IT-3708: Subsymbolic AI Methods

- Instructor: Boye Annfelt Høverstad
- Office: Room 361 IT-Vest
- E-mail: hoversta@idi.ntnu.no
- URL: http://www.idi.ntnu.no/emner/it3708/

Course Details

- Sign up for mailing list, and check the website frequently!
- All slides available from the course web site…but not all at once.

- Grading
  - 50% 4 group projects (group size = 1-2)
  - 50% final essay
  - No late work accepted! Always hand in what you have.

- Course Outline
  - Part I: Artificial Neural Networks (ANN)
  - Part II: Evolutionary Algorithms (EA)

- Projects outline
  1. Build an ANN framework, then train and test a specific network type (e.g. Hopfield).
  2. Build an ANN that learns using back-propagation, then use it to solve 2 different problems.
  3. Build an EA and test on a trivial problem.
  4. Combine EA with ANN.

Physical Symbol Systems

- Newell & Simon (1976)
- *A physical symbol system is a machine that produces through time an evolving collection of symbol structures. Such a system exists in a world of objects wider than just these symbolic expressions themselves.*

The Physical Symbol System Hypothesis

- *A physical symbol system has the necessary and sufficient means for general intelligent action.*

PSS => Intelligence
Well…ok…

Intelligence => PSS
Forget it!!

Symbols & The Main Goals of AI

- Engineering: Build intelligent systems
  - Lots of fantastic symbolic AI systems for a multitude of specialized tasks….and many more to come!
  - But general intelligent systems are a major problem, since common sense is hard to represent and reason with symbolically.

- Science: Understand natural intelligence via computers
  - Cognitive Science founded by symbolic AI researchers.
  - But they took the metaphor too far.
    - Organisms clearly compute, but not necessarily:
      - as Von Neumann computers (i.e. serially)
      - as Logic theorem provers (i.e. mathematically complete & consistent)
      - with symbols!! …We can interpret the reasoning process as symbolic, but the underlying mechanism may not be. The ends don’t explain the means!

Symbols & Common Sense

Qualitative Physics (1970's - ?) - represent everyday physical concepts in a qualitative symbolic manner so that common-sense physical reasoning can be automated using symbolic AI inference methods.

"Naïve Physics I: Ontology of Liquids" (Hayes, 1979)

movement(h) => arriving(in(h)) ^ leaving(out(h)) ^ inward(in(h),h) ^ inward(out(h),h) ^ height(in(h)) > height(out(h)) ^ direction(h) = (in(h),out(h))

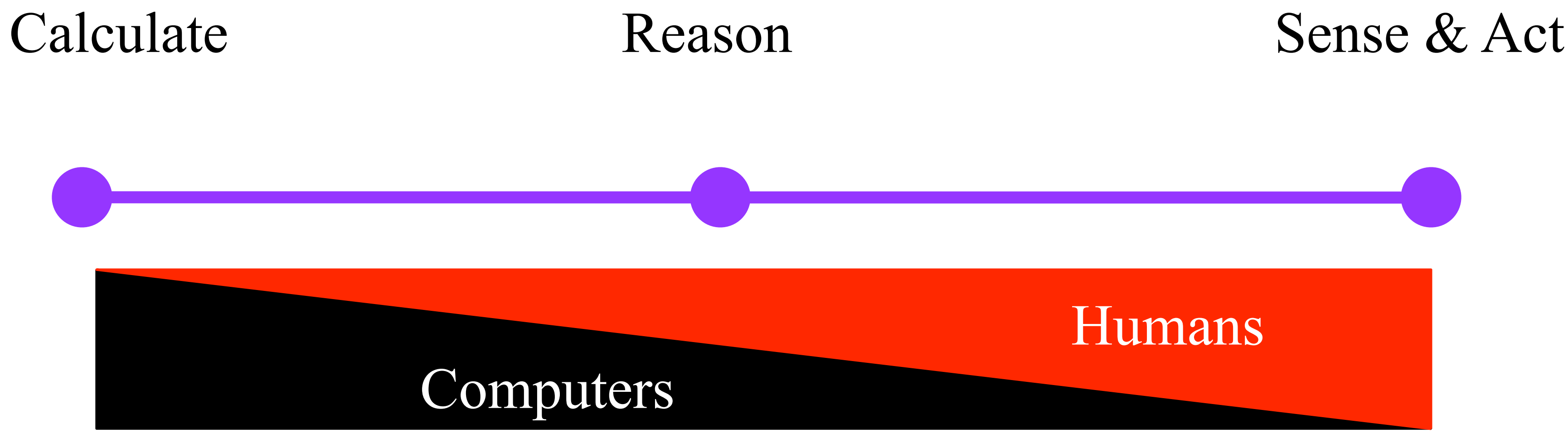
Is this common-sense axiomatization (in any symbolic language) useful for knowledge engineers?

…Maybe

Does the brain really represent information this way? … Probably not

The Intelligence Spectrum

- *Robot* (Moravec ,1999)



<u>On the fringes:</u>

- Humans are slow, error-prone calculators.
- Robots sense and act no better (and much slower) than frogs.

<u>The battle for the middle ground:</u>

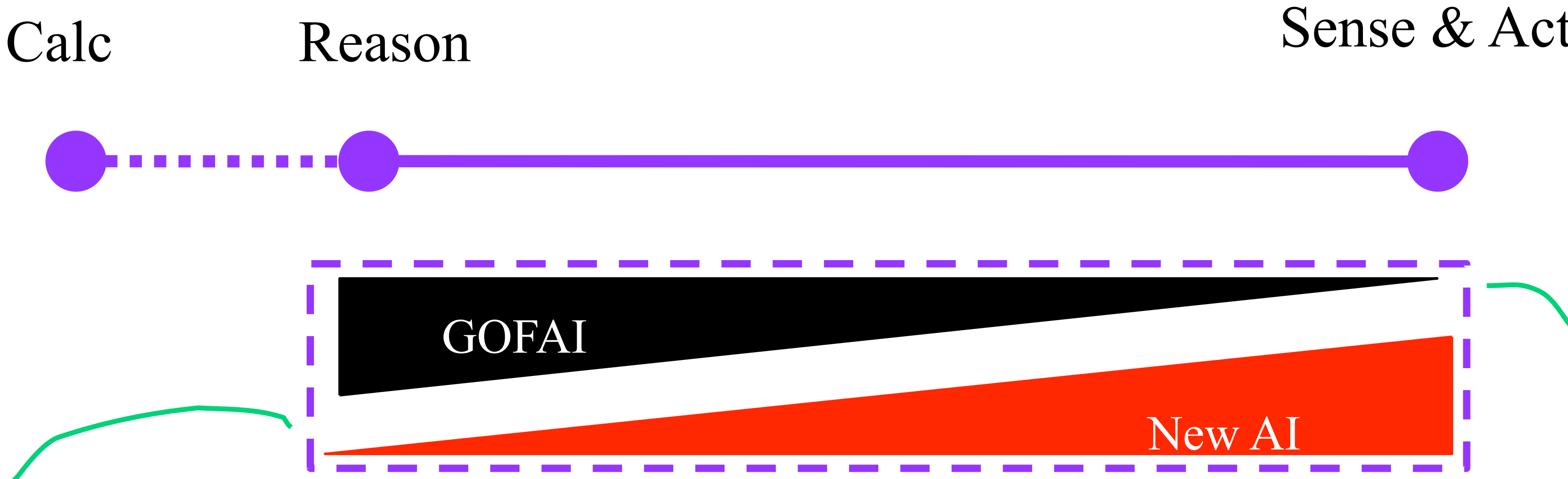
- Deep Blue beat the best human chess player.
- But minimax search ≠ “reasoning”.
- Should we care?

Evolutionary Progressions along the Intelligence Spectrum

	Living organisms	Computers
Sense & Act:	**10,000,000**+ years.	15+ years
Reason:	**100,000**+ years.	30+ years
Calculate:	**1,000**+ years	50+ years

- Evolution of reasoning was tightly constrained and influenced by sensorimotor capabilities. Else extinction!
- GOFAI systems are often **in their own little worlds,** making unreasonable assumptions about independent sensorimotor apparatus.
- To achieve AI's scientific goal of understanding human intelligence, the road from sense-and-act to reasoning via simulated evolution may be the only way.
- But, to achieve AI's engineering goals, both approaches seem important. E.g. Deep Blue (minimax search) for chess, Samuel's -vs- Blondie-24 in checkers, etc.

GOFAI -vs- The New AI



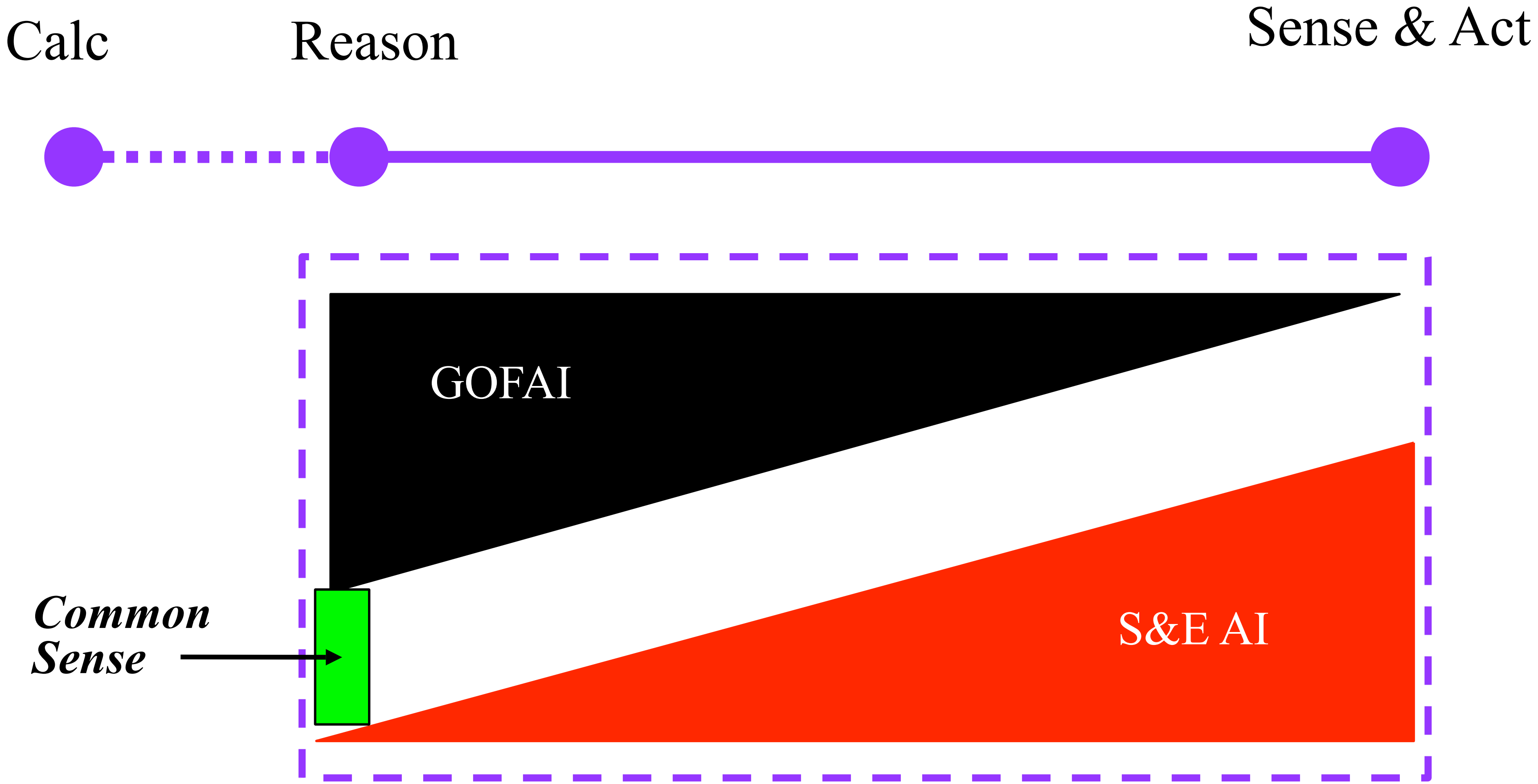
GOFAI

- Disembodied reasoning systems can't plug-and-play on robots.
- Lack of common sense => no general human reasoning abilities.

New AI

- Embodied S&A gives basis for common sense but has not yet scaled up to sophisticated human-like abstract reasoning.

GOFAI -vs- Situated & Embodied AI

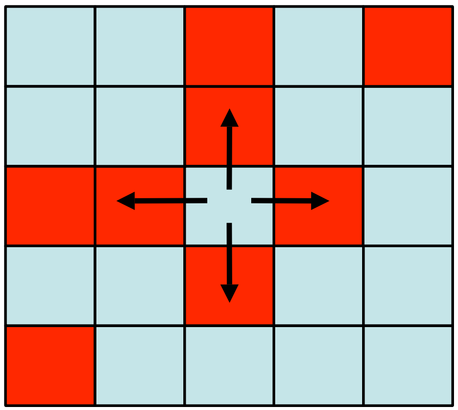


The Situated & Embodied AI Hypothesis

- Complex intelligence is better understood and more successfully embodied in artifacts by working up from low-level sensory-motor agents than from abstract cognitive mechanisms of rationality (e.g. logic, means-ends analysis, etc.).
- **Cognitive Incrementalism:** Cognition (and hence common sense) is an extension of sensorimotor behavior.
- Brooks, Steels, Pfeifer, Scheier, Beer, Nolfi, Floreano…

The Artificial Life Approach to AI

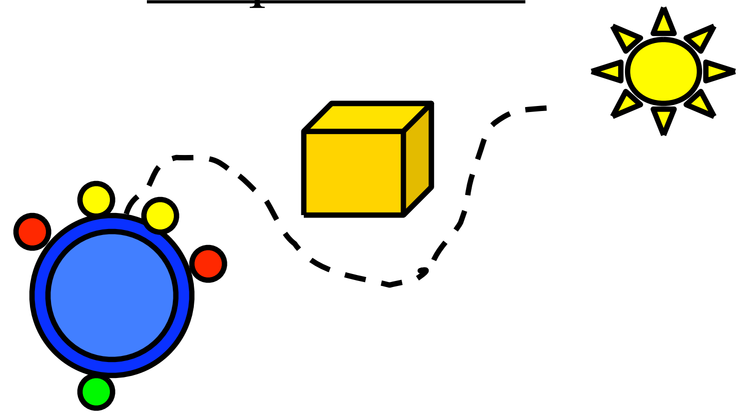
Cellular Automata



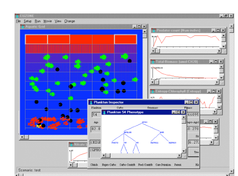
```
  2 2 2 2 2 2 2 2
2 1 7 0 1 4 0 1 4 2
2 0 2 2 2 2 2 2 0 2
2 7 2         2 1 2
2 1 2         2 1 2
2 0 2         2 1 2
2 7 2         2 1 2
2 1 2 2 2 2 2 2 1 2 2 2 2 2
2 0 7 1 0 7 1 0 7 1 1 1 1 1
  2 2 2 2 2 2 2 2 2 2 2 2 2
```

Langton Loop

Simple Robots



Simulated Real Worlds



Computer Memories

MOV_CD
PUSH_AX
INC_C
ADRB
JMPB

Properties of Alife Systems

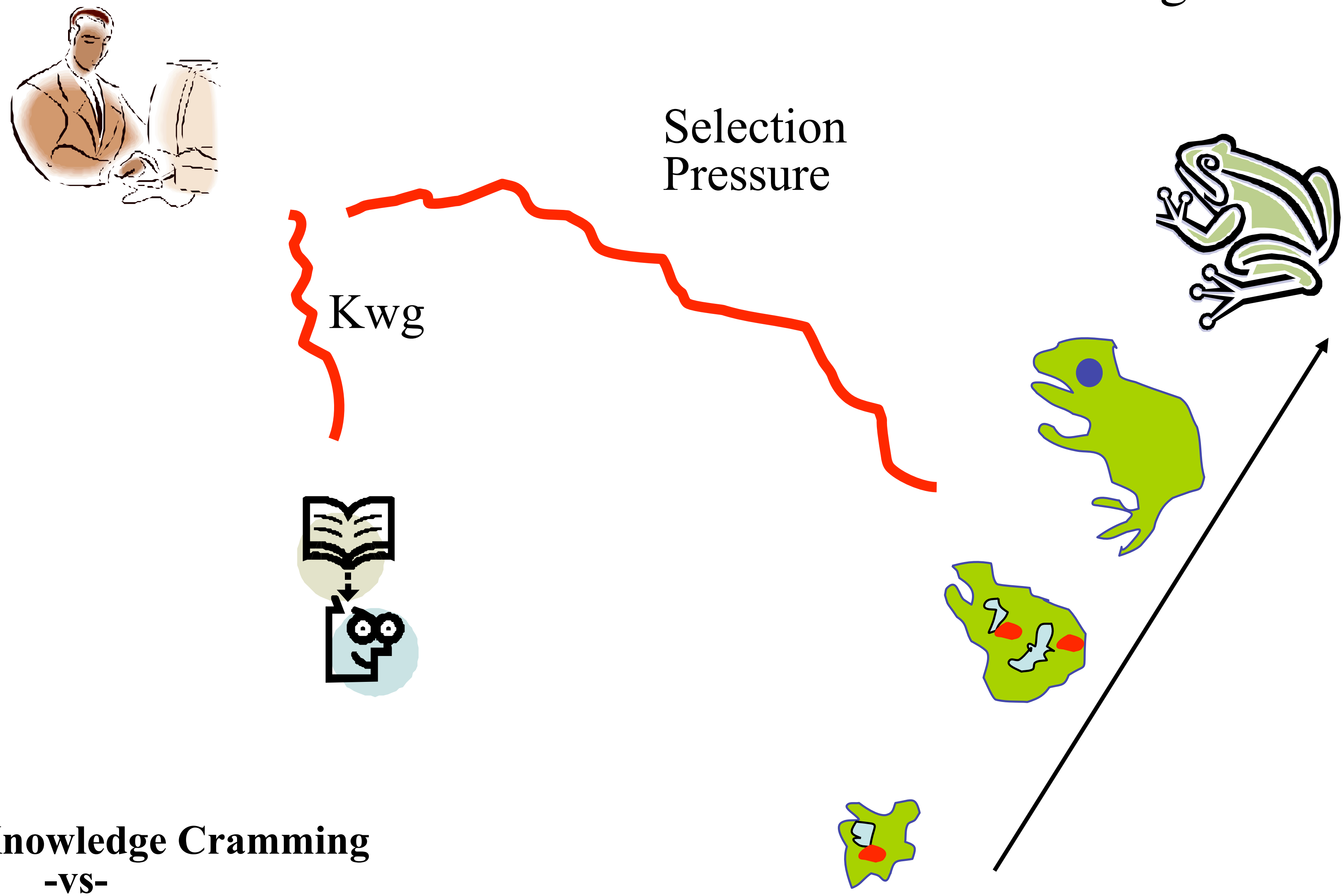
- Synthetic: Bottom-up, multiple interacting agents
- Self-Organizing: Global structure is emergent.
- Self-Regulating: No global/centralized control.
- Adaptive: Learning and/or evolving
- Complex: On the edge of chaos; dissipative

Adaptation

- Key focus of Situated & Embodied AI (I.e., Alife AI)
  - But now, often at level of simple organisms (ants, flies, frogs, etc.)
- Machine Learning (ML) is also a key part of GOFAI.
- Alife AI is very interested in subsymbolic ML techniques:
  - Artificial Neural Networks (ANNs)
  - Evolutionary Algorithms (EAs)

- <u>Learning:</u> agents modify their own behavior (normally to improve performance) in their lifetime.
- <u>Evolution:</u> populations of agents change their behavior over the course of many generations.
- <u>Both:</u> Evolving populations of learning agents
  - Classifier Systems
  - EA + ANN * (it-3708)
  - EA + Reinforcement Learning

AI State of the Art: Nerds or Well-Rounded Frogs?



Knowledge Cramming
-vs-
Adaptive (Learning & Evolving) Systems

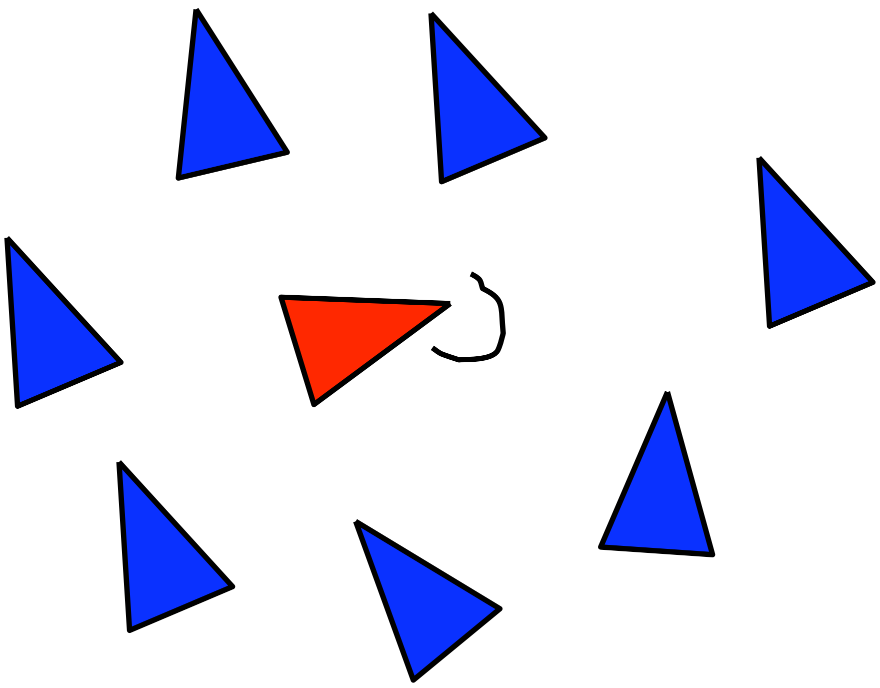
The Low Road to Intelligence

Why the ALife approach to AI is worth pursuing?

3. Intelligence (rationality, intentionality, cognition) are (often only) in the eye of the **observer**.

5. Mind, body and environment are very tightly coupled, with cognition built on top of the sensorimotor apparatus. Sensing and acting come first (in both evolution and human development), so our understanding of cognition is enhanced by knowing how it arises from and interacts with sensing and acting.

