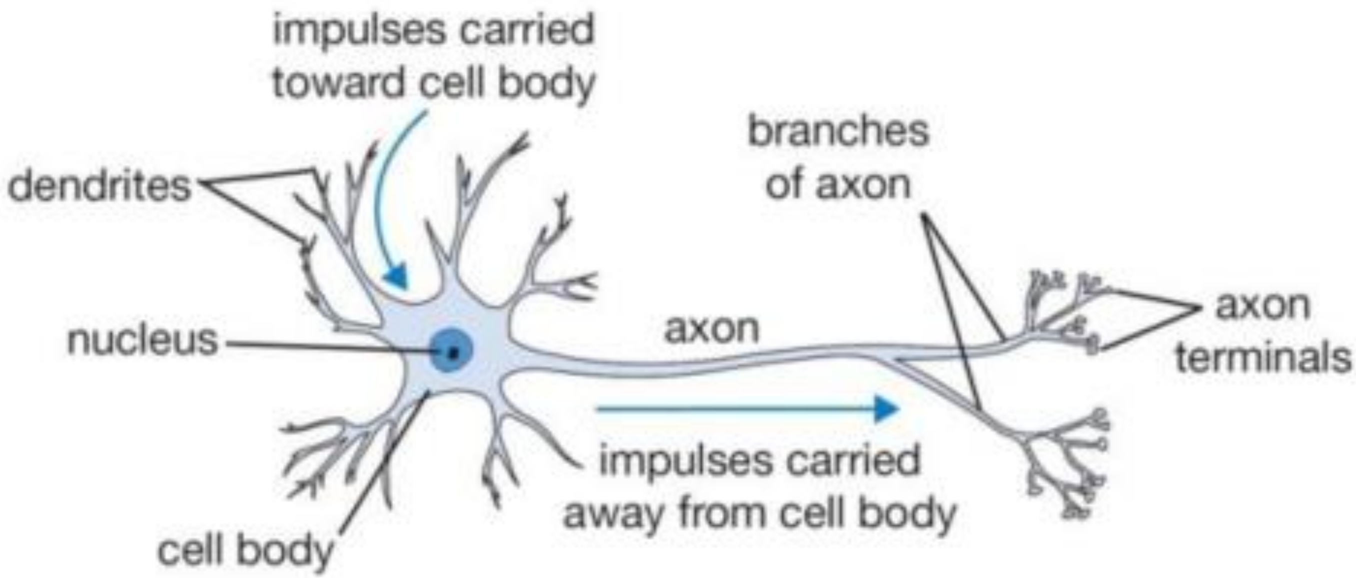
