dot{I} = GT(TI, npI, np(1 - p)I) - I$
- If τE is sufficiently larger than τI, an E - I balance will be reached after a few up and down oscillations



-

2.16 How fruit flies remember smells

- Each of the ~ 50 odorant receptor neuron type (ORN) is excited by a different amount, so you have a vector
- Odor $\to$ 50 odorant receptor neuron (ORN) types $\to \sim 50$ projection neurons (PNs) $\to \sim 2000$ Kenyon cells (KCs) $\to$ APL
- APL inhibitory
- Random 100 winners (Kenyon cells) take all
- Is PN $\to$ KC a random bipartite graph? How to know?
  - Columbia experiment
  - Sampled about 200 Kenyon cells and did all possible random tests to show it is random
  - Generated adjacency matrix!
  - This means the genome has a space for random number generation
  - Got complete connectome of fly

2.17 Flys algorithm (RP&C) outperforms LSH (2017)

- But not in 2018