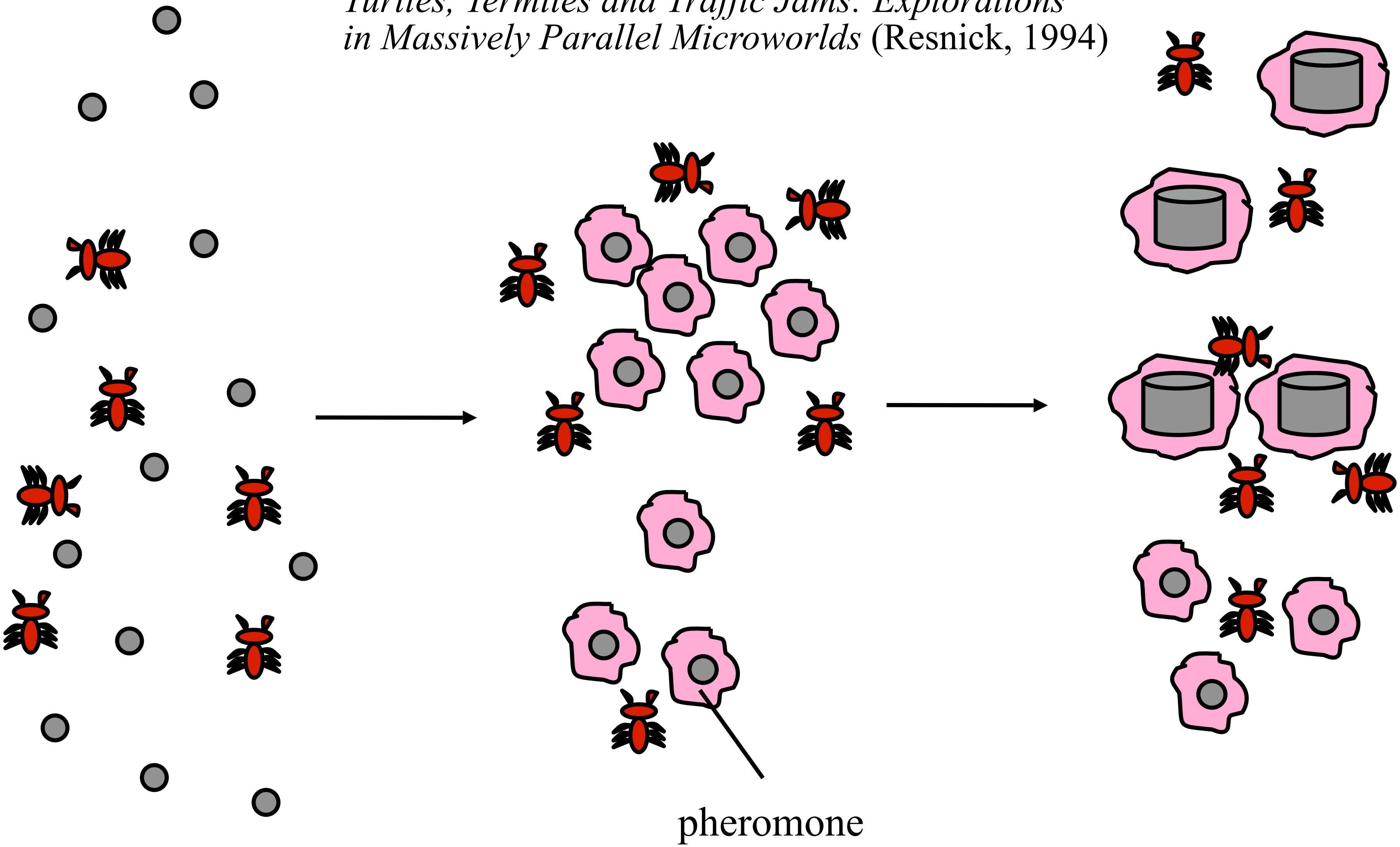
BOIDS



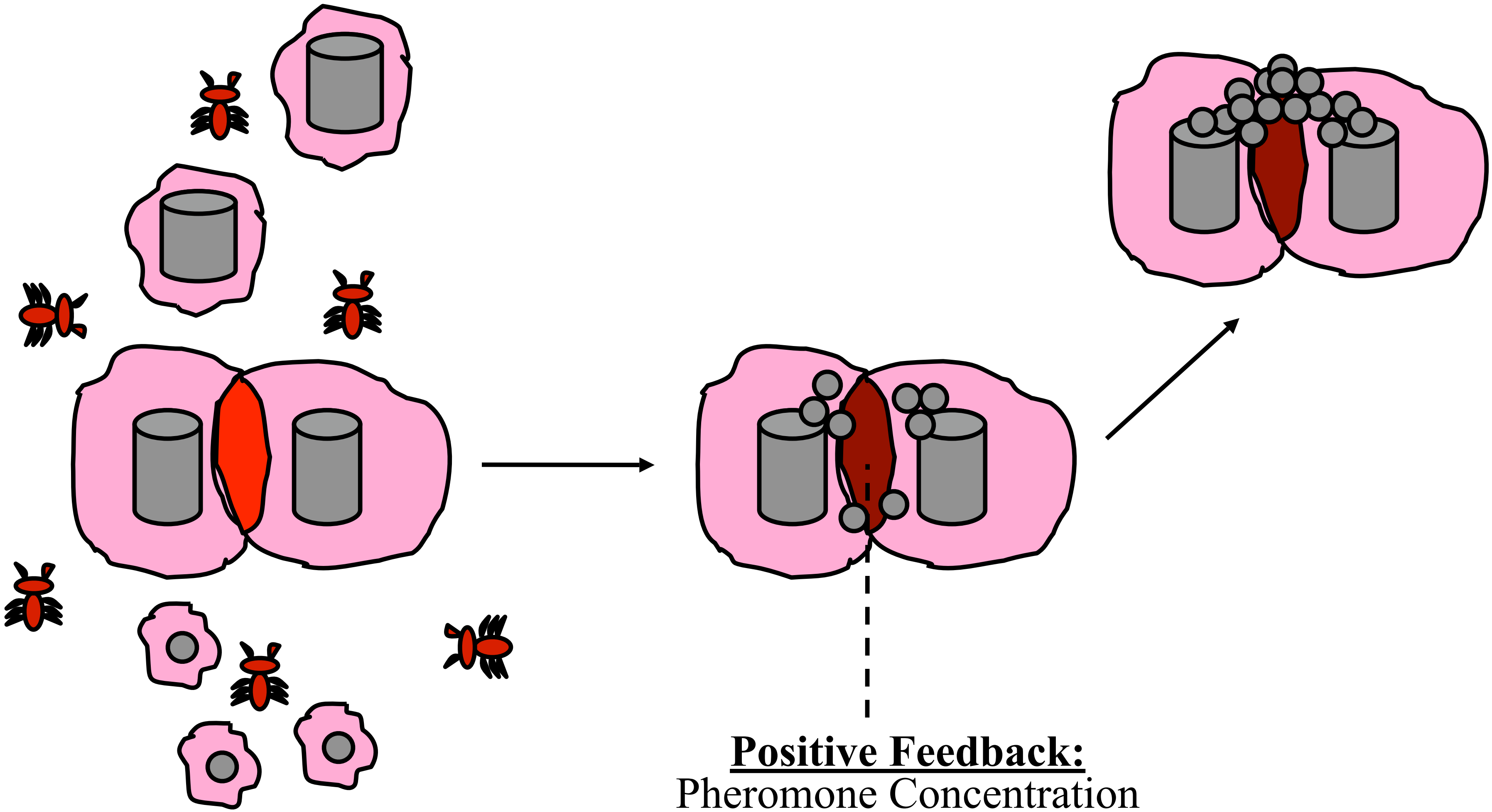
- http://www.red3d.com/cwr/boids/

Termite Arch-Building (Stigmergy)

Turtles, Termites and Traffic Jams: Explorations in Massively Parallel Microworlds (Resnick, 1994)



Columns to Arches



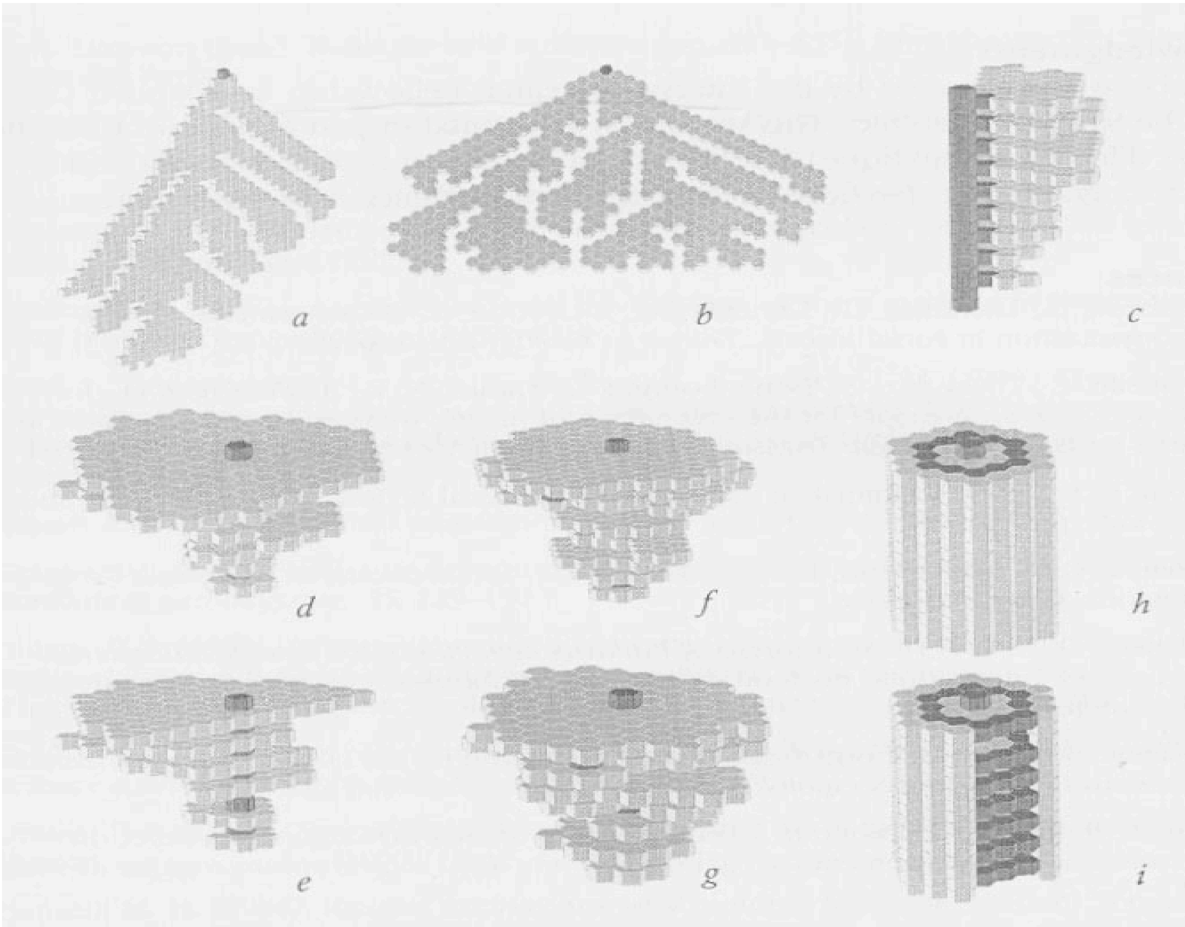
Positive Feedback:
Pheromone Concentration
in middle gets higher and higher
as more dirt balls are added.

Alite and S&E Examples

Emergent Nest Building

Simulated Wasps (Theraulaz & Bonabeau, 1999)
A few dozen simple behavioral rules:

Add a new cell at a 3-walled site location with prob = .25



Dumb Cricket? Smart Cricket? *(Clark, 2001)*

Fact:

Female crickets (somehow) ”recognize” the chirping songs of males in their species and move toward them. They do not move toward males of other species.

Classic Symbolic AI explanation:

Female cricket must:

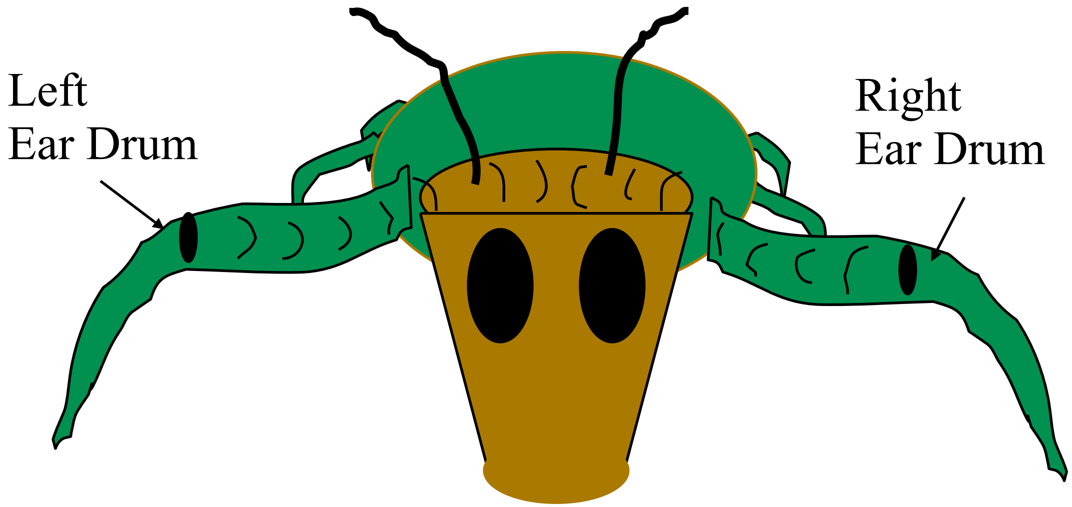
1. Hear songs and classify them as *same*- or *different*-species.
2. Determine the location of the conspecific male.
3. Move toward that male.
4. Maybe even **plan** a route to that male before moving.

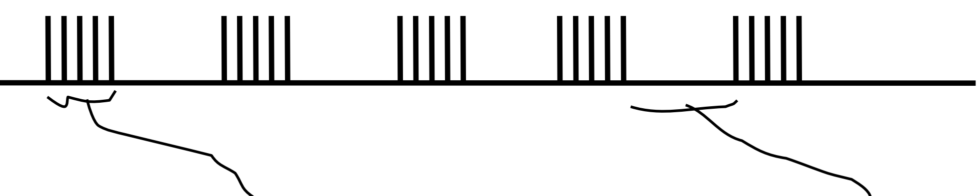
Situated and Embodied AI explanation = Biological explanation

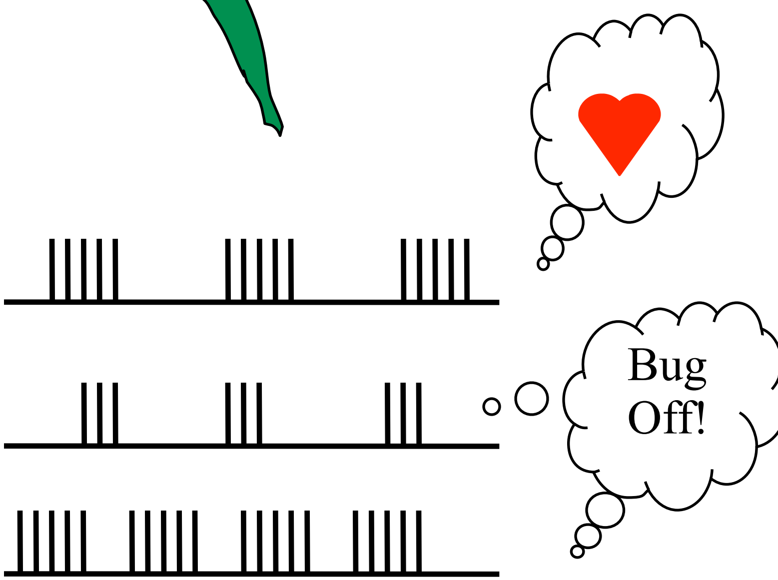
The physical interactions between the song (environment), auditory canal (body) and relatively simple neural circuitry (brain) of the cricket lead to approach behavior in only those special cases where the song of the male is compatible with the body and brain of the female.

Cricket Phonotaxis

- Webb, B. (2001). Biorobotics: Methods & Applications, Ch. 1.
- Female Crickets only respond to songs with particular **<u>carrier frequencies</u>** and **<u>syllable durations</u>**.



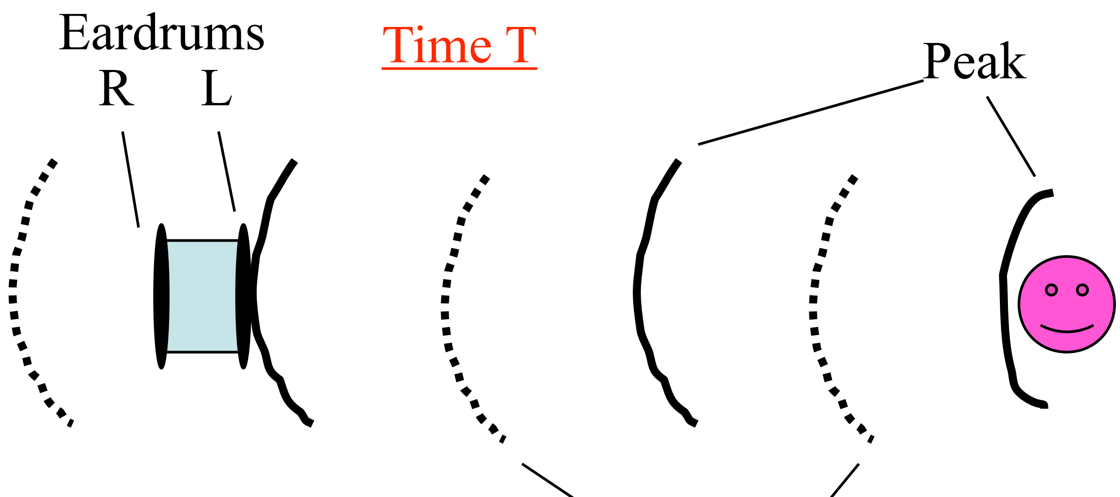


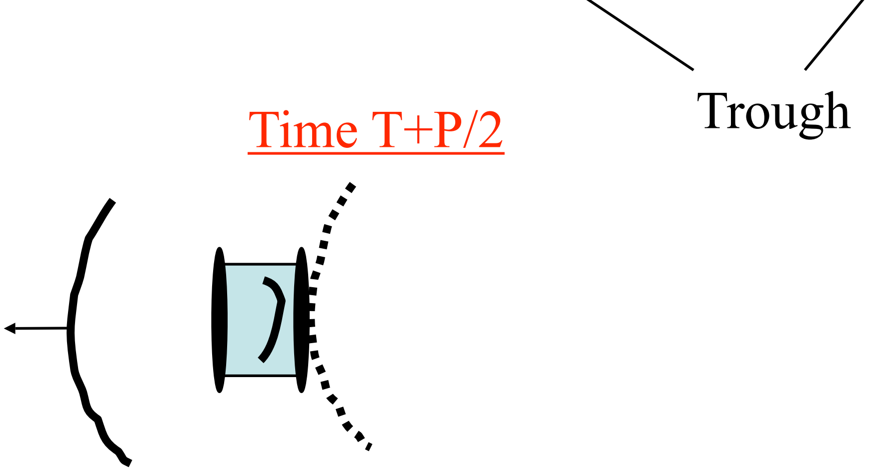


- Syllable Duration
- Carrying Frequency = 1/Inter-syllable period

Preferred Carrier Frequency

Distance between the two ear-drums is the critical determinant. If it's ONE QUARTER the song's inter-syllable wavelength, then the eardrums vibrate most strongly. Here P = period of the sound wave.

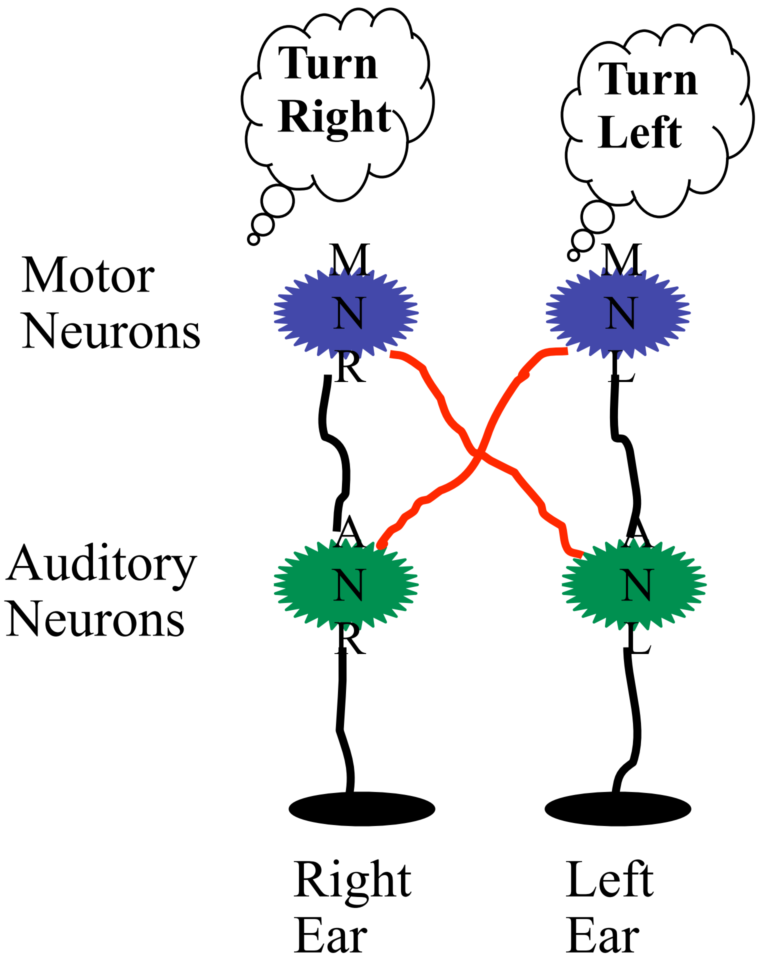




- From T to T+P/4, the peak travels across the body and meets the right eardrum, causing it to vibrate, thus generating a new peak.
- From T+P/4 to T+P/2, the new peak travels exactly 1/4 wavelength = ear-to-ear distance.
- At time T+P/2, the left ear has a) a trough on the outside, and b) a peak on the inside.
- That's a max pressure difference => the eardrum is maximally stimulated.
- The cricket is happy!!

Preferred Syllable Duration

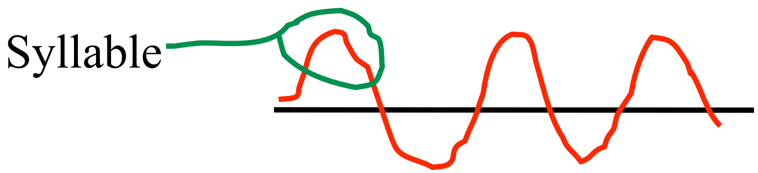
- Appears to be determined in the brain, but details only partially known.
- Biorobotics researchers (Webb et. al.) provide minimal ANNs that are sufficient explanations.



- Each auditory neuron stimulates the corresponding motor neuron and inhibits the opposite motor neuron.
- Each of the 4 neurons has a very detailed (but standard) model: ***leaky integrate-and-fire***
- AN => MN synapses are temporarily depressed after the AN fires

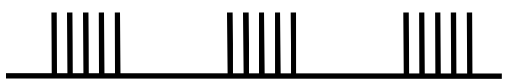
Preferred Syllable Duration

- Assume a stimulus on the left side of the cricket.
- High frequency (short wavelength) sound has a quickly-decaying amplitude with distance, so the left ear gets a stronger signal than the right.



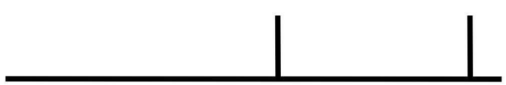
Incoming sound

- Neuron ANL integrates the inputs from the left ear drum and fires groups of pulses with durations = syllable durations.



ANL Response

- This inhibits motor neuron MNR but stimulates MNL, which integrates the inputs from ANL and eventually begins to fire. However, it integrates more slowly than ANL and therefore fires less frequently.

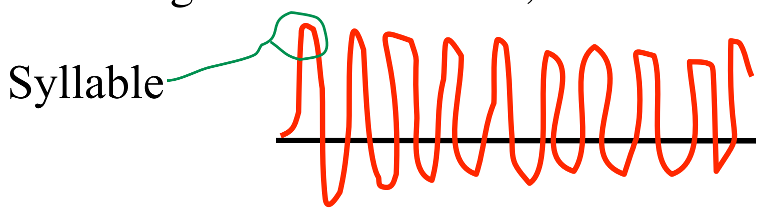


MNL Response

- The cricket turns left. It is attracted to the song.

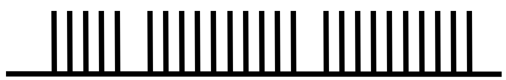
Null Poeng

- Stimulus again from left side, but now the syllables are very short and frequent..



Incoming sound

- Neuron ANL integrates the inputs from the left ear drum and fires constantly, with very few significant gaps.