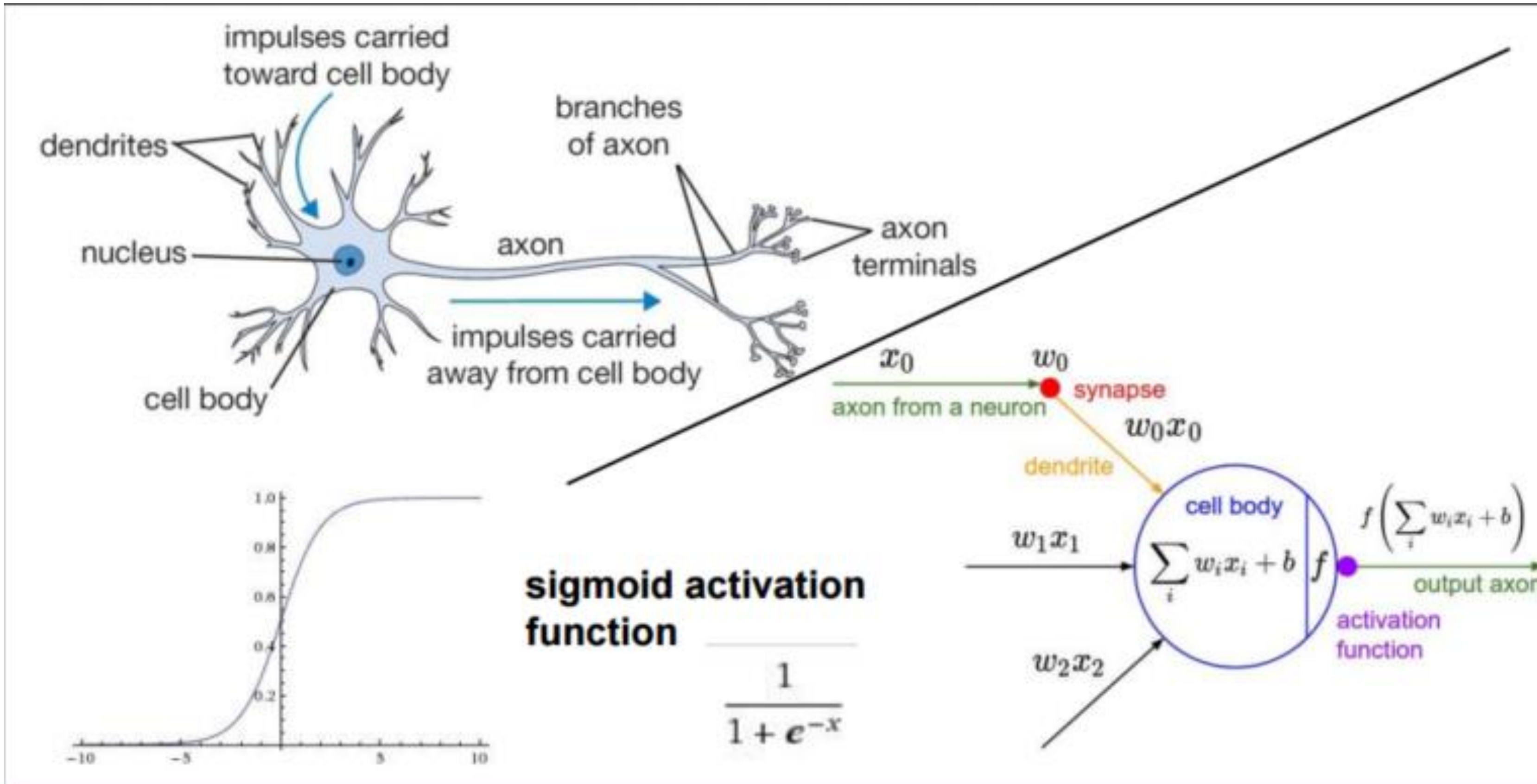
A Bit About Neuroscience