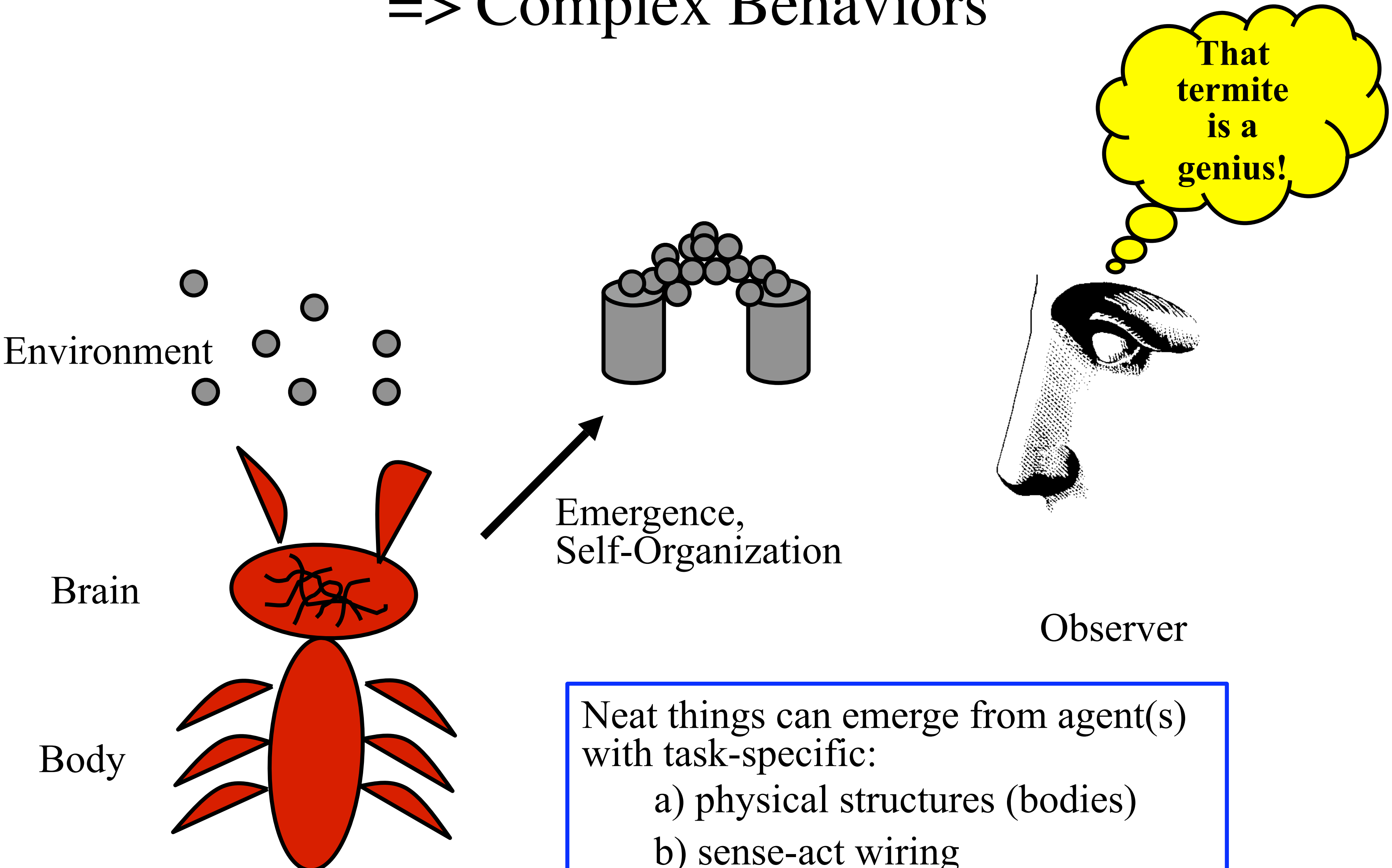
ANL Response

- This inhibits motor neuron MNR and stimulates MNL.
- But, now the ANL-MNL synapse **<u>habituates</u>** due to the constant firing of ANL (and hence no break in which to regain strength).
- So the signals that ANL sends to MNL are WEAK, and MNL never integrates enough charge to fire.

MNL Response

- The cricket is not interested.

Complex Worlds & Bodies + Simple Brains => Complex Behaviors



Neat things can emerge from agent(s) with task-specific:

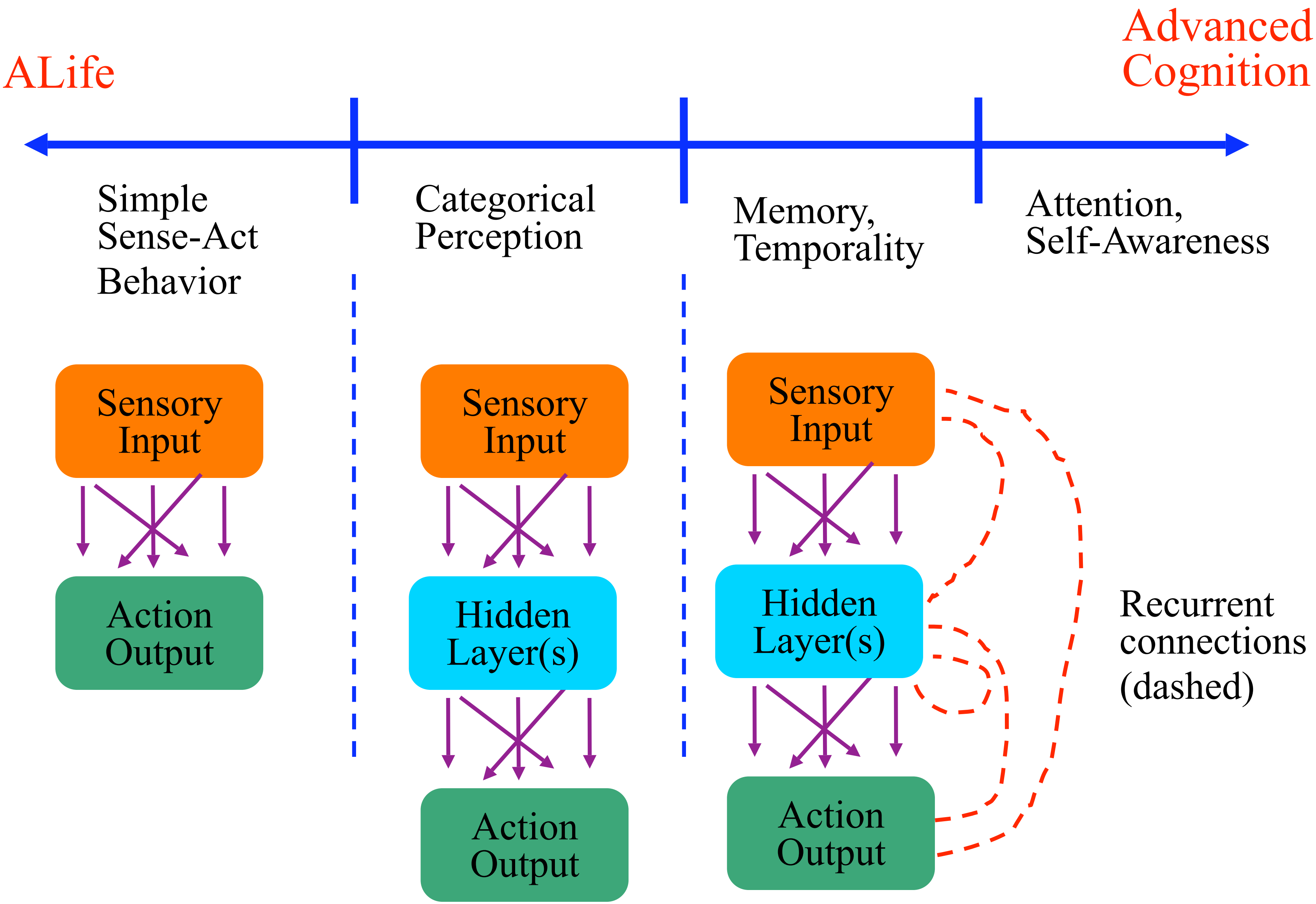
a) physical structures (bodies)

b) sense-act wiring

Brains of ALife Agents

- Sets of If-Then rules
- Logical Networks
- Lisp-like programs
- Artificial Neural Networks (ANNs)
  - Very general; much less biased than symbolic representations
  - Loosely inspired by biology
  - 3-way adaptive (POE)
    - evolution (phylogenesis)
    - development (ontogenesis)
    - learning (epigenesis)

Intelligence Spectrum & ANN Architectures



What's the Right Net for the Job??

- Hand-coding of **topology** (i.e. layers, nodes and arcs) and **weights** is extremely difficult except for very simple networks.
- Normal approach = hand-code topology + use machine-learning to find the weights.
  - 2-layered feed-forward nets: delta rule
  - n-layered Ffwd nets: backpropagation (generalized delta rule)
  - n-layered recurrent nets: backprop through time (BPTT)
- Problems
  - Backprop often gets caught in local minima
  - BPTT is very hairy
  - Neither BP nor BPTT works in typical ALife situations, where there is not a user-defined correct action for each sensory input.
- Popular Solution
  - Use evolutionary algorithms (EAs) to evolve the weights.
  - Bonus! You can also use EAs to evolve topologies.

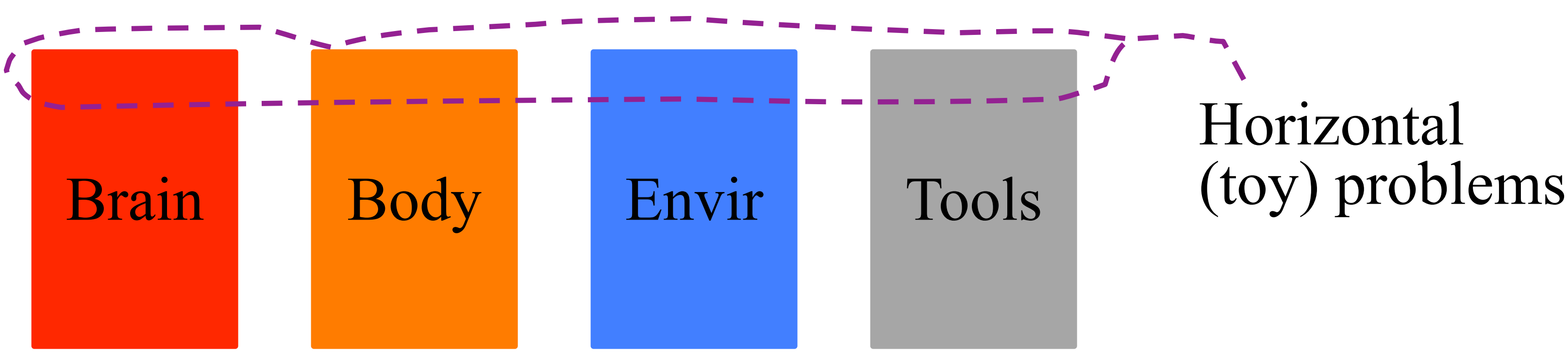
Minimally Cognitive Behaviors

Randall Beer (1996++)

- Categorical Perception
- Focusing Attention
- Discriminating self from non-self

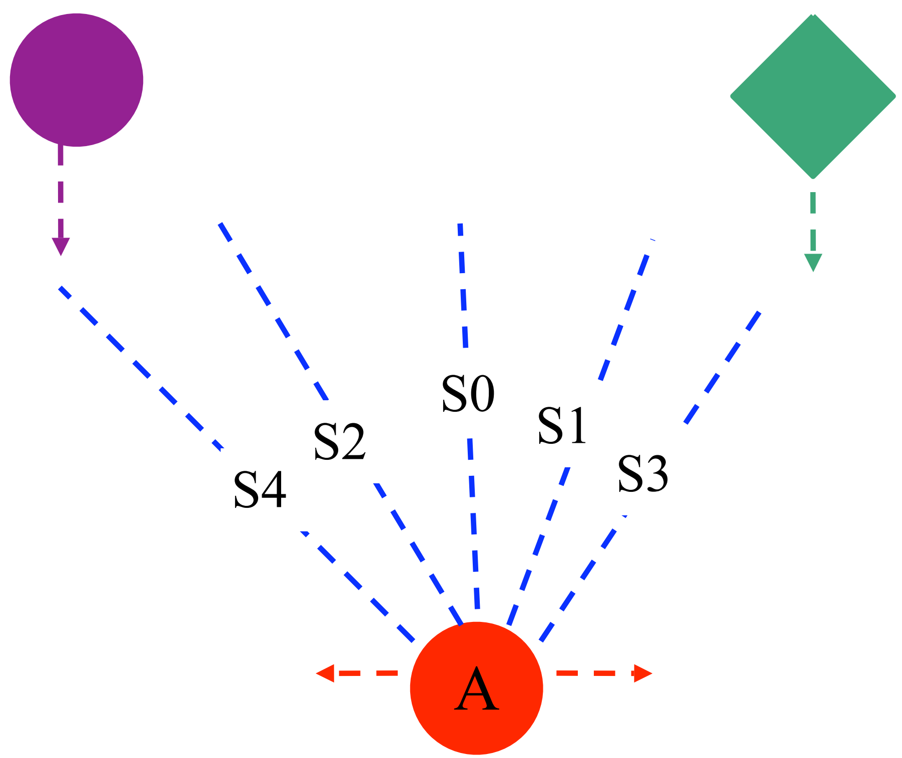
Look for links between sensorimotor behavior and these primitive cognitive acts (in nature and in simulations).

Synthetic approach to testing cognitive incrementalism.



Perceptual Categorization

- Beer (1996 + 2004)
- Agent with brain = recursive ANN must move to falling circles and flee from falling diamonds.
- The ANN topology is hand-designed, but the weights are evolved using a genetic algorithm.



Categories

Classic Objective (Philosophy, Symbolic AI)

A circle is all points in a 2-dimensional plane that are equidistant from a center common center point.

Situated & Embodied:

A circle is any object that, when the agent stands directly under its center point, causes sensor activations of the form:

$S0 > S1 = S2 > S3 = S4$

where $S0 - S1 < S1 - S3$ and $S0 - S2 < S2 - S4$

Sensorimotor (very subjective, often history dependent)

A circle is any object that invokes the following sequence of sensory inputs and motor outputs in me (the particular agent):

1) Move: right ; Sense: S0 = ... S5 = ..
2) Move: right; Sense: S0 = ...S5 = ..
3) Move: left; Sense: S0 = ... S5 = ..

Beer's agents formed these types of classes

(Weak) AI Evidence for Cognitive Incrementalism

- Active Categorical Perception (Randall Beer, 1996 + 2004)
  - Categorical Perception = *minimally cognitive behavior*
  - Differential classification based on dynamics of interaction between environment, agent, and its ANN.
  - **Internal state** is key to nonreactive behavior, not **representations**.

- Darwin V (Almassy, Sporns, Edelman; 2000)
  - Foraging robot, with (very complex) neural-net brain.
  - Learns to classify food and poison.
  - However, when robot cannot move but is presented with sensory inputs that are identical to those in the mobile scenarios, it fails to learn food/poison.
  - Motricity is essential for concept formation (in Darwin V)!

- Pfeifer & Scheier (1999), Nolfi & Floreano (2000) and others

Bio + Psych Evidence for Cognitive Incrementalism

- Inverted Prism Glasses - only subjects who are allowed to **move** can adapt to wearing them.
- Much of our abstract reasoning involves metaphors grounded in our bodies and the environment (*Where Mathematics Comes From*, Lakoff & Nunez, 2000)
- Neurobiological evidence (final lecture for it3708)
  - Classic “motor” brain regions such as cerebellum and basal ganglia are also very active during cognitive activities.
  - The prefrontal cortex is the highest level of both symbol processing (we DO do this, just not like theorem provers!) and motor control.

Where's the AI (in EAs and ANNs)?

Type I AI: Understanding Organic Intelligence

- Sensorimotor Behavior is the evolutionary basis for higher (cognitive) activities.

- ANNs are very effective model brains for a wide range of sensorimotor agents.

- EAs are one of the best techniques for designing and tuning ANNs in non-supervised tasks (such as ALife simulations).

- Plus, simulated evolution of these agents in general (regardless of their brain models) can give insights into how intelligence evolved (and hence into some of the *whats* and *whys* of intelligence).

Where's the AI?

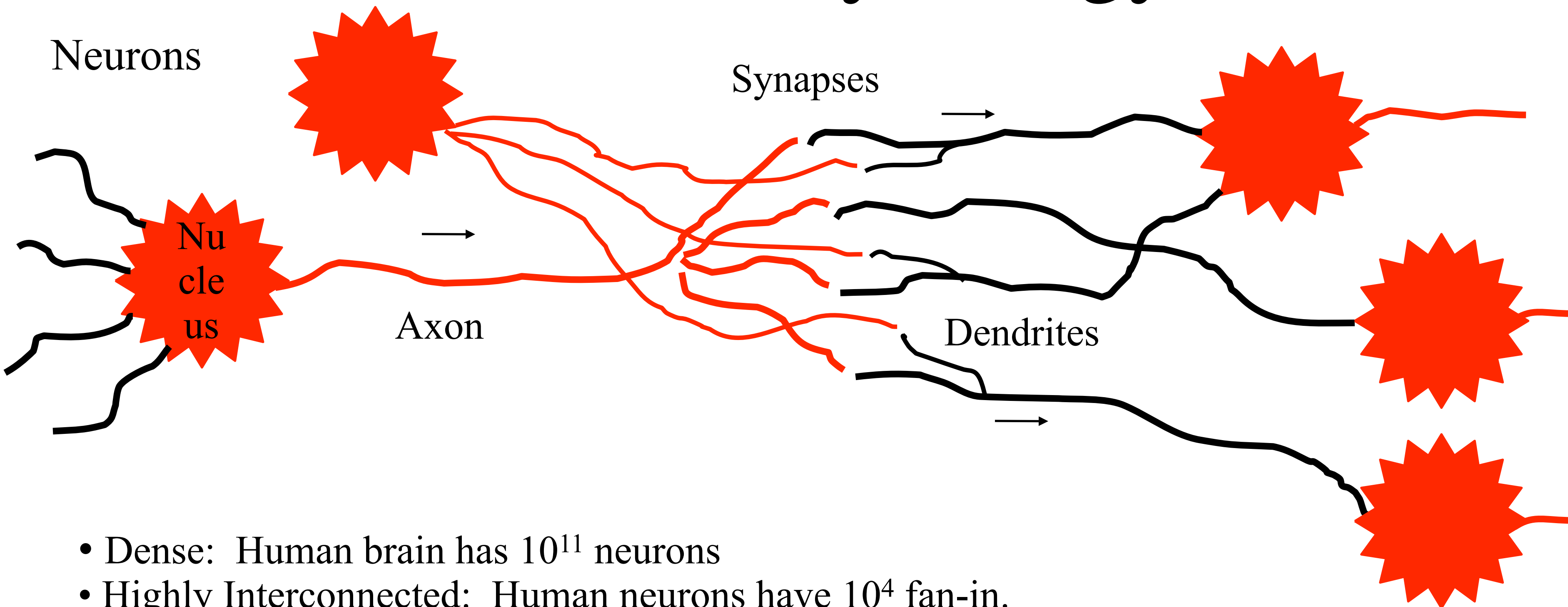
Type II AI: Building Smarter Machines

- AI problems are classically couched as SEARCH problems.
- EAs are very powerful and general search techniques for large, rough solution spaces.
- ANNs are very useful for finding a general function (mapping from inputs to outputs) for a particular set of input-output pairs, and many problems can be viewed as function regression.
- ANNs and Bayesian Networks are two standard statistical machine-learning techniques.
- ANNs provide a very flexible form of knowledge representation that:
  - a) enables the system to LEARN the salient features of a domain, but
  - b) the representations used by a learning ANN can be nearly impossible for a human to comprehend! Typically engineering task decomposition $\neq$ neural network modules.

Artificial Neural Networks (ANN)

- Distributed representational and computational mechanism based (very roughly) on neurophysiology.
- A collection of simple interconnected processors (neurons) that can learn complex behaviors & solve difficult problems.
- Wide range of applications:
  - Supervised Learning
    - Function Learning (Correct mapping from inputs to outputs)
      - Time-Series Analysis, Forecasting, Controller Design
    - Concept Learning
      - Standard Machine Learning Classification tasks: Features => Class
  - Unsupervised Learning
    - Pattern Recognition (Associative Memory models)
      - Words, Sounds, Faces, etc.
    - Data Clustering
      - Unsupervised Concept Learning

NeuroPhysiology



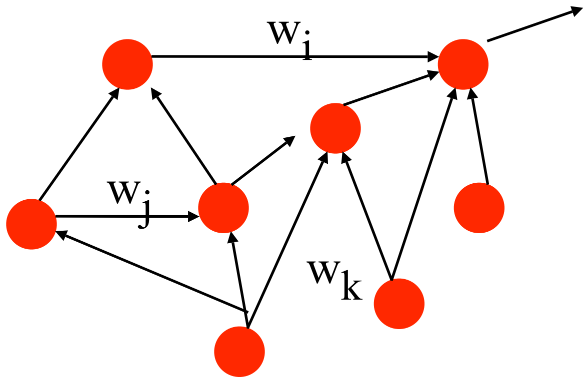
- Dense: Human brain has 10^{11} neurons
- Highly Interconnected: Human neurons have 10^4 fan-in.
- Neurons firing: send action potentials (APs) down the axons when sufficiently stimulated by SUM of incoming APs along the dendrites.
- Neurons can either stimulate or inhibit other neurons.
- Synapses vary in transmission efficiency

Development: Formation of basic connection topology

Learning: Fine-tuning of topology + Major synaptic-efficiency changes.

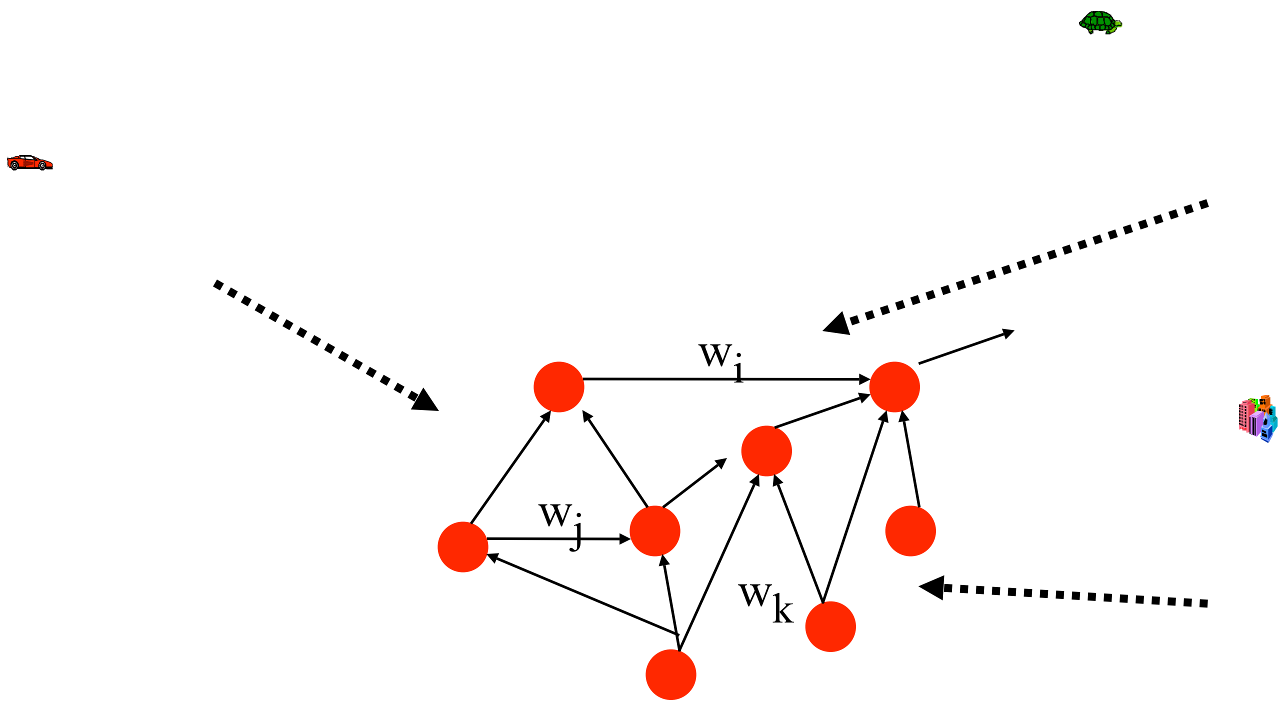
The matrix IS the intelligence!