Slides from Geoffrey Hinton, Andrej Karpathy, Wlodzimierz Wojcik, Vedran Lovic

CSC321: Intro to Machine Learning and Neural Networks, Winter 2016

Michael Guerzhoy

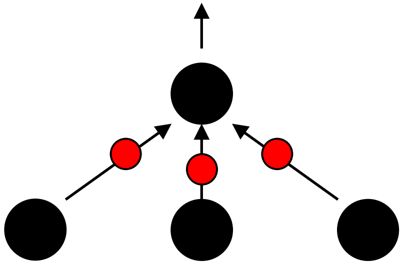


Andrej Karpathy

Only an Analogy

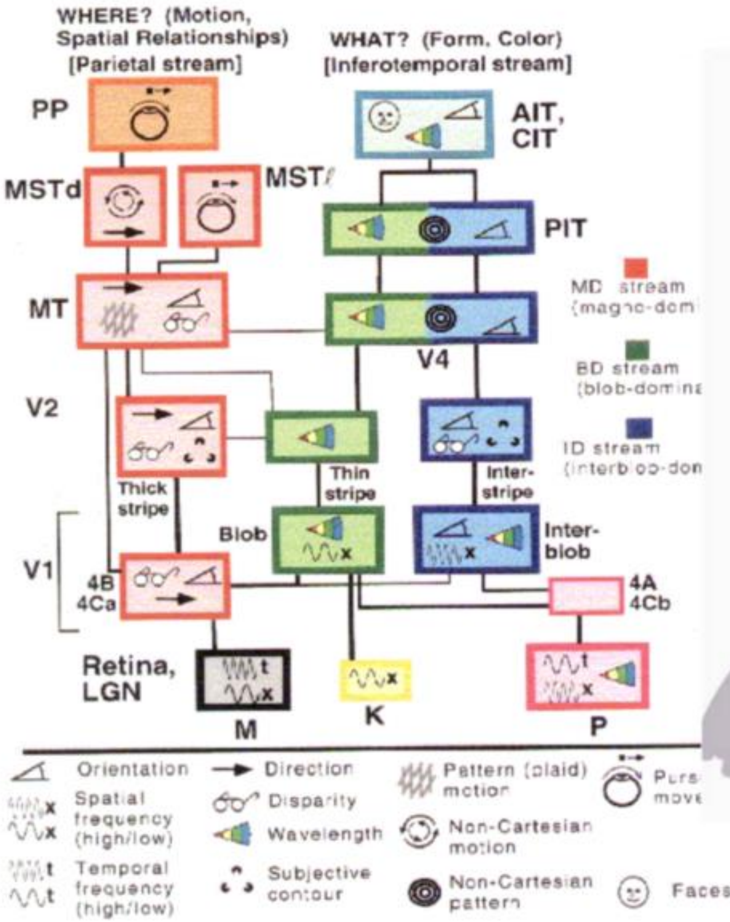
- Many different types of neurons
- The computations are not simply linear combinations of inputs transformed by the same activation functions
- Synapses are more complicated than a single weight
- The neurons don’t output a real number: instead, they “fire” spikes at a (somewhat) regular rate

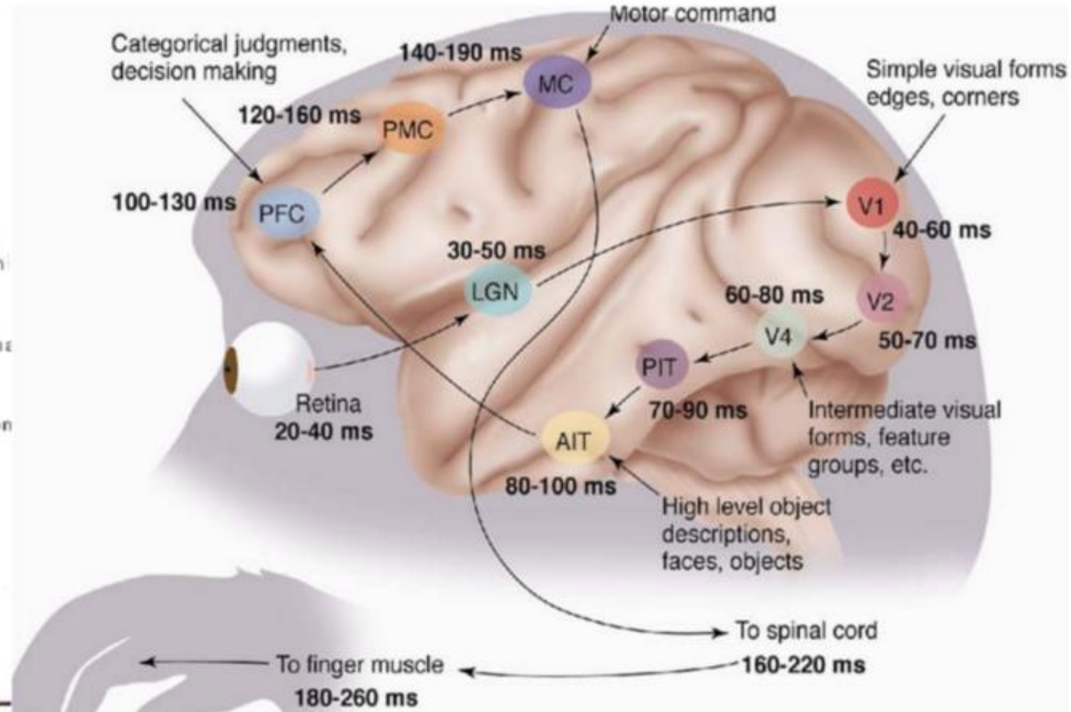
How the Brain Works



- Each neuron receives inputs from other neurons
  - Some neurons also connect to receptors
  - Cortical neurons use spikes to communicate
  - The timing of spikes is important
- The effect of each input line on the neuron is controlled by a synaptic weight
  - The weights can be positive or negative
- The synaptic weights adapt so that the whole network learns to perform useful computations
  - Recognizing objects, understanding language, making plans, controlling the body
- You have about 10^{11} neurons each with about 10^{3} weights
  - A huge number of weights can affect the computation in a very short time. Much better parallelism than a computer.