NeuroComputing



- Nodes fire when sum (weighted inputs) > threshold.
  - Other varieties common: unthresholded linear, sigmoidal, etc.
- Connection topologies vary widely across applications
- Weights vary in magnitude & sign (stimulate or inhibit)
- Learning = Finding proper topology & weights
  - Search process in the space of possible topologies & weights
  - Most ANN applications assume a fixed topology.
- The matrix **IS** the learning machine!

Distributed Information Storage & Processing

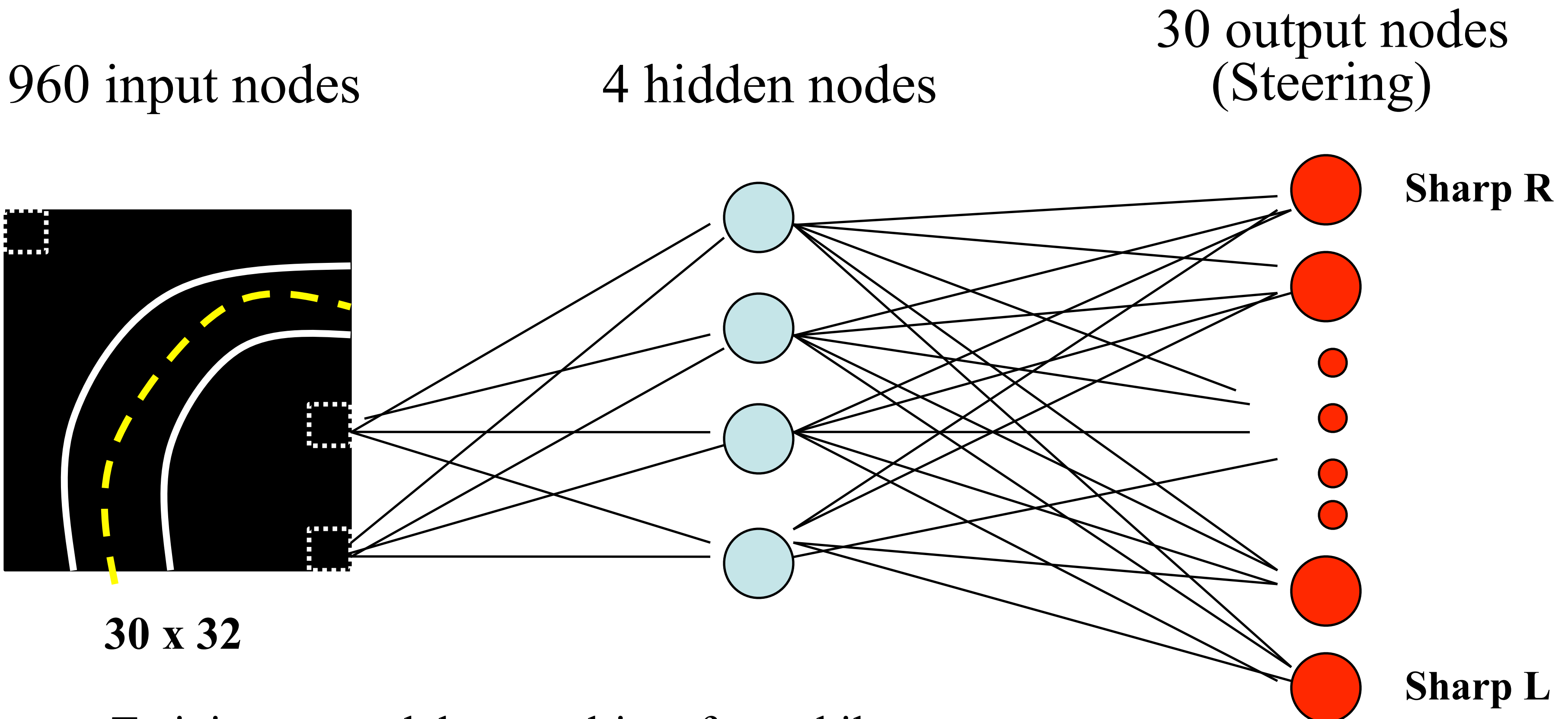


<u>Information is stored in the weights with:</u>

- Concepts/Patterns spread over many weights, and nodes.
- Individual weights can hold info for many different concepts

Driving Miss ANN

- ALVINN (Pomerleau, 1993)



- Training = watch human driver for awhile
- Testing = ANN drives alone…up to 90 miles
- Drove 98% alone across USA
- Symbolic methods for this task can't compare!

Speech Generation

<u>Given</u>: written text

<u>Produce:</u> the appropriate sounds

<u>Mapping:</u> Letters (26) ==> Phonemes (79)

<u>Context Sensitivity:</u> Makes the problem difficult.

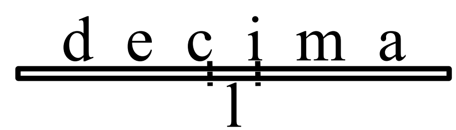
The a in ”bad” is pronounced differently than in ”car”.

English is particularly difficult in this respect.

<u>DECtalk</u>

- Digital Equipment Corporation (DEC) produced software to read books aloud - very useful for the blind.

- Uses a context window of 7 characters.



- Complex set of rules + database of exceptions

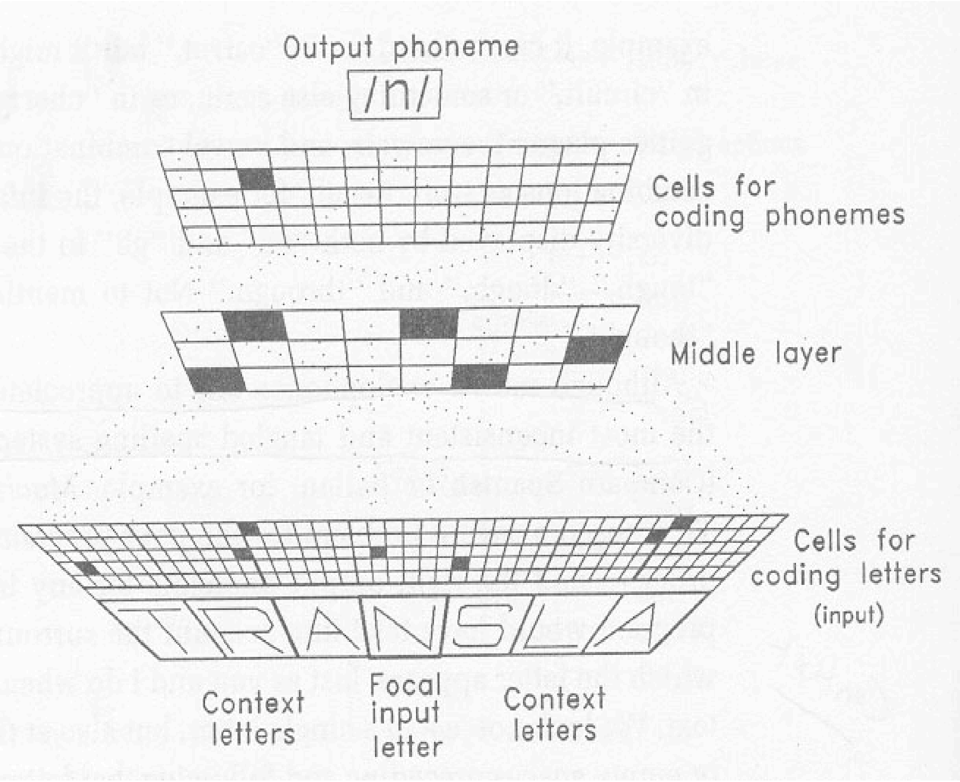
- Several <u>man years</u> of programming effort

- Classic symbolic AI approach.

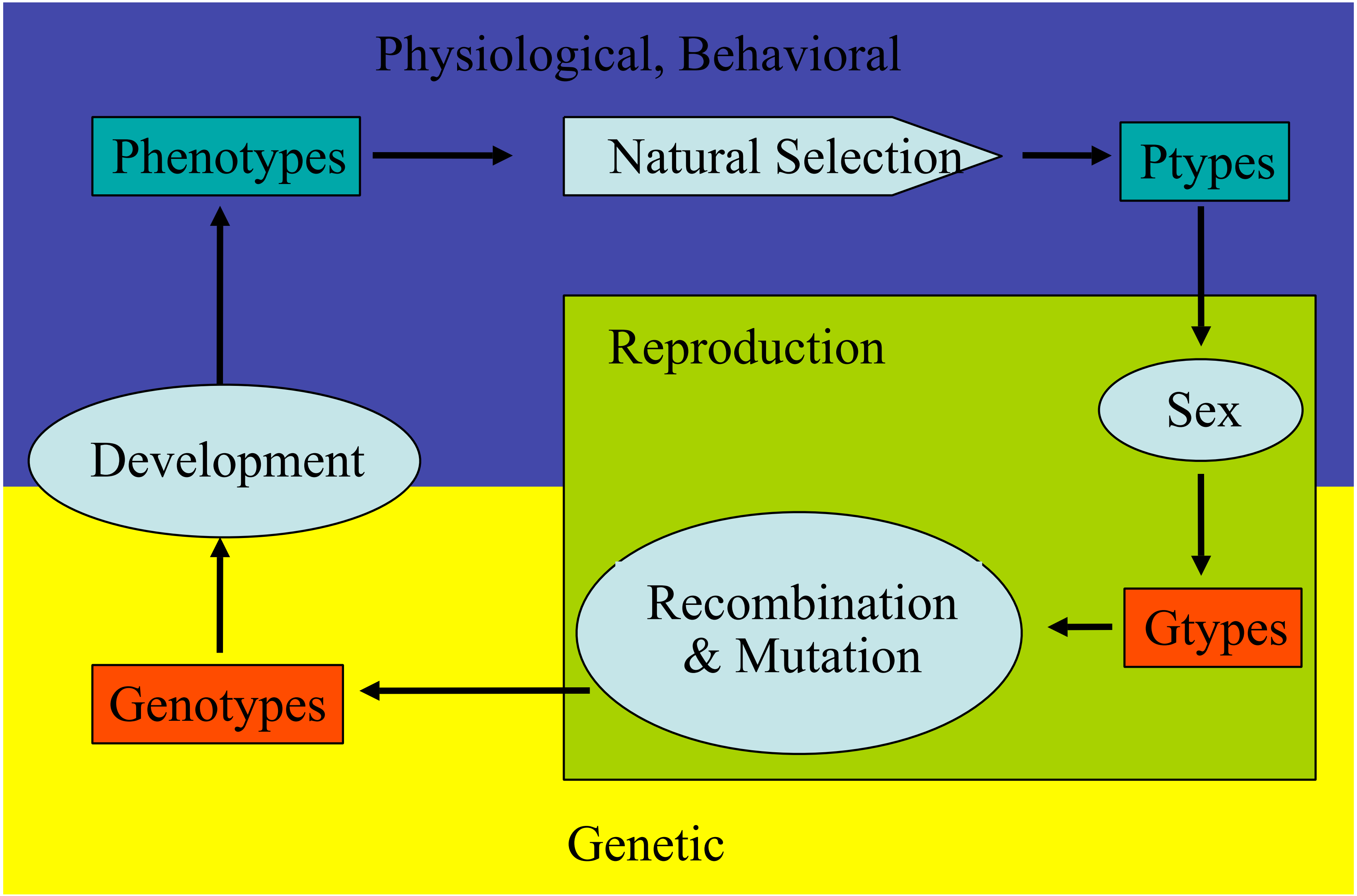
NETtalk

- Terry Sejnowski & Charles Rosenberg (1987).
- Simple 3-layered feed-forward network - trained with backprop
- Training:
  - Moved the 7-letter window over a 1000-word text and saved (window-contents, phoneme) pairs as the training set.
  - Ran epochs for <u>10 hours</u>.
  - 95% accuracy on training set.

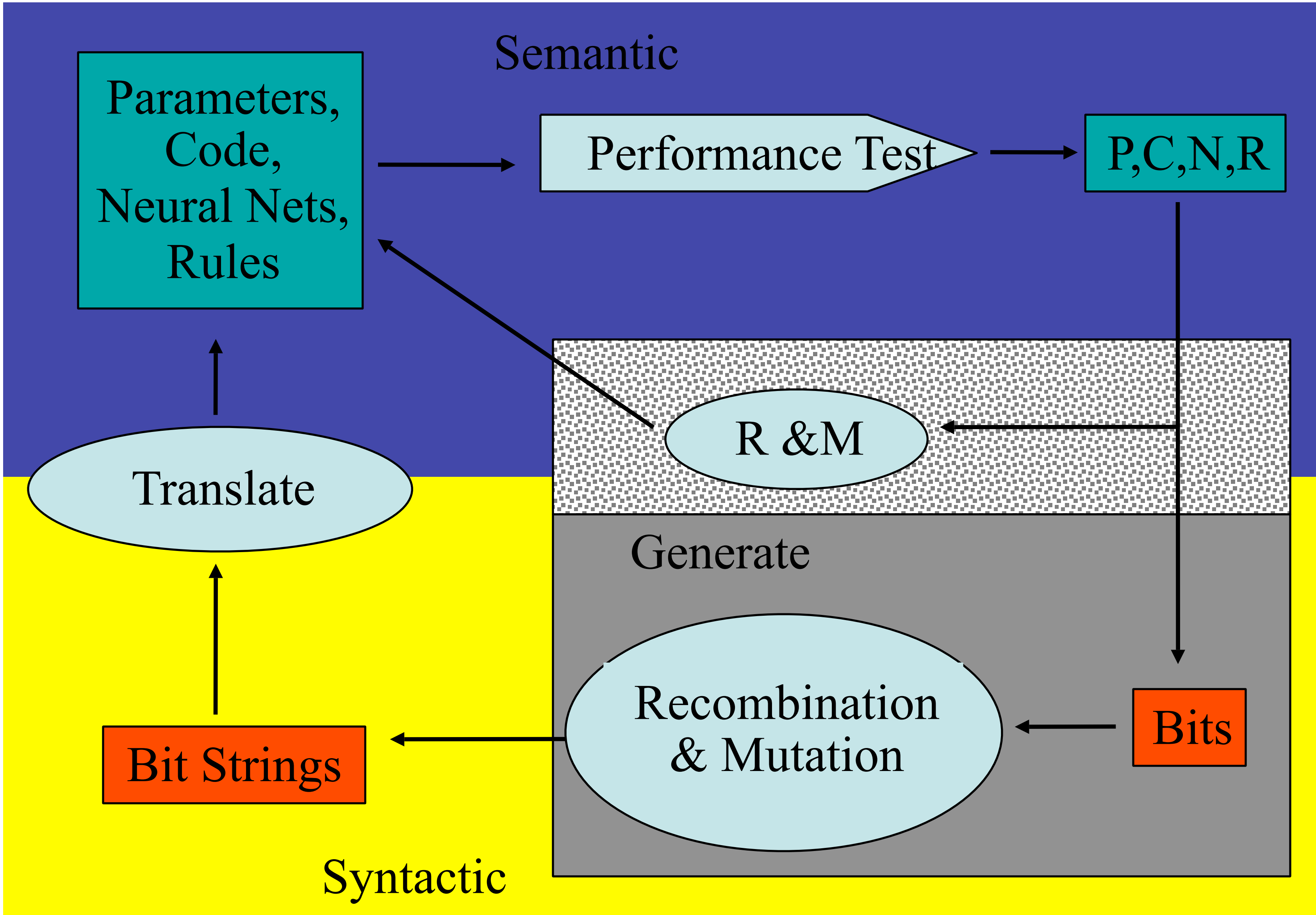
- Testing:
  - 78% accuracy
  - But the speech was still highly understandable.
  - 97.5% accuracy given larger training sets!!!



Darwinian Evolution



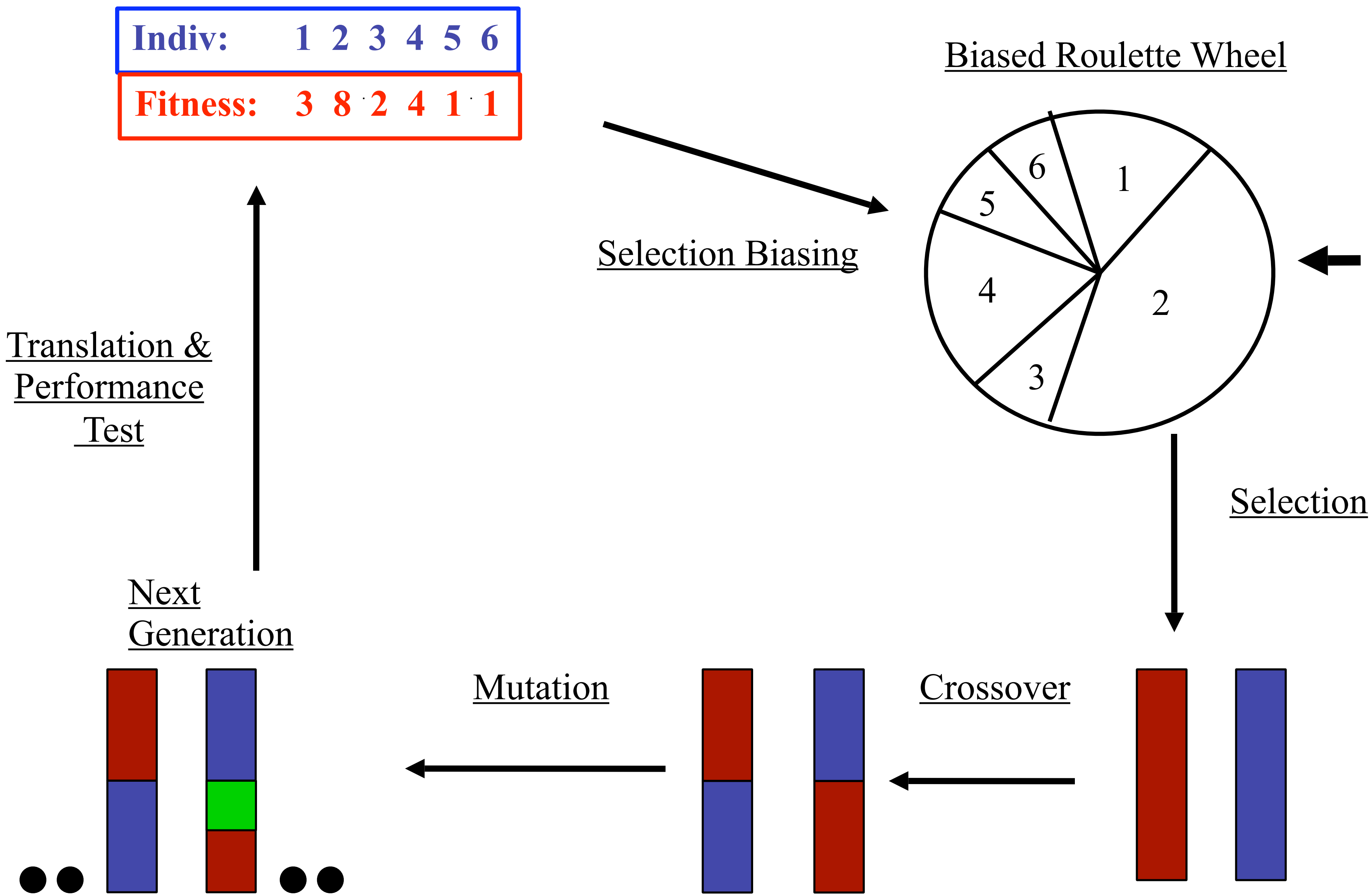
Evolutionary Algorithms



EA Application Areas (Large Search Spaces)

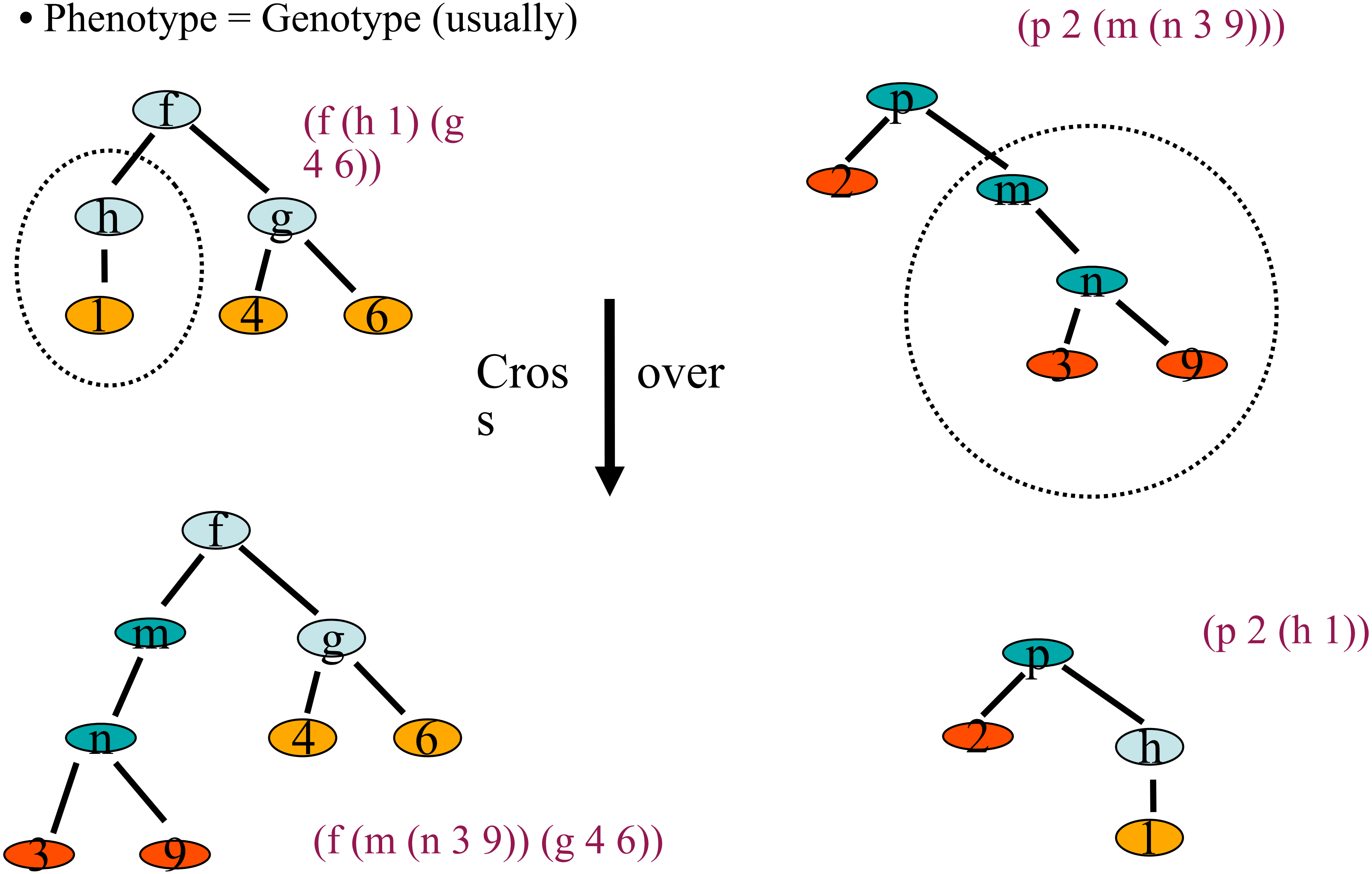
- Optimization: Controllers, Job Schedules, Networks(TSP)
- Electronics: Circuit Design (GP)
- Finance: Stock time-series analysis & prediction
- Economics: Emergence of Markets, Pricing & Purchasing Strategies
- Sociology: cooperation, communication, ANTS!
- Computer Science
  - Machine Learning: Classification, Prediction…
  - Algorithm design: Sorting networks
- Biology
  - Immunology: natural & virtual (computer immune system)
  - Ecology: arms races, coevolution
  - Population genetics: roles of mutation, crossover & inversion
  - Evolution & Learning: Baldwin Effect, Lamarckism…