Visual Cortex





[picture from Simon Thorpe]

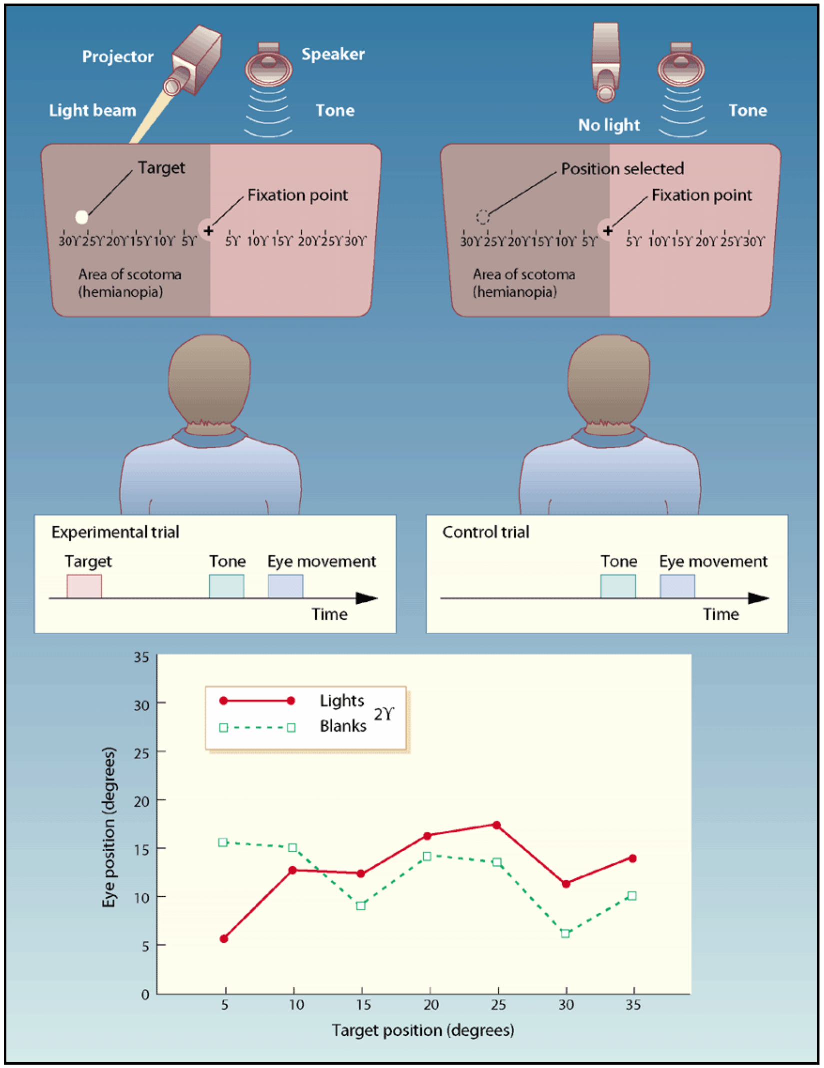
[Gallant & Van Essen]

Modularity and the Brain

- Different bits of the cortex do different things.
  - Local damage to the brain has specific effects
  - Specific tasks increase the blood flow to specific regions.
- But cortex looks pretty much the same all over.
  - Early brain damage makes functions relocate
- Cortex is made of general purpose stuff that has the ability to turn into special purpose hardware in response to experience.
  - This gives rapid parallel computation plus flexibility
  - Conventional computers get flexibility by having stored programs, but this requires very fast central processors to perform large computations.