Evolutionary Computation = Parallel Stochastic Search



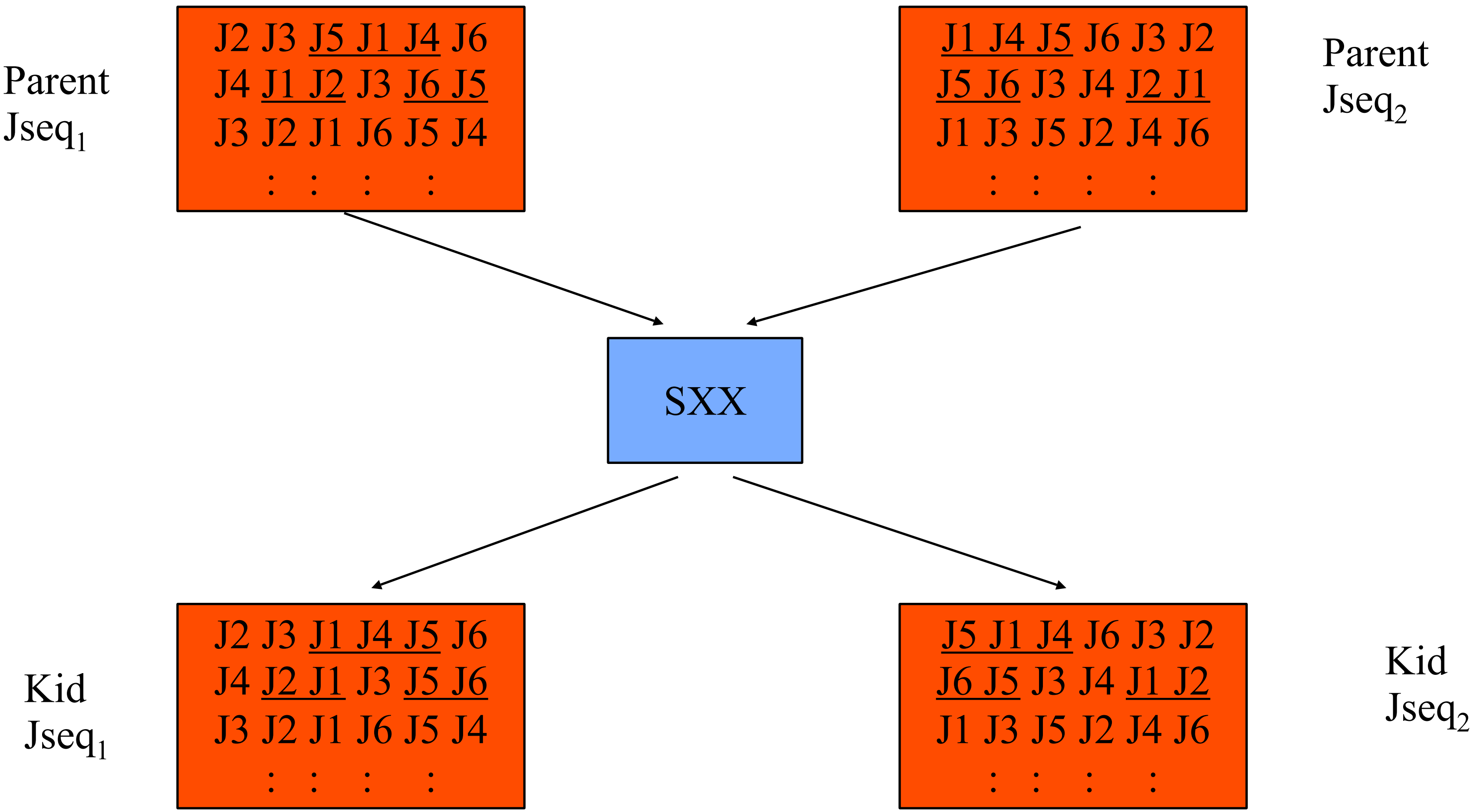
Genetic Programming (GP)

- Genotype = a computer program (often in Lisp)
- Phenotype = Genotype (usually)



Job-Shop Scheduling using EAs

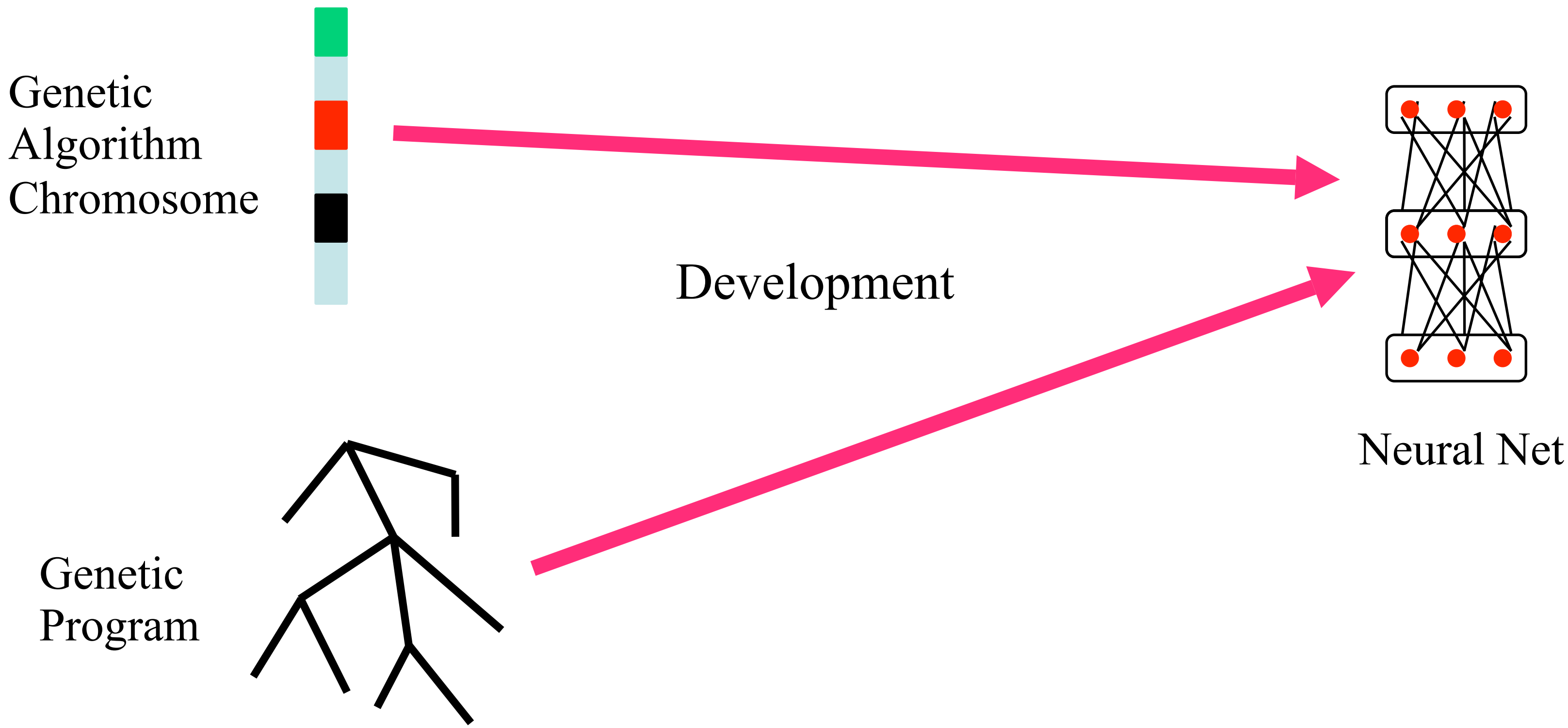
Swap similar (but permuted) subsets of parent job sequences (Kobayashi et. al, (1995))



Creative Evolutionary Systems

- Evolutionary Music
  - Chromosomes encode:
    - sequences of pitch, duration, loudness, etc.
  - User-based fitness:
    - Direct - selection of best melodies
    - Indirect- adjusting fitness-function parameters
- Evolutionary Art
  - Chromosomes encode:
    - Mappings: pixel-coordinates => color
    - Parameters for pattern-drawing routines
    - Programmed sequences (GP) of drawing activities
  - Fitness: humans select favorites of each generation
- Evolutionary Design
  - Chromosomes encode:
    - Blueprints describing tables, electronic circuits, antennae, ANNs, etc.
    - Recipes for generating " " " " " "
  - Fitness evaluated on-line by simulated performance tests.

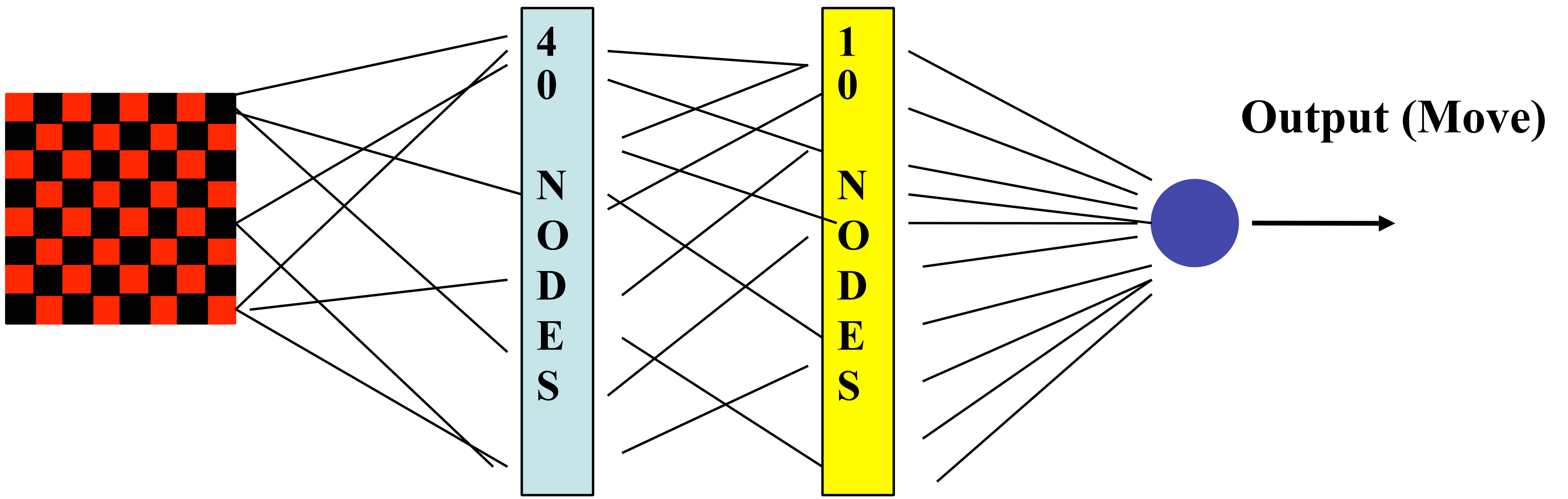
Evolving Neural Networks



- Combining evolution and learning
  - Evolve weights, topologies, neuron-type distributions, learning rules, etc.
  - Powerful adaptivity
- Replacing (weight) learning with evolution
  - Popular in un-supervised and low-supervision ANN applications such as evolutionary robotics.

Blondie 24

- **Fogel & Chellapilla (2002)**
- **Chromosome: the weights for a fixed-topology ANN.**
- **Fitness test: coevolution among a population of ANN checkers players.**



- Took months of co-evolution to train
- Beat Chinook (world champion - also a computer program)..once.
- Achieved very high rating on zone.com (continuous on-line checkers tournament)

Evolutionary Robotics

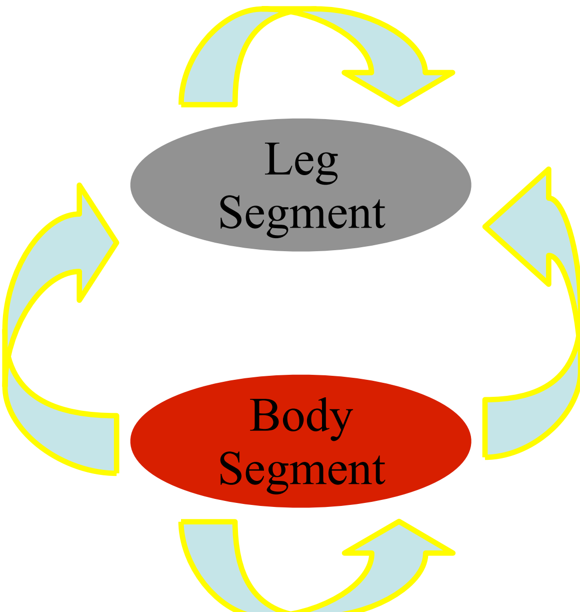
- Using EAs to design controllers (and occasionally morphologies) of autonomous robots.
- Controllers are often ANNs.
- Follows the behavior-based robotics paradigm
  - Bottom up approach
  - Minimal representations of the world. *The world is its own best model* (Rodney Brooks, MIT).
  - Distributed, self-organizing; no fixed command hierarchy.
- Chromosomes encode
  - Weights, topologies, learning rules for ANNs
- Fitness
  - Place real or simulated robot in an environment and give fitness points based on its ability to perform a task.
  - Often evolve hundreds of generations in simulation, then download the best controllers into real robots and evolve a little more on-board.

Evolving Morphology & Control (Karl Sims)

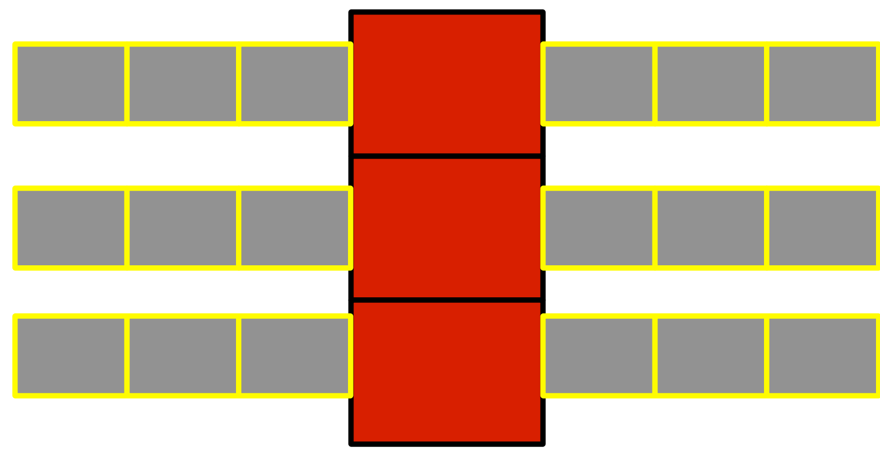
Morpholog
y

Genotype

Phenotype



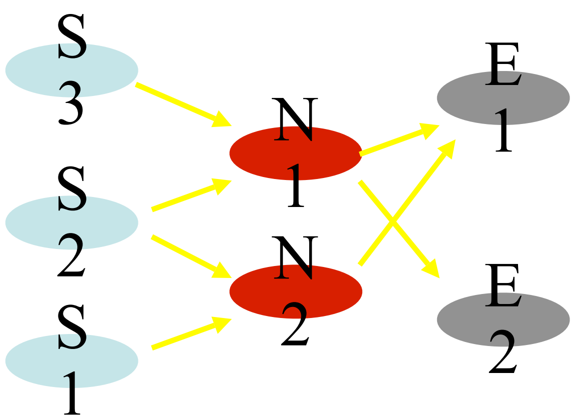
Segment Graph

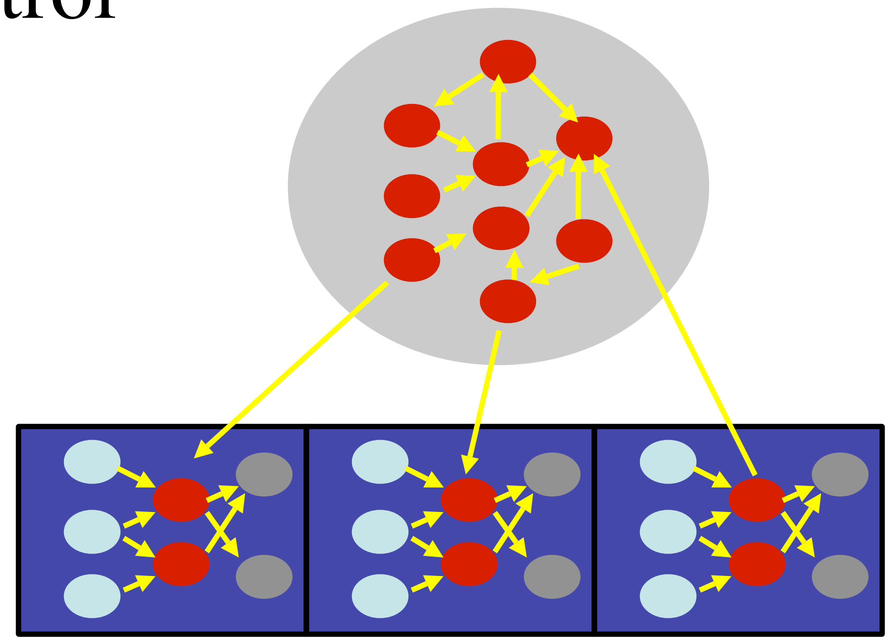


3d Segments & Joints

Control

Genotype & Phenotype





Neural Circuit Graphs

- Joint-Angle Sensors
- Joint-Force Effectors

Each segment & brain
have a controller

Tasks: Swimming, Running, Jumping, Capturing
Search: Evolution - Selection, Mutation, Crossover

Neural Coding

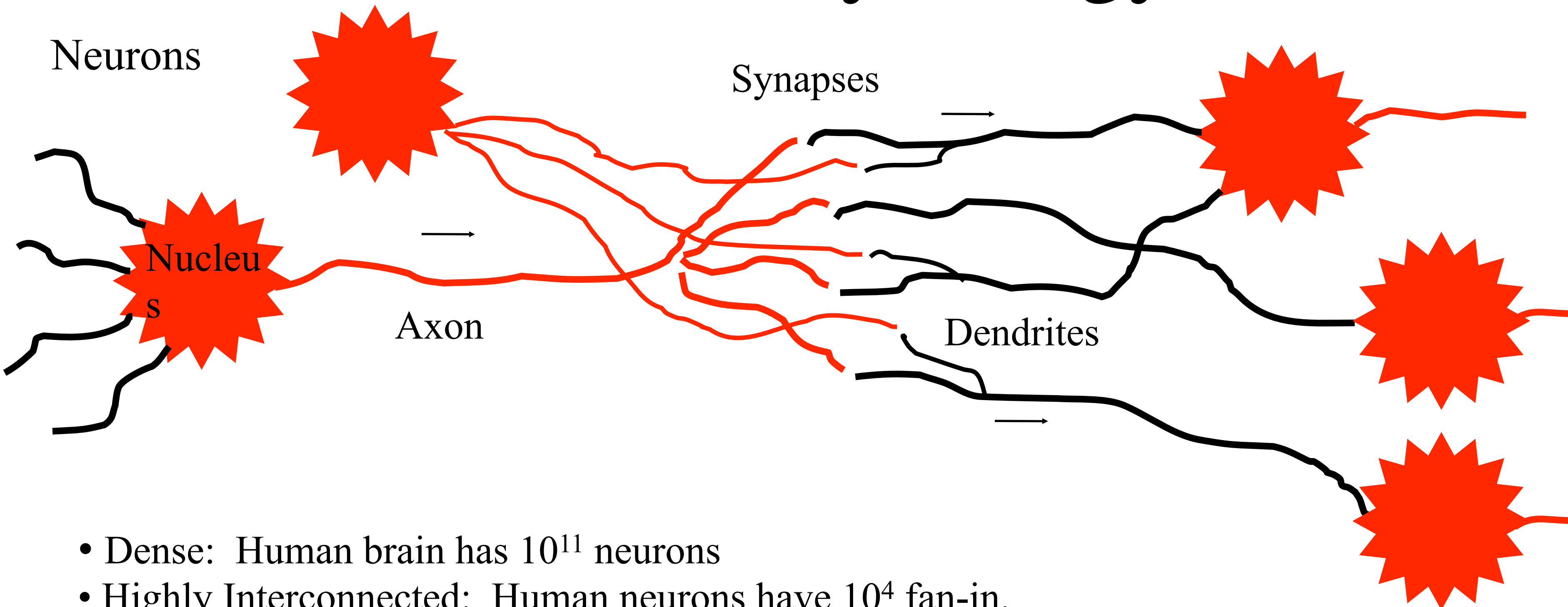
What Kind of Information is Represented in a Neural Network and **How**?

Outline

- ANN Basics
- Local -vs- Distributed Representations
- The Emergence/Learning of Salient Features in Neural Networks
- Feed-Forward Neural Networks Embody Mappings
- Linearly Separable Mappings
- Classification in Spaces that are NOT Linearly Separable
- Coding Heuristics

ANN Basics

NeuroPhysiology



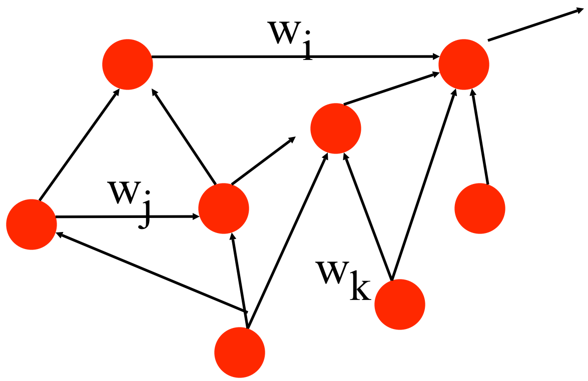
- Dense: Human brain has 10^{11} neurons
- Highly Interconnected: Human neurons have 10^4 fan-in.
- Neurons firing: send action potentials (APs) down the axons when sufficiently stimulated by SUM of incoming APs along the dendrites.
- Neurons can either stimulate or inhibit other neurons.
- Synapses vary in transmission efficiency

Development: Formation of basic connection topology

Learning: Fine-tuning of topology + Major synaptic-efficiency changes.

The matrix IS the intelligence!

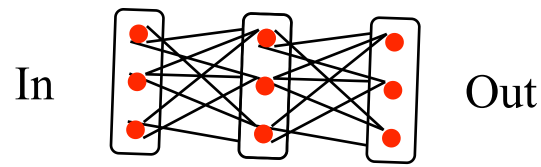
NeuroComputing



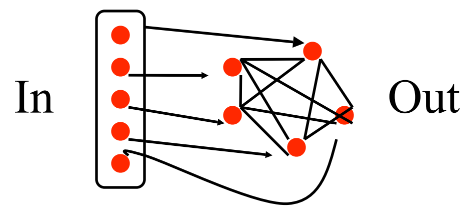
- Nodes fire when sum (weighted inputs) > threshold.
  - Other varieties common: unthresholded linear, sigmoidal, etc.
- Connection topologies vary widely across applications
- Weights vary in magnitude & sign (stimulate or inhibit)
- Learning = Finding proper topology & weights
  - Search process in the space of possible topologies & weights
  - Some ANN applications assume a fixed topology, some others dynamic.
- The weight matrix **IS** the learning machine!