Blindsight

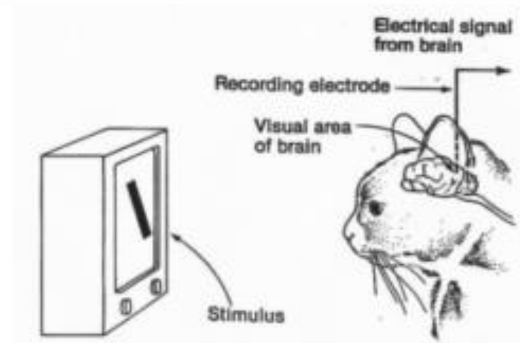
- Case D.B.
- Area around the right calcarine fissure was removed for treatment of angioma
- Reported not seeing anything in the left visual field
- Able to pointing out where the light was in the left visual field
- Blindsight – residual visual abilities within a field defect in the absence of acknowledged awareness

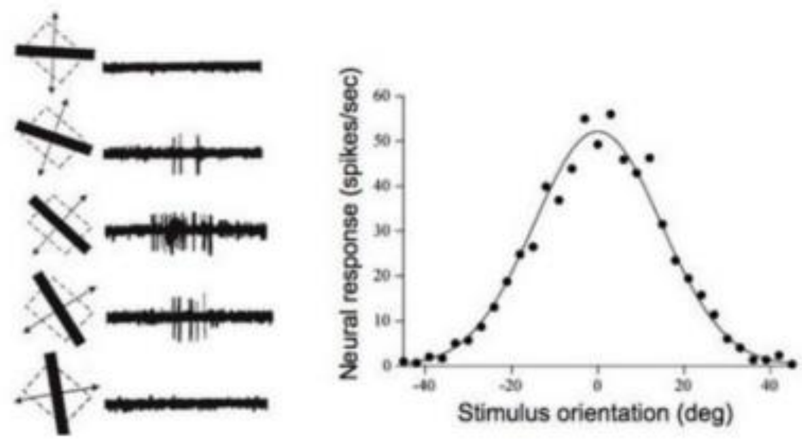


Phineas Gage

What Does V1 do?

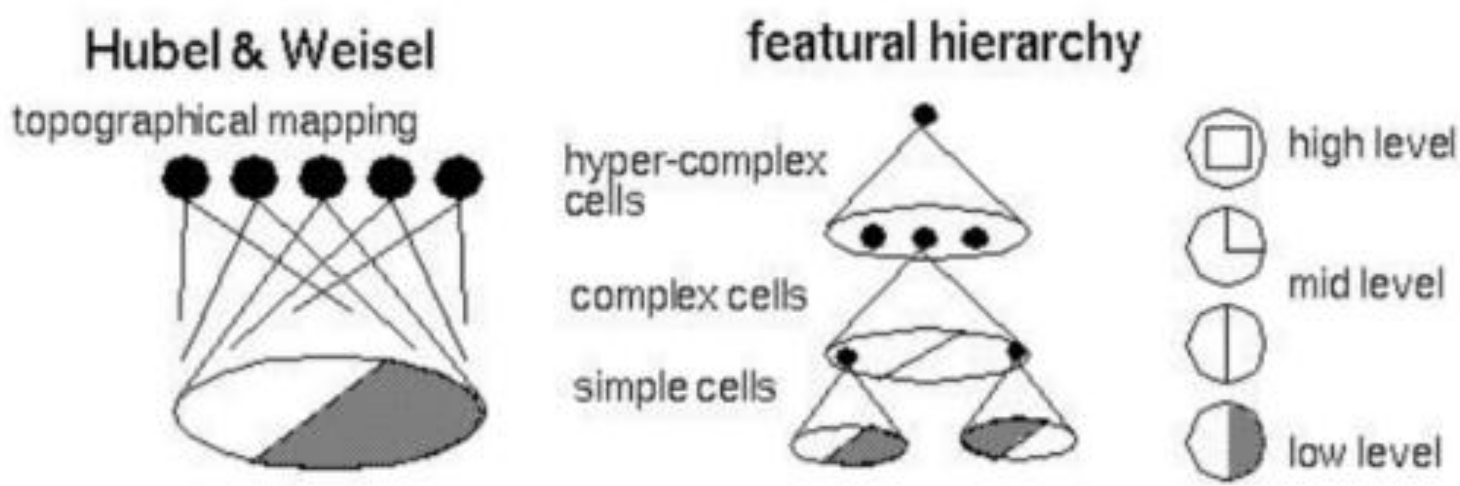
- David Hubel and Torsten Wiesel (Nobel Prize recipients, 1981) showed that individual ("simple cell") neurons in a cat's V1 cortex fire in reaction to seeing lines at a certain angle in a certain location
- Other ("complex cell") neurons fired at lines regardless of orientation