Tasks & Architectures

- Supervised Learning
  - Feed-Forward networks
    - Concept Learning: **Inputs** = properties, **Outputs** = classification
    - Controller Design: **Inputs** = sensor readings, **Outputs** = effector actions
    - Prediction: **Inputs** = previous X values, **Outputs** = predicted future X value
  - Learn proper weights via back-propagation

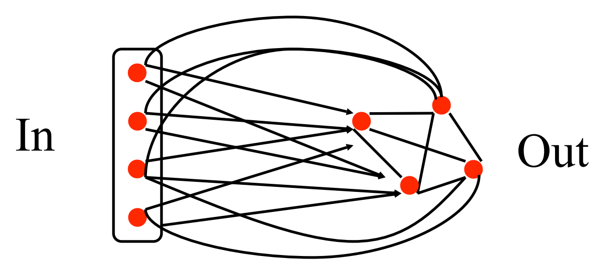


- Unsupervised Learning
  - Pattern Recognition
    - Hopfield Networks



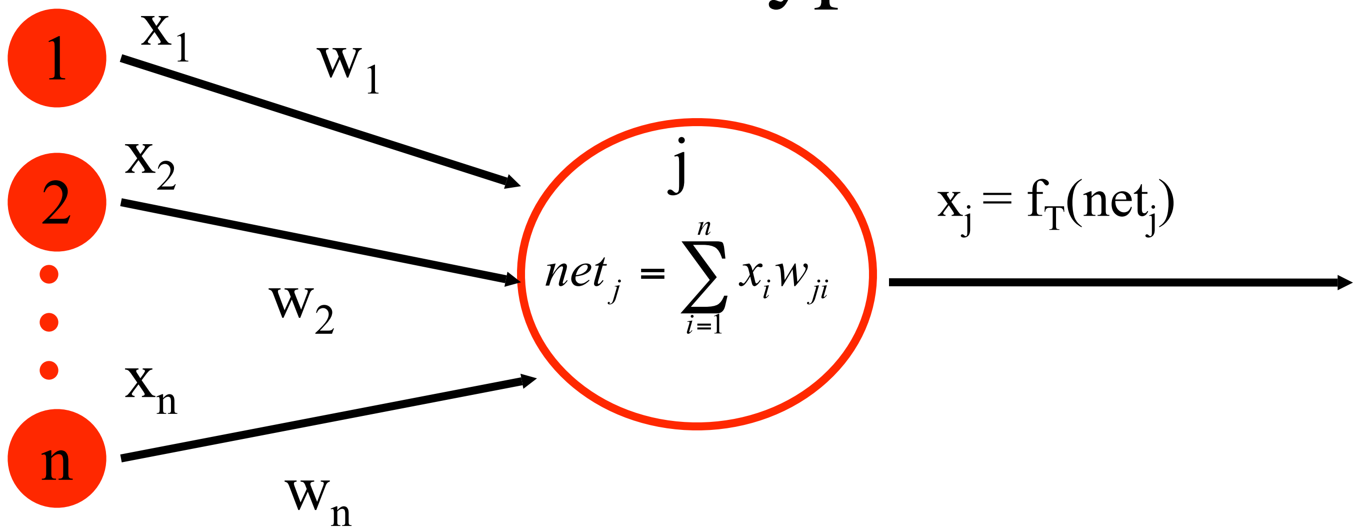
Excitatory & Inhibitory Arcs in the Clique

  - Data Clustering
    - Competitive Networks



Maxnet: Clique = only inhibitory arcs

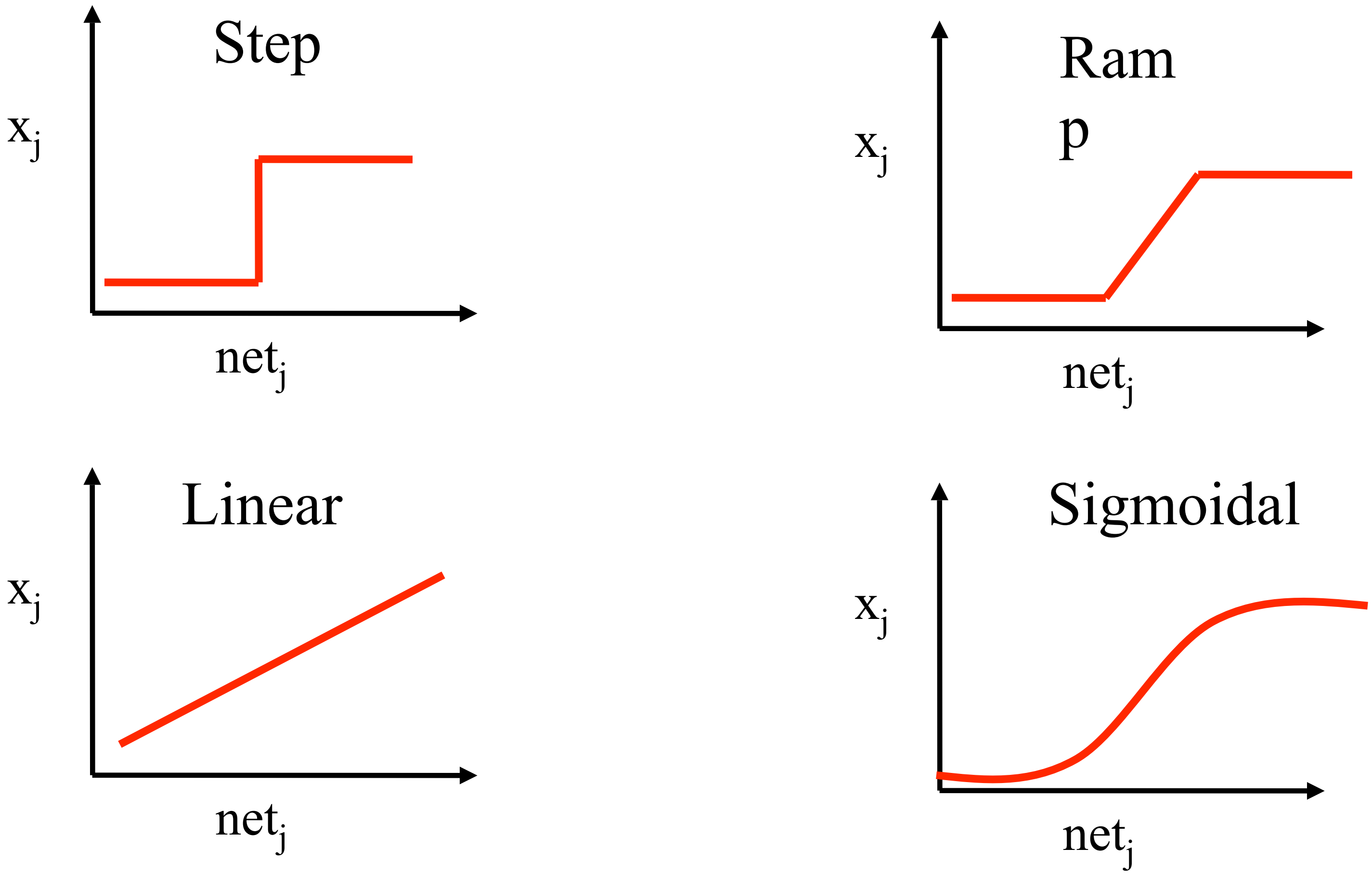
Node Types



- Most ANNs use nodes that sum the weighted inputs.
- But many types of transfer functions, f_T are used.

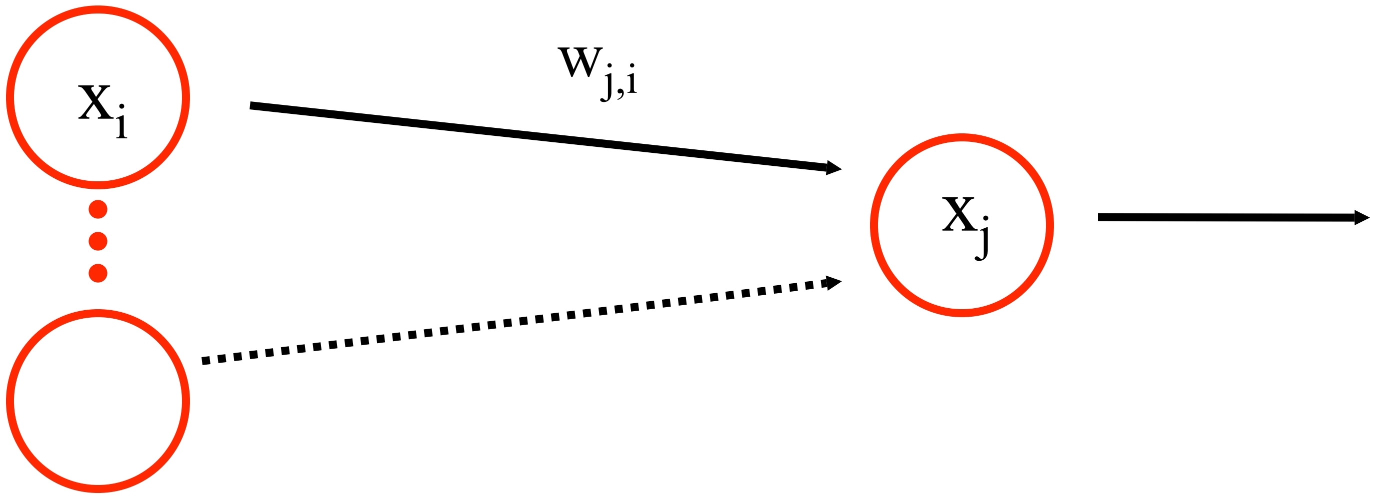
  - Thresholded (Discontinuous)
    - Step
    - Ramp
  - Non-thresholded (Continuous, Differentiable)
    - Linear
    - Sigmoid

Transfer Functions



- Step functions are useful in classifier nets, where data partitioning is important.
- Linear & Sigmoidal are everywhere differentiable, thus popular for backprop nets.
- Sigmoidal has most biological plausibility.

Learning = Weight Adjustment



- **Generalized Hebbian Weight Adjustment:**

  When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes place in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.

- **For ANN, the simplest form of weight modification rule can be stated as:**

  $w_{j,i} = cx_ix_j$

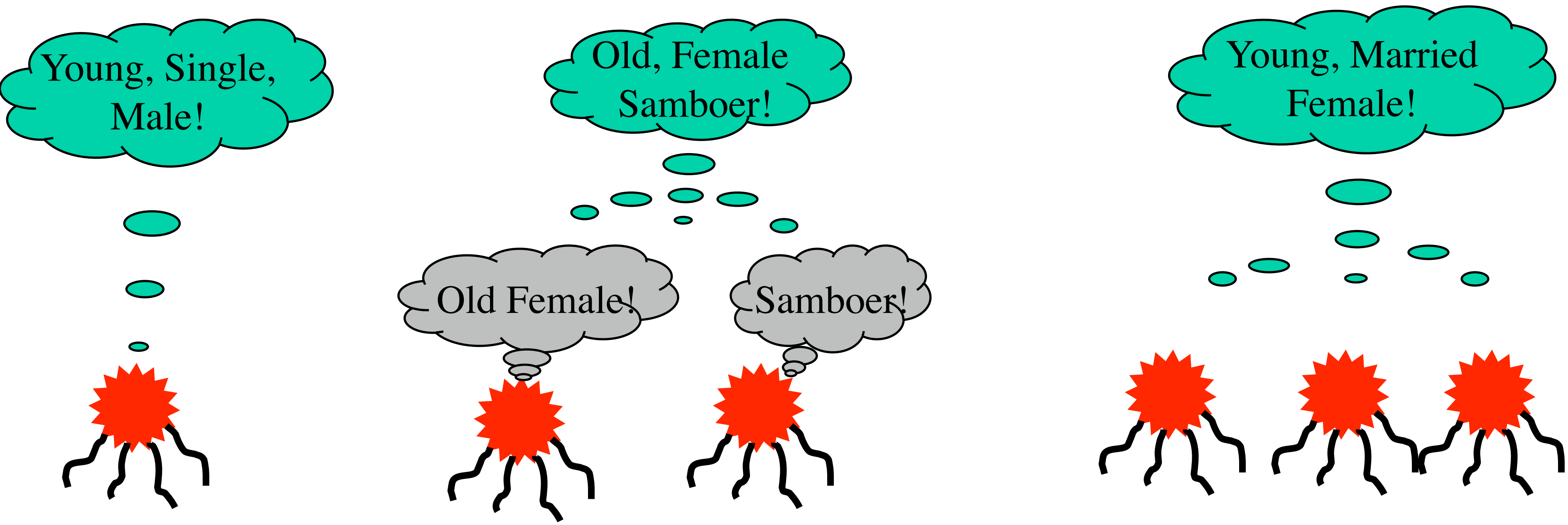
  c: some small constant that denotes the strength of the connection from node j to node i.

  x_i, x_j : activation levels of these nodes.

Distributed Representations

Local -vs- Distributed Representations

- Assume examples/concepts have 3 features:
  - Age : {Young, Middle, Old}
  - Sex: {Male, Female}
  - Marital Status: {Single, Samboer, Married}



Local: One neuron represents an entire conjuctive concept.

Semi-Local: Together they rep a conjunctive concept, and each neuron reps **one or a few** conjuncts - i.e. concept broken into clean pieces.

Distributed: Together they rep a conjunctive concept, but the individual conjuncts cannot necessarily be localized to single neurons

Local

- Size requirements to represent the whole set of 18 3-feature concepts - assuming binary neurons (on/off)
  - Local: 3x3x2 = 18
    - Instance is EXACTLY 1 of 18 neurons being on.

Young, married, male

Young, married, female

Young, unmarried, male

Young unmarried, female

.

.

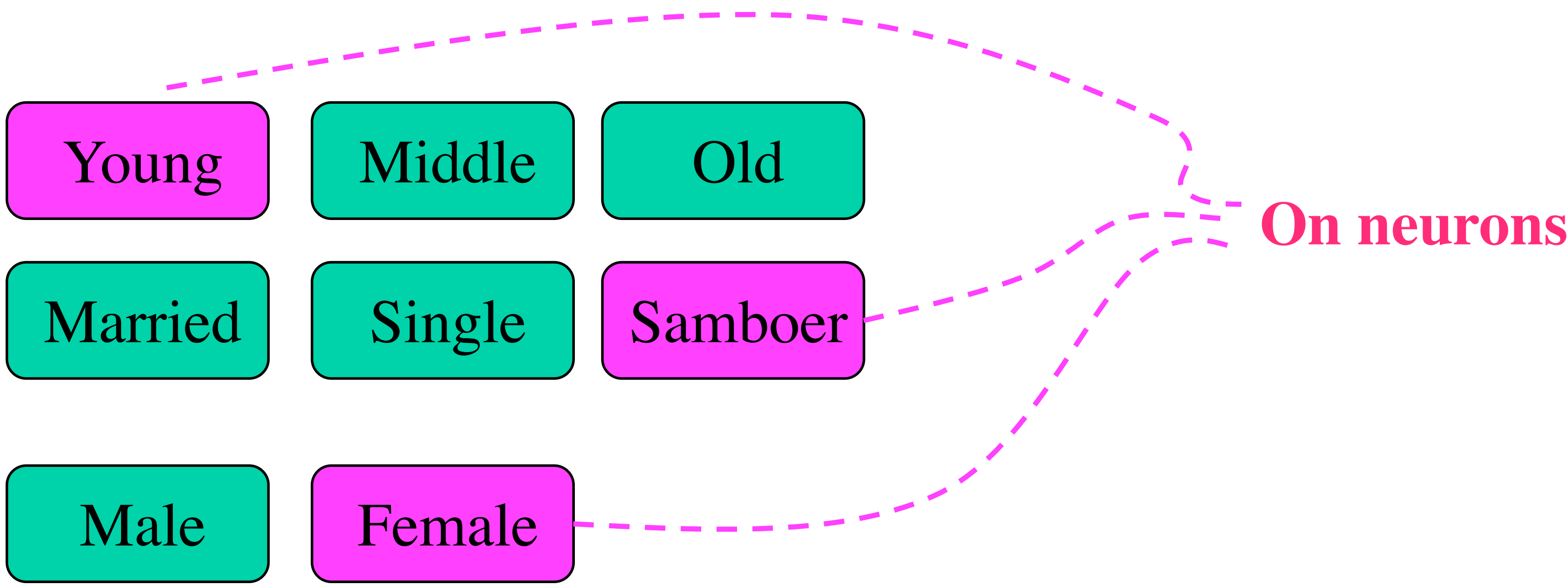
.

18 different triples! Each triple is represented by a single neuron.

On neuron

Semi-Local

- Semi-Local: 3+3+2 = 8 (Assume one feature value per neuron)
  - Instance is EXACTLY 3 of 8 neurons being on.



Distributed

– Distributed: $\log_2 18 = 5$
  - Instance is any combination of on/off neurons
  - An example distibuted assignment can be as follows:

x x x **0 0** (**Married**) **x = don’t care**

x x x **0 1** (**Single**)

x x x **1 1** (**Samboer**)

x **0 0** x x (**Young**)

x **0 1** x x (**Old**)

x **1 1** x x (**Middle**)

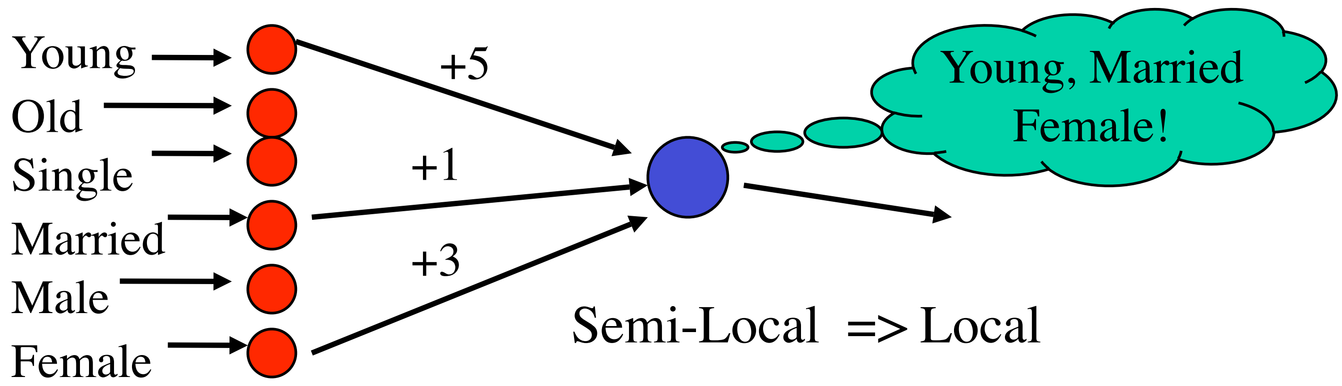
0 x x x x (**Male**)

1 x x x x (**Female**)

  - Add 1 bit and DOUBLE the representational capacity, so each concept can be represented by 2 different codes (redundancy).

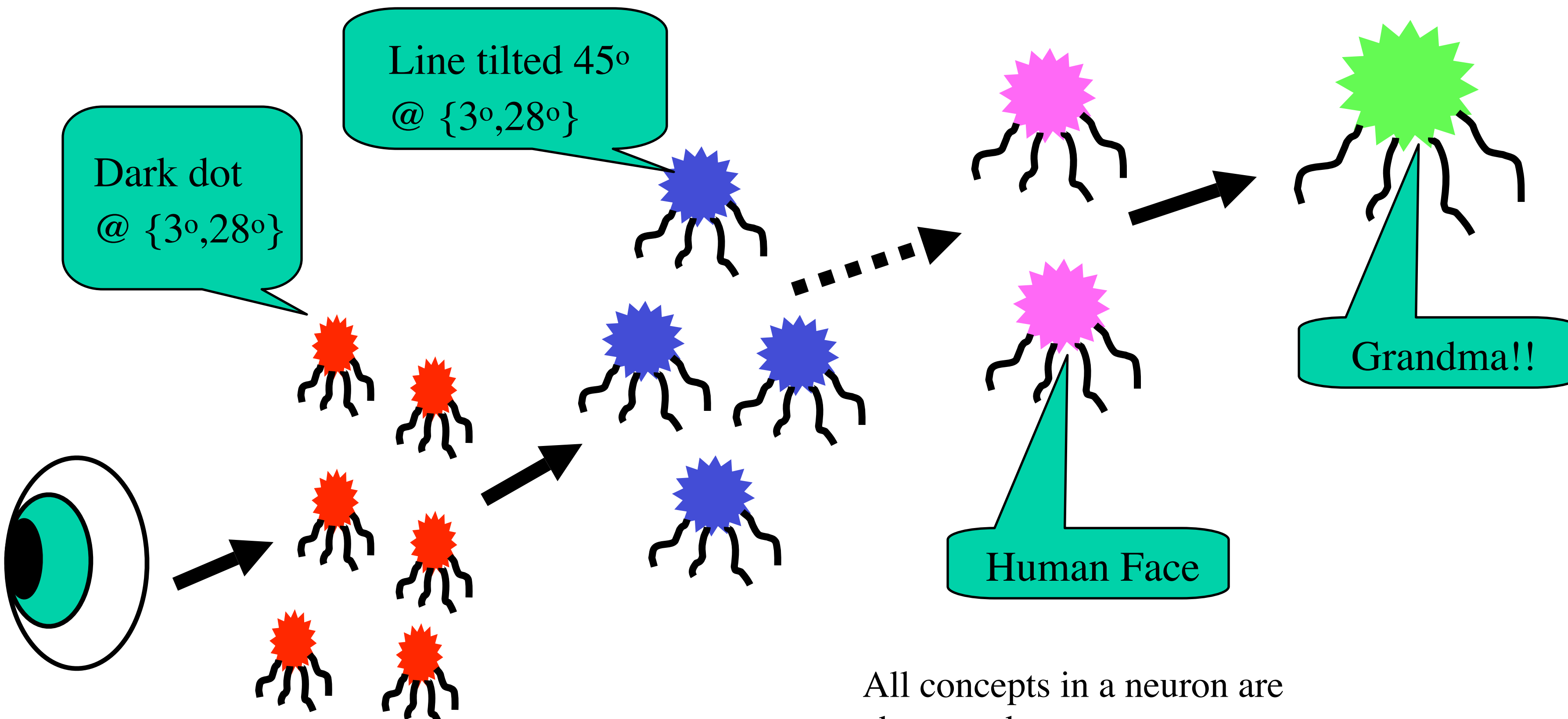
Varying code type in different regions of a network

- The same neural network (artificial or real) may have different types of coding in different regions of the network.



Brain Representational Hierarchies

- In the brain, neurons involved in early processing are often semi-local, while neurons occuring later along the processing path (i.e. higher level neurons), are often local.
- In simpler animals, there appears to be a lot of local coding. In humans, it is still debatable.



All concepts in a neuron are abstracted.

Binding Problem

- (How to represent two concepts that occur simultaneously): Local (EASY! - two active nodes), Distributed (HARD - but may be possible by quick shifts back and forth between the 2 activation patterns)

  E.g. “Where’s Waldo”: Easy to pick out a human face among a bunch of round objects, or your mother’s face among a bunch of other faces, thus indicating that we probably have relatively local codes for these all-important concepts. But, it’s VERY HARD to find Waldo (i.e. a generic-faced cartoon man with a red-and-white striped shirt) in a crowd of several hundred generic cartoon characters wearing all sorts of colors & patterns. Why? **“Red-and-white stripes” is probably not locally coded in the human brain and hence not quickly/ effortlessly detected**. **It doesn’t reside as a whole in one neuron.** It probably shares neurons with concepts such as “stripe” “red”, “white”, etc.