- https://www.youtube.com/watch?v=4nwpU7GFYe8

Hierarchical Organization of Cells

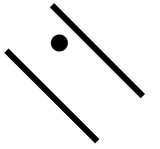


The Invariance Problem

- Our perceptual systems are very good at dealing with invariances
  - translation, rotation, scaling
  - deformation, contrast, lighting, rate
- We are so good at this that its hard to appreciate how difficult it is.
  - Its one of the main difficulties in making computers perceive.
  - We still don’t have generally accepted solutions.

The Invariant Feature Approach

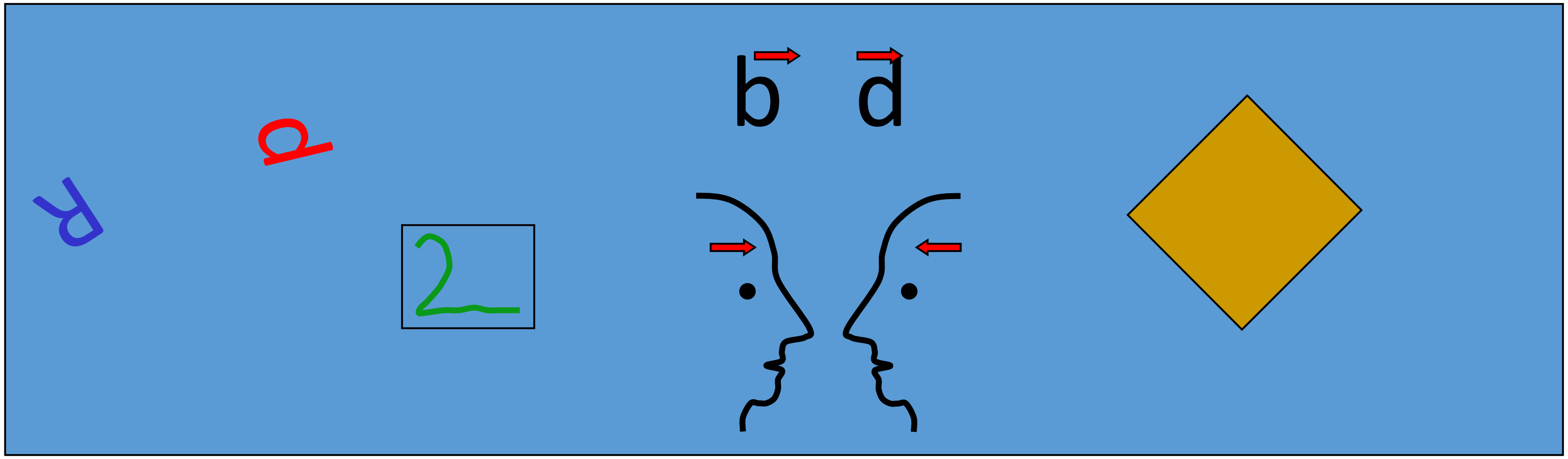
- Extract a large, redundant set of features that are invariant under transformations
  - e.g. “pair of parallel lines with a dot between them.



- With enough of these features, there is only one way to assemble them into an object.

The Normalization Approach

- Do preprocessing to normalize the data
  - e. g. put a box around an object and represent the locations of its pieces relative to this box.
  - Eliminates as many degrees of freedom as the box has.
    - translation, rotation, scale, shear, elongation
  - But its not always easy to choose the box



The Replicated Feature Approach

- Use many different copies of the same feature detector.
  - The copies all have slightly different positions.
  - Could also replicate across scale and orientation.
    - Tricky and expensive
  - Replication reduces the number of free parameters to be learned.
- Use several different feature types, each with its own replicated pool of detectors.
  - Allows each patch of image to be represented in several ways.

The red (and green/blue, respectively) connections all have the same weight.