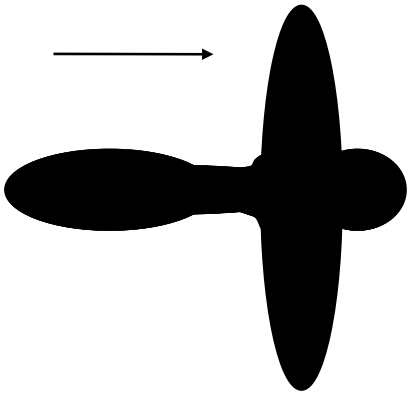
- In more complex animals, probably local representations are used for the ***most salient concepts*** for that organism.

and Learning of

Salient Features in

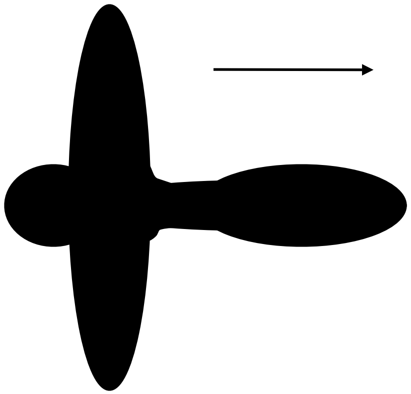
Species-Specific Saliency

- The key stimuli for an organism are often **locally (one neuron represents an entire conjuctive concept)** or **semi-locally (one neuron represents one or more features)** encoded, with direct connections from the detector neuron(s) to a motor (action-inducing) neuron.



The movement of this simple pattern resembles a **hawk** and scares small chickens.

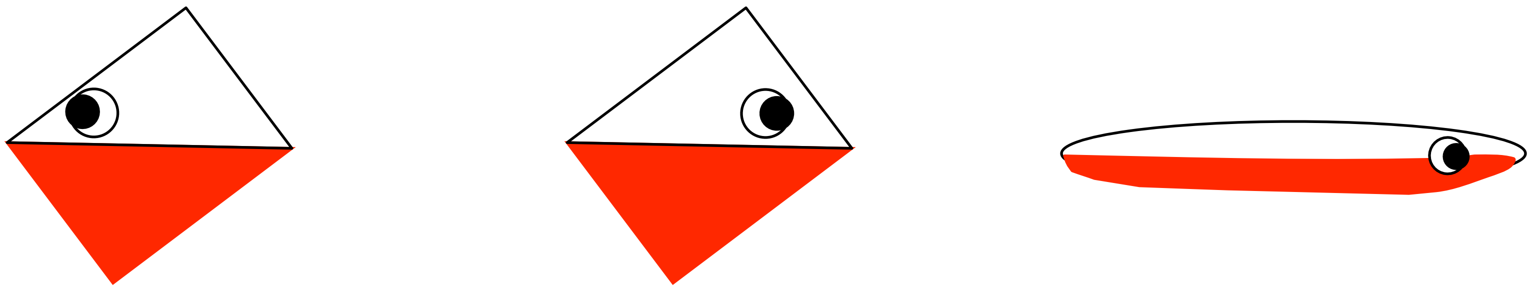
Salient feature: Flying is always with wings front



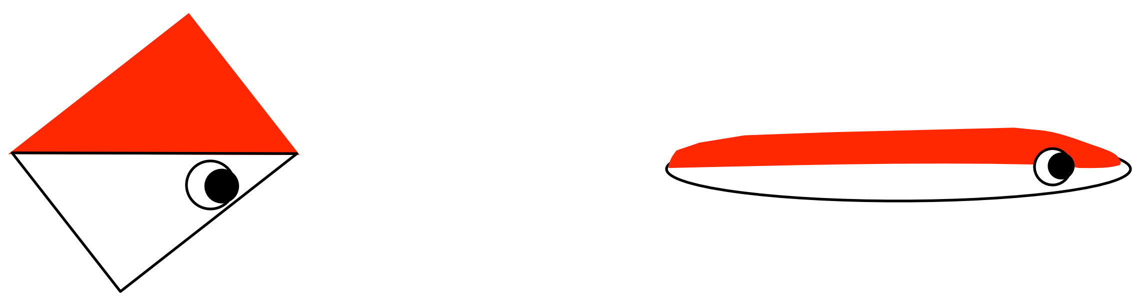
The movement of the reverse pattern ressembles a **goose** and elicits no response from the chicks.

Fish Dinner

- Three-spined sticklebacks respond to these simple stimuli:



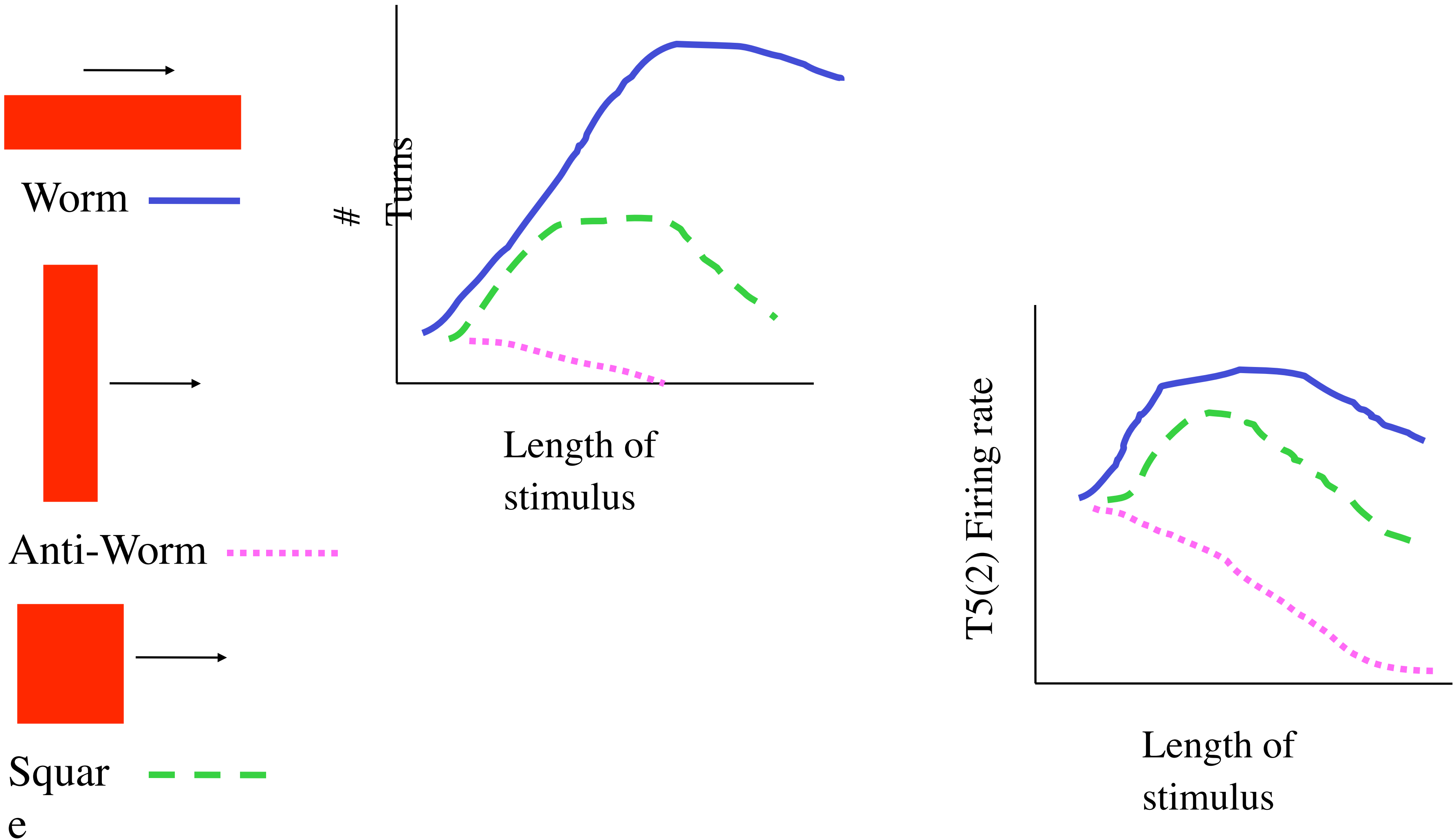
- But not these:



- Salient feature: Red belly! Other similarities are not detected if also representation is local or semi-local.

Toad Turn-ons

- The behavioral response (i.e. number of times that it turns around per minute) of a toad as a function of **the length of the stimulus (salient feature)** is mirrored by the firing rates of neurons in the T5(2) region of its brain.



Emergent Salience

- Animal bodies **and brains** have evolved to maximize the odds of survival and reproduction (i.e., fitness). **Both** are tailored to the survival task at hand.

- Hence salient features will emerge (via evolution and learning) as the activating conditions for various neurons. When fired, those neurons will then help to initiate the proper (motor) response to a salient input. Others which couldn't emerge salient features disappeared.

- Similarly, if an ANN is given a task and the ability to adapt (i.e. learn and/or evolve), **the salient features of that task will emerge as the activating conditions for hidden-layer and output neurons**. Otherwise an ANN will be disqualified to survive!!!

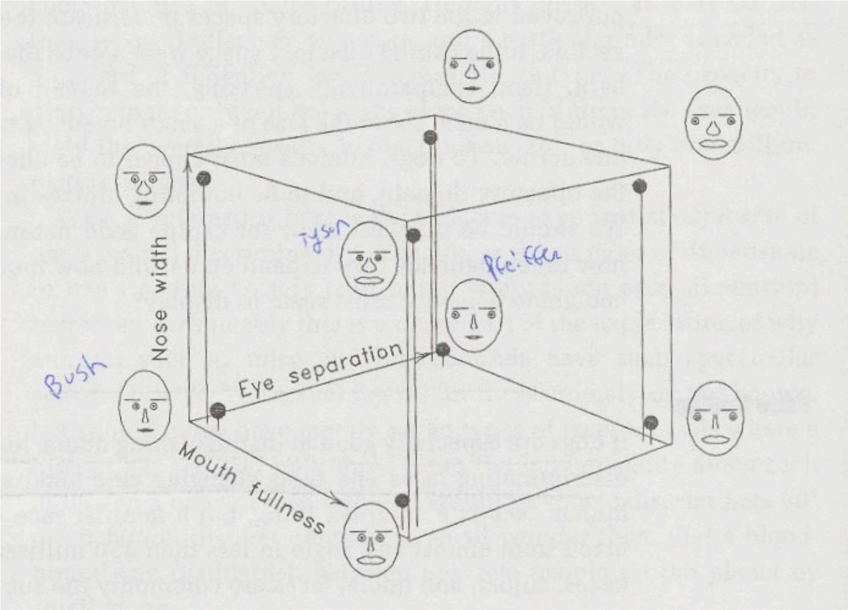
- Salient features can then be read from the input weights to those neurons.

- So, the only features that need to be given to the ANN are the very primitive ones at the input layer. The rest are **discovered** by **emergence**!

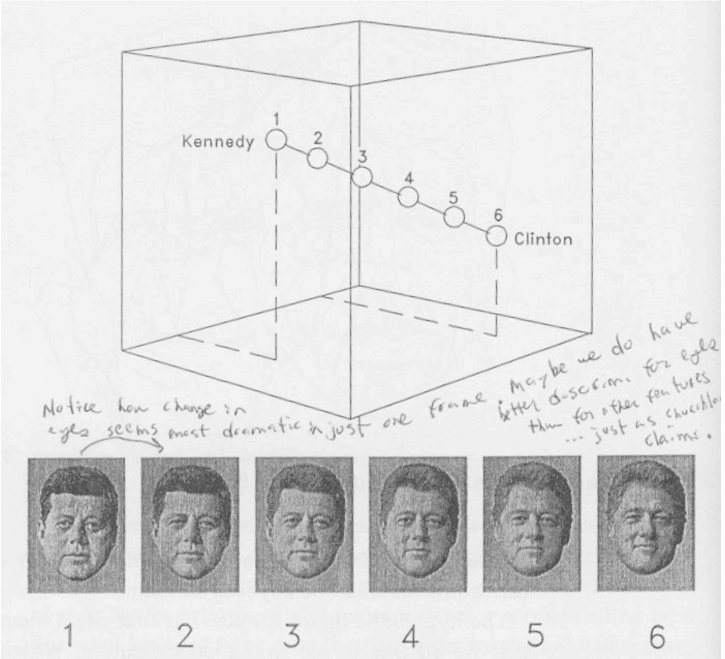
Face Recognition

- Animals differ as to their abilities to disciminate sounds, tastes, smells, colors , etc.
- Humans are very good at disciminating faces, at least faces of the type that they grow up around. **(can detect salient features of the faces very easily).**
- Hypothesized # dimensions in face-coding space = 20 (Churchland)

Face Space



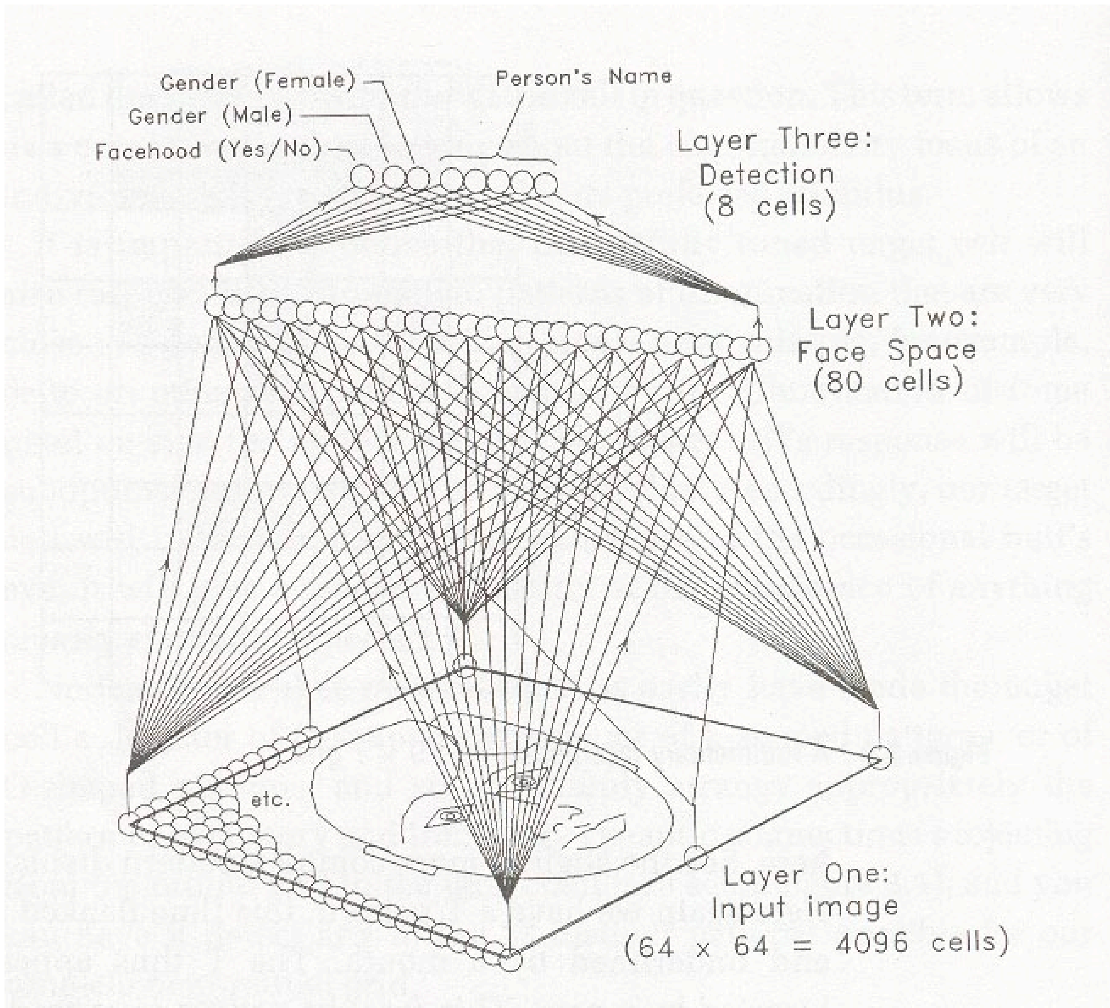
Morphing



Choose evenly-spaced points along the vector that connects the source & target faces

ANN for Face Recognition

- Garrison Cottrell et. al. (1991)
- Feed-forward net with backprop learning



Training & Testing

- Training: 64 photos of 11 different faces + 13 non-face photos
- Performance Criteria: Classify each picture as to:
  - face or non-face?
  - male or female?
  - Name?
- Results:
  - Training Accuracy: 100%
  - Test with same faces but new pictures: 98%
  - Test with new faces : 100% (face-non-face?), 81% (male-female?)
  - Test with known face but with 20% of picture erased:
    - Vector completion: **the firing patterns of middle-layer neurons are very similar to those patterns when the non-erased image is presented**. Hence, in its *understanding* of the pictures, the ANN fills in the missing parts.
    - Generally good performance, but erased foreheads caused problems (71% recognition).
  - Holons: Middle-layer nodes represent generic combi-faces instead of individual features.

Facial Holons

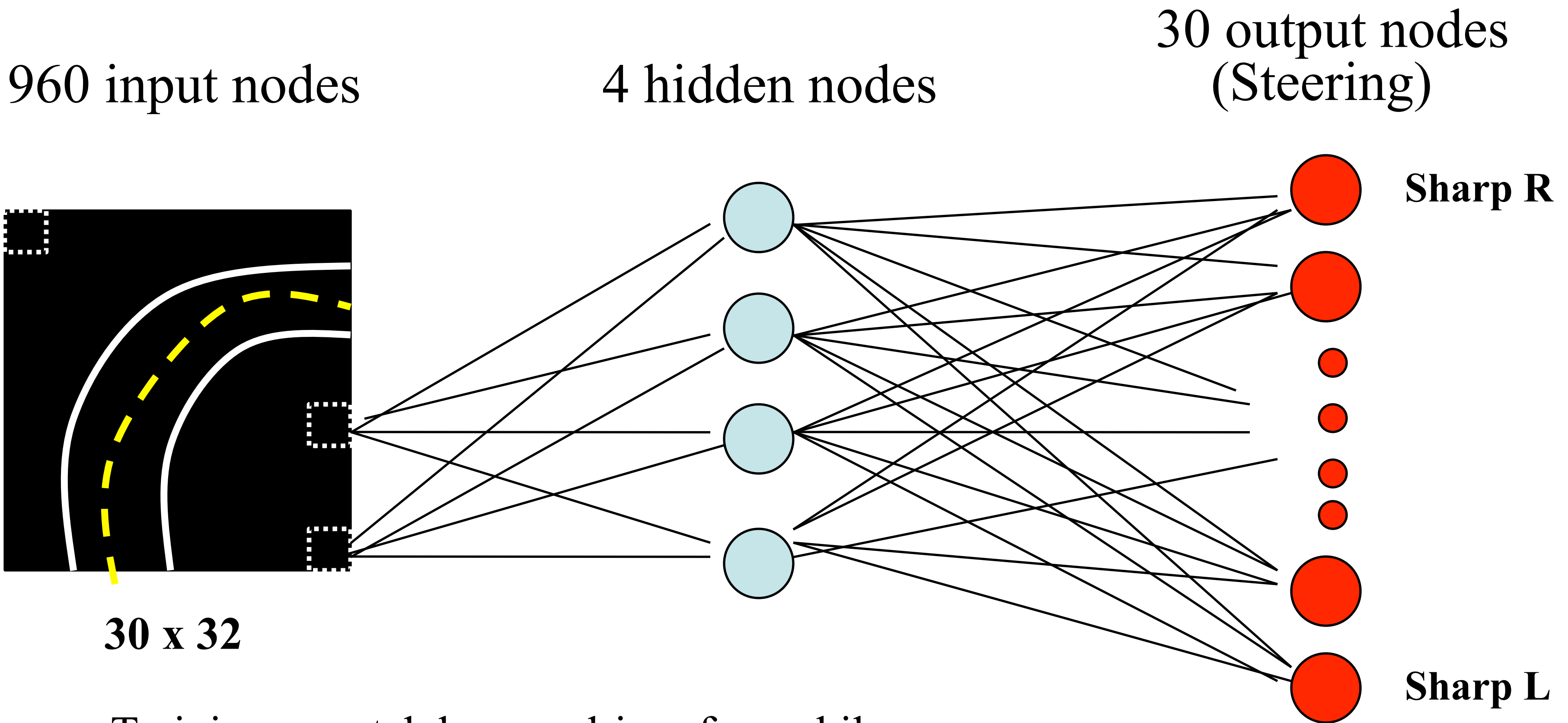
Each input case satisfies a subset of the 80 holons. i.e., if each input case is a combination of holons

- Prefered stimuli: By looking at the signs of the input weights to a hidden node, we can construct a prototypical input vector that the node would fire on. E.g. If $w_{ji} > 0$, then $x_i > 0$ is desired, and if $w_{ji} < 0$, then $x_i < 0$ is desired (**supply positive contribution to hidden nodes**).
- Doing this for each of the 80 hidden nodes of the face net yields an interesting set of hybrid faces as prefered stimuli (**What are the corresponding inputs for positive contribution**???).
- Combine input faces for a subset of the hidden nodes. That subset of hidden nodes is one holon.
- Enhanced robustness: since recognition of particular features is now spread over many hidden nodes/holons, the network can still successfully recognize faces if a node or two are inoperable.
- Preferred stimuli might be teaching new features to the hidden nodes during training phase.

Applications

Driving Miss ANN

- ALVINN (Pomerleau, 1993)



- Training = watch human driver for awhile
- Testing = ANN drives alone…up to 90 miles
- Drove 98% alone across USA
- Symbolic methods for this task can't compare!

Speech Generation

<u>Given</u>: written text

<u>Produce:</u> the appropriate sounds

<u>Mapping:</u> Letters (26) ==> Phonemes (79)

<u>Context Sensitivity:</u> Makes the problem difficult.

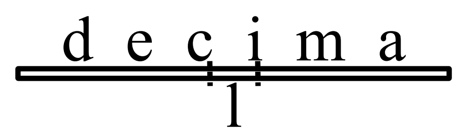
The a in ”bad” is pronounced differently than in ”car”.

English is particularly difficult in this respect.

<u>DECtalk</u>

- Digital Equipment Corporation (DEC) produced software to read books aloud - very useful for the blind.

- Uses a context window of 7 characters.



- Complex set of rules + database of exceptions

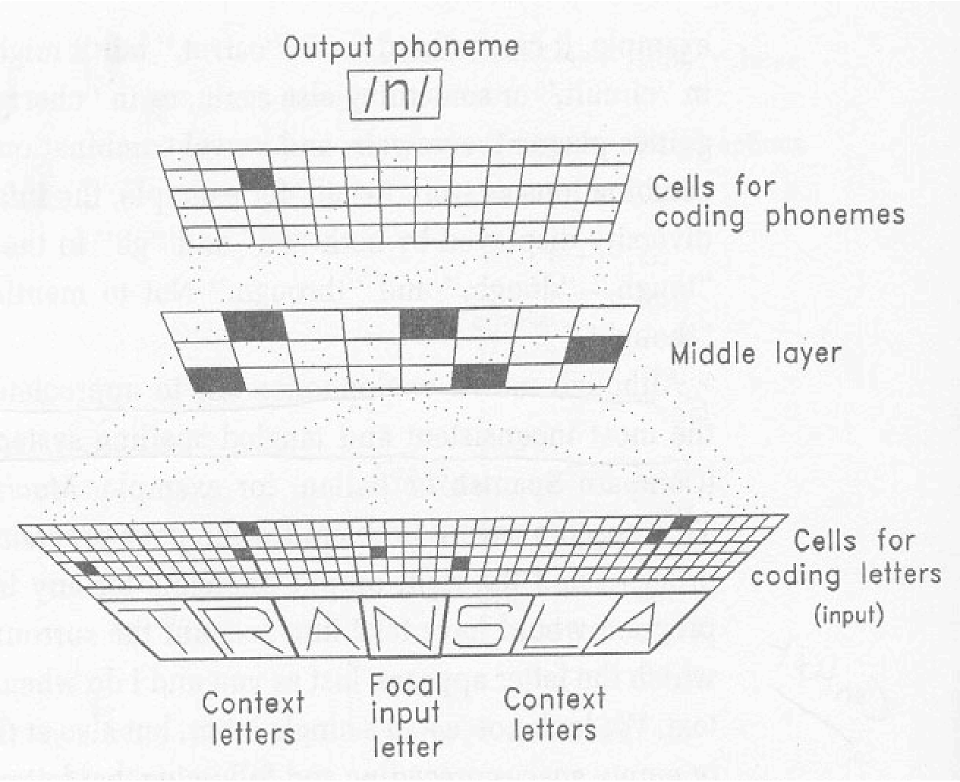
- Several <u>man years</u> of programming effort

- Classic symbolic AI approach.

NETtalk

- **Terry Sejnowski & Charles Rosenberg (1987).**
- **Simple 3-layered feed-forward network - trained with backprop**
- **Training:**
  - **Moved the 7-letter window over a 1000-word text and saved (window-contents, phoneme) pairs as the training set.**
  - **Ran epochs for <u>10 hours</u>.**
  - **95% accuracy on training set.**

- **Testing:**
  - **78% accuracy**
  - **But the speech was still highly understandable.**
  - **97.5% accuracy given larger training sets!!!**



Hebbian Learning

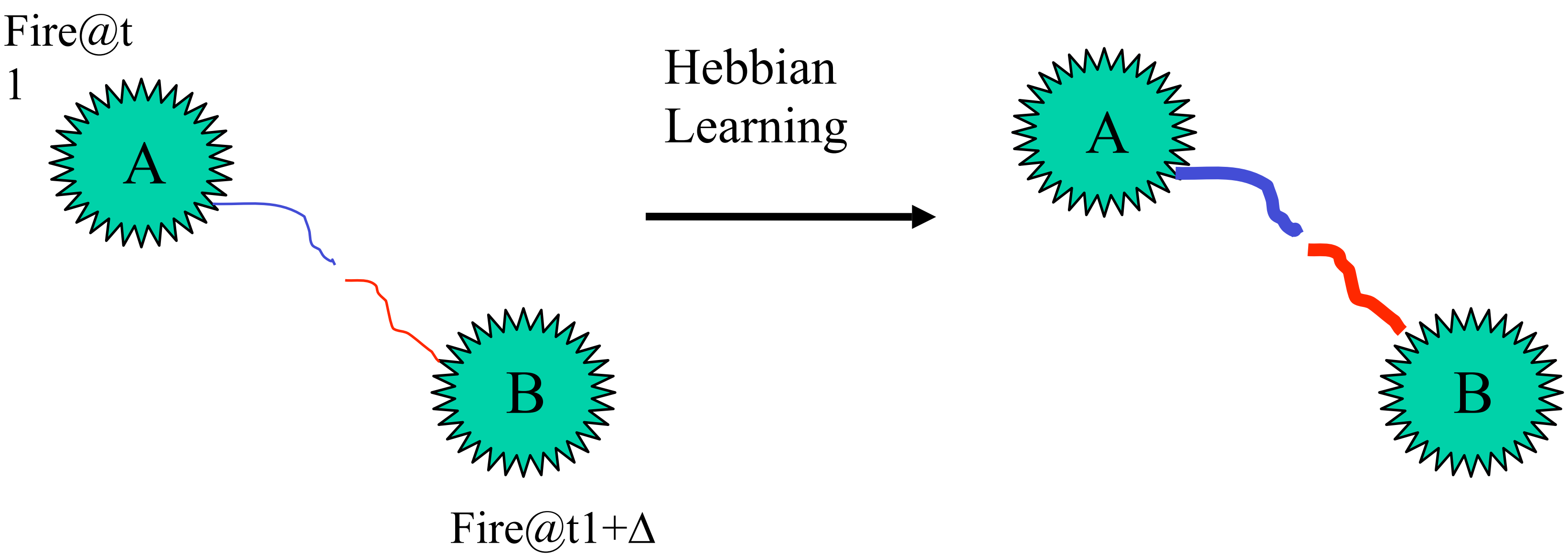
"Fire together, wire together"
(Donald Hebb, 1949)

Hebb's Rule

Correlations in Firing => Changes in Synaptic Strength

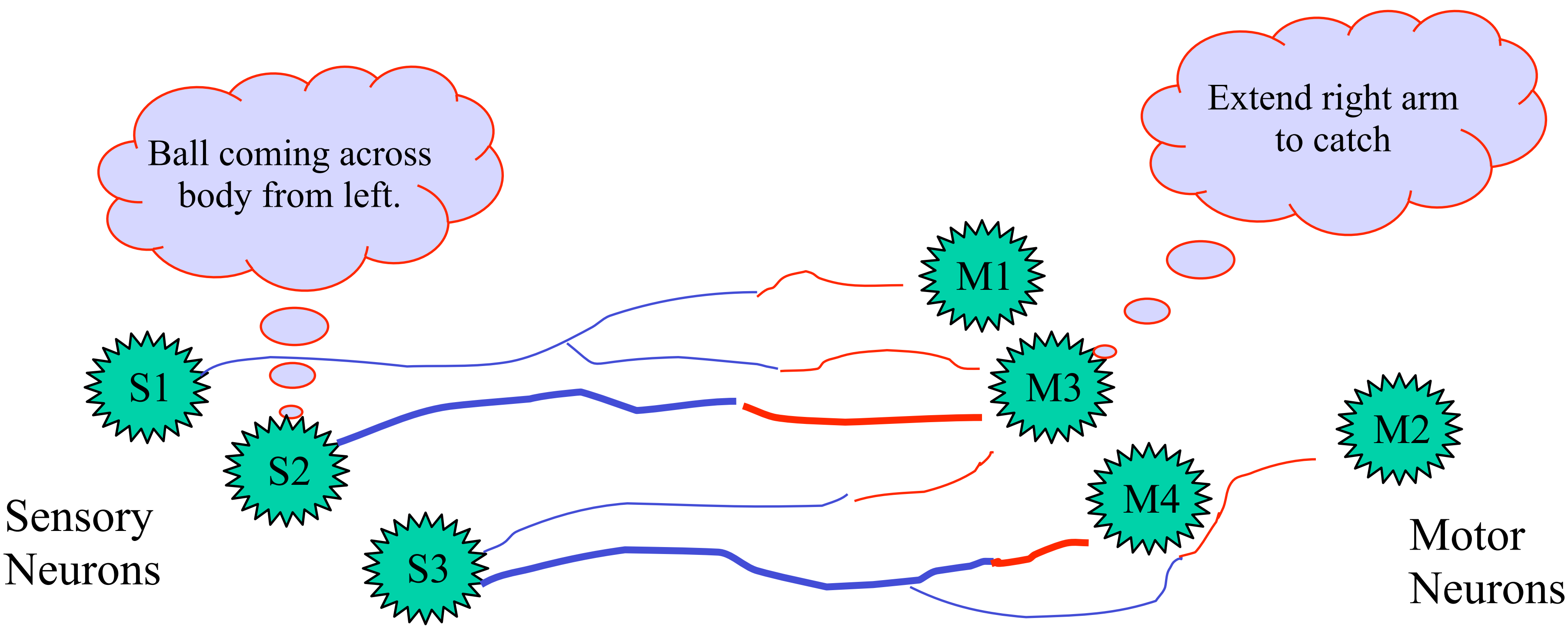
``When one cell repeatedly assists in firing another, the axon of the first cell develops synaptic knobs (or enlarges them if they already exist) in contact with the soma of the second cell." (Hebb, 1949)

``When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells, such that A's efficiency as one of the cells firing B, is increased." (Hebb, 1949)



Associative Learning

- The basic Hebbian mechanism enables the learning of associations between:
  - Portions of an image
  - Features and classes
  - Words and images
  - Sensory input (perceptions) and motor output (actions)



Ion Channels

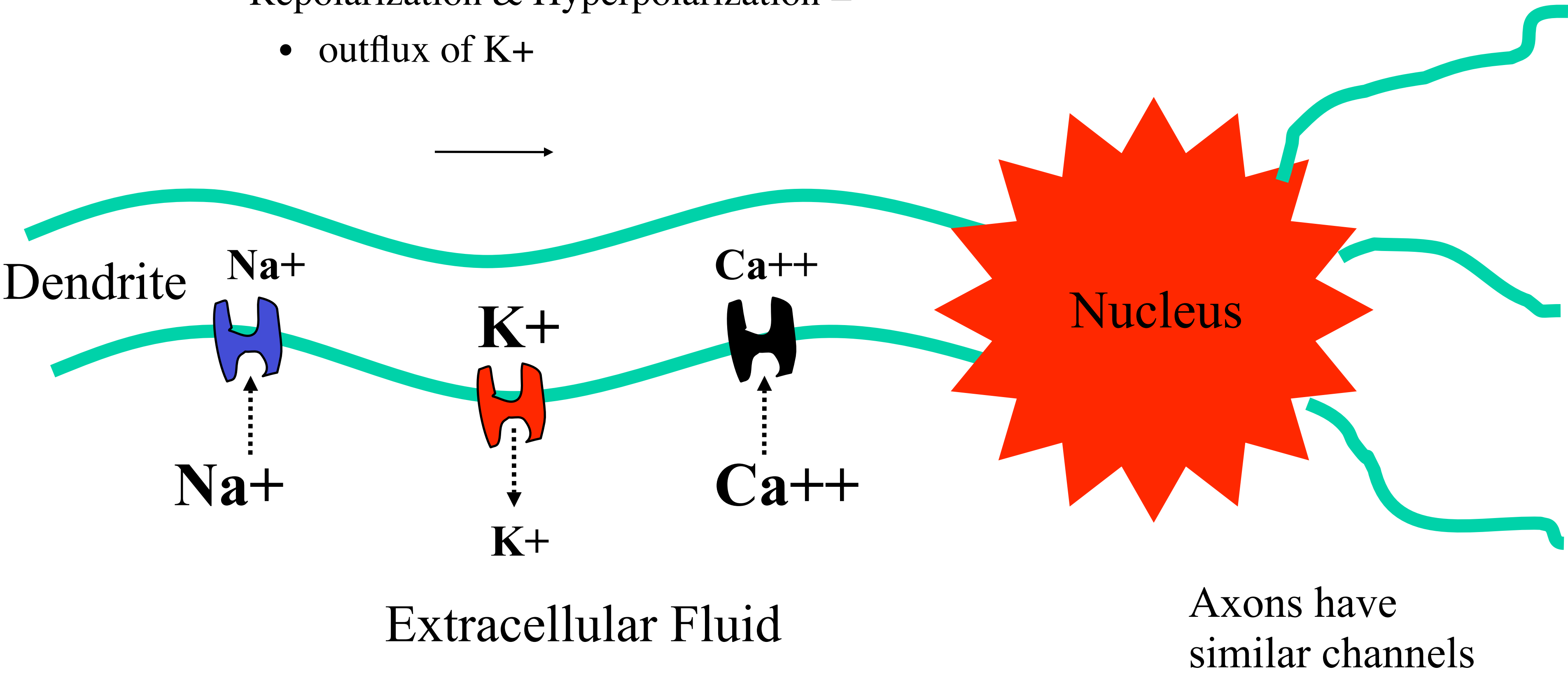
- Different ions are exchanged between the interior of a neuron and the extracellular medium (area between cells). The main ions are:
  - Potassium (K+)
  - Sodium (Na+)
  - Calcium (Ca++)
  - Chlorine (Cl-) {ignored in rest of discussion}

- Gated channels control the inflow/outflow of these ions.

- During the resting (polarized) state, there is:
  - More K+ inside than outside the cell
  - Less Na+ and Ca++ inside than outside the cell

Ion Flows

- In the standard case:
  - Depolarization =
    - influx of Na+ (along axons & dendrites)
    - influx of Ca++ (at pre-synaptic terminals)
  - Repolarization & Hyperpolarization =
    - outflux of K+

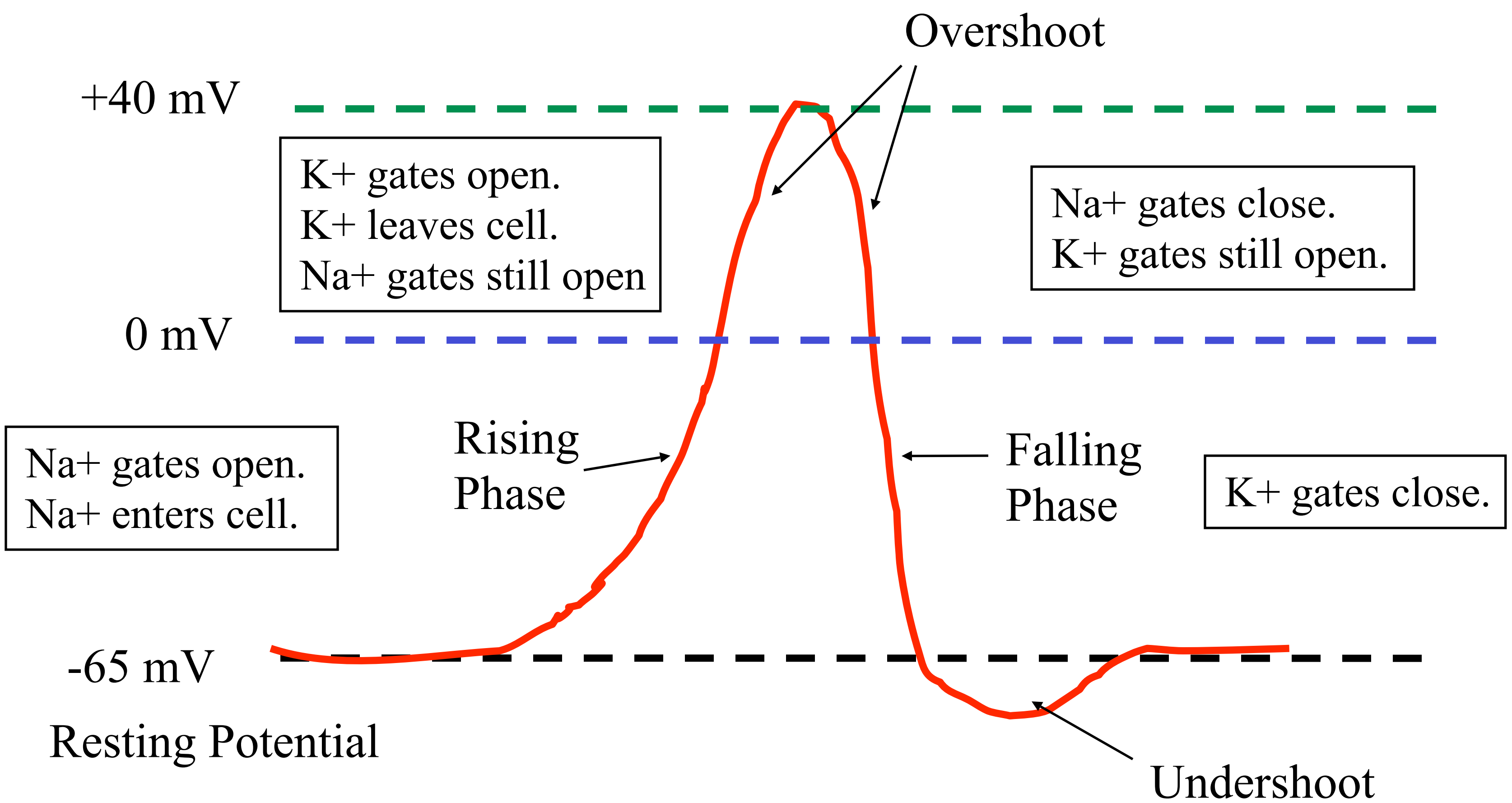


Action Potentials (AP)

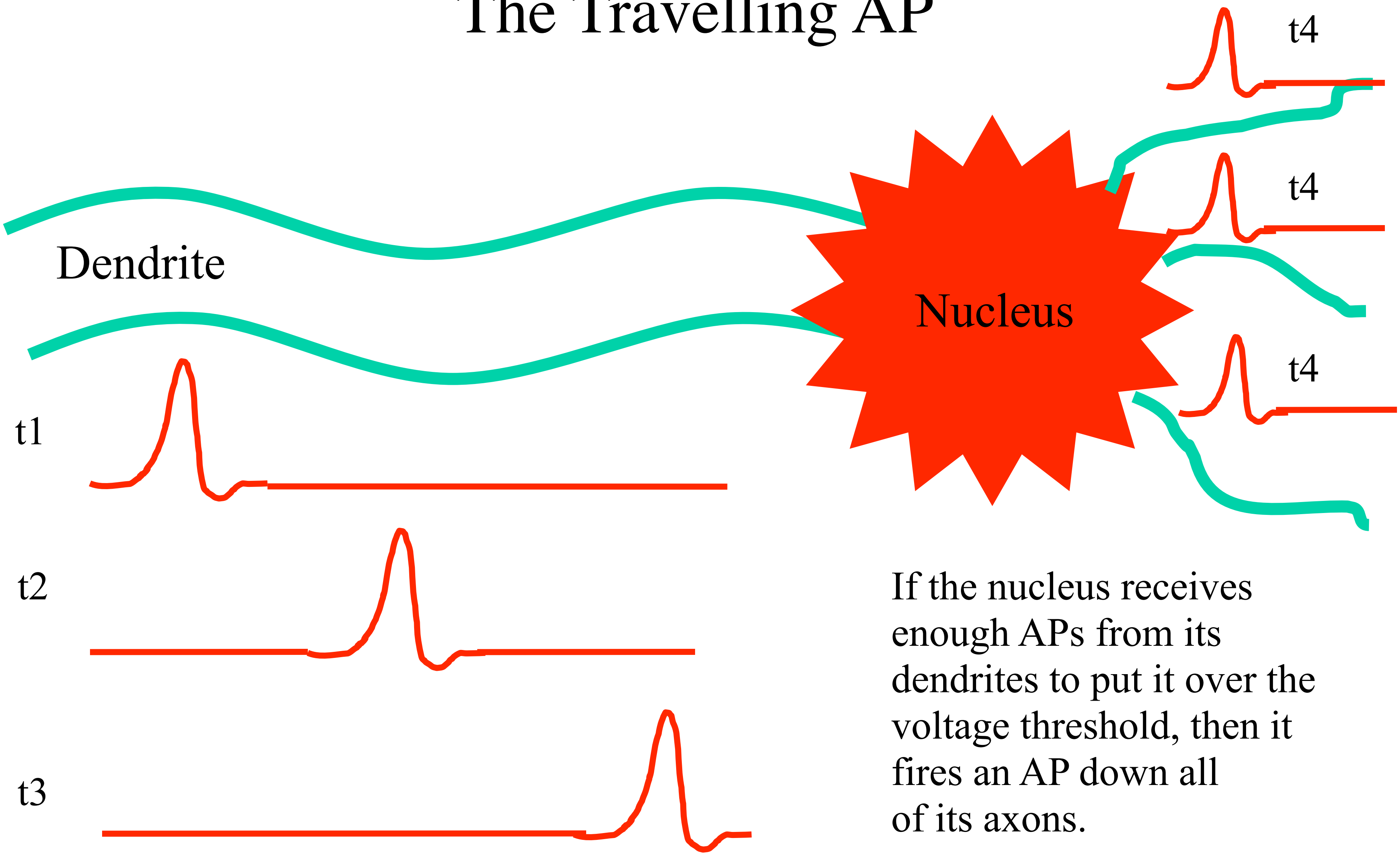
- At rest, the interior of a neuron has a net negative charge: it is **polarized**.
- Some stimuli to the nerve cell can increase the flow of **positive** ions **into** it, thus leading to **depolarization**.
- If the depolarization is strong enough and moves the net charge above a threshold, the cell fires an action potential (AP), which travels away from the cell center and down its axons.
- Conversely, some stimuli can increase the flow of positive ions **out** of the cell, thus leading to **hyperpolarization**. A hyperpolarized cell is HARDER to stimulate to the threshold level where it sends out an AP. Such as cell is **inhibited**.
- Propagating APs are a main communication mechanism between neurons.

AP = Voltage Spike

- Although the voltage of a neuron changes constantly, only large abrupt changes (action potentials) can be transmitted to other neurons.



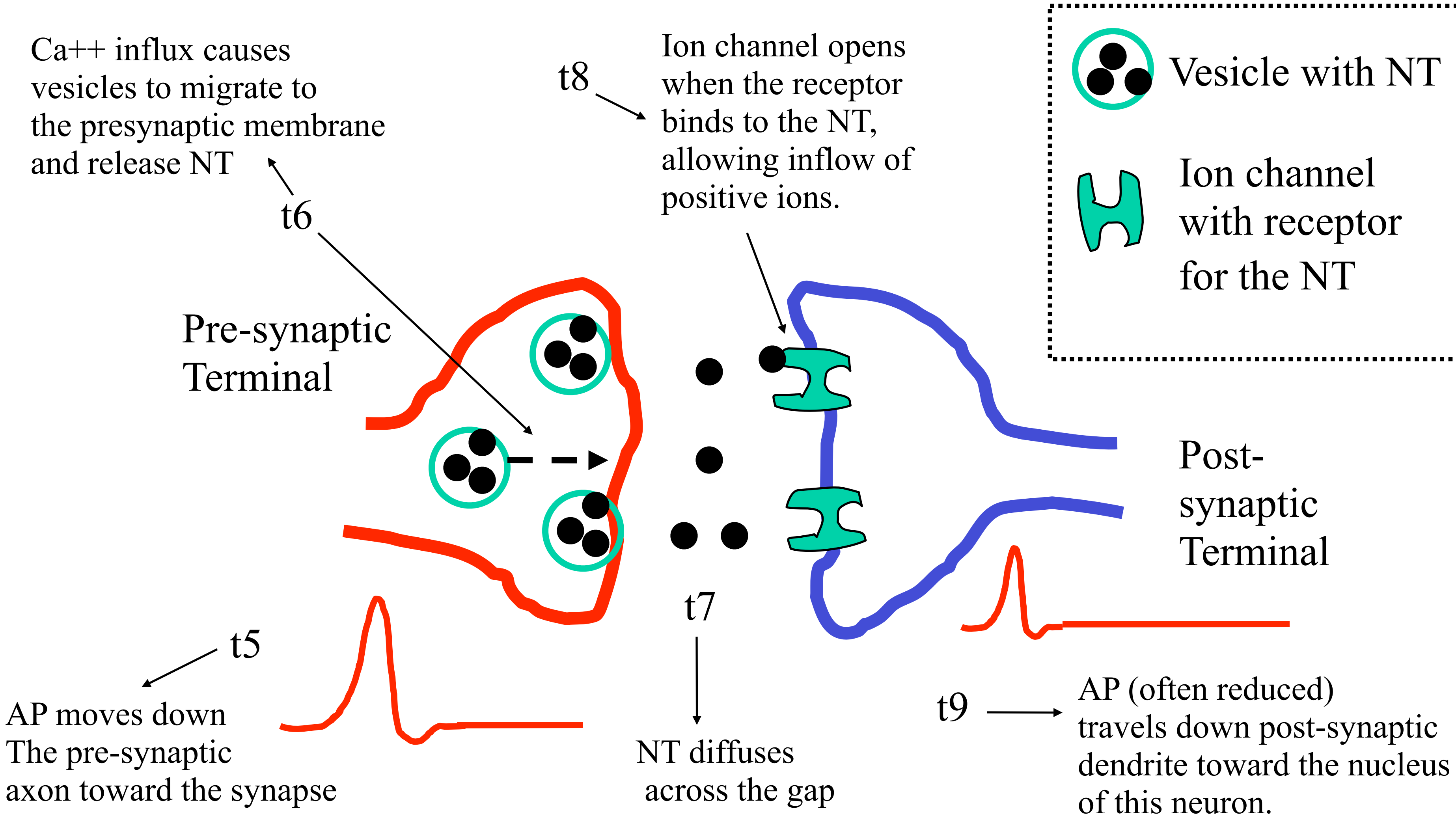
The Travelling AP



If the nucleus receives enough APs from its dendrites to put it over the voltage threshold, then it fires an AP down all of its axons.

Crossing a Chemical Synapse

Synapse = Tiny gap between two neurons (normally the axon of the pre-synaptic neuron and the dendrite of the post-synaptic neuron) across which signals (APs) are transferred by either chemical neurotransmitter (NT) or direct electrical means (**gap junctions**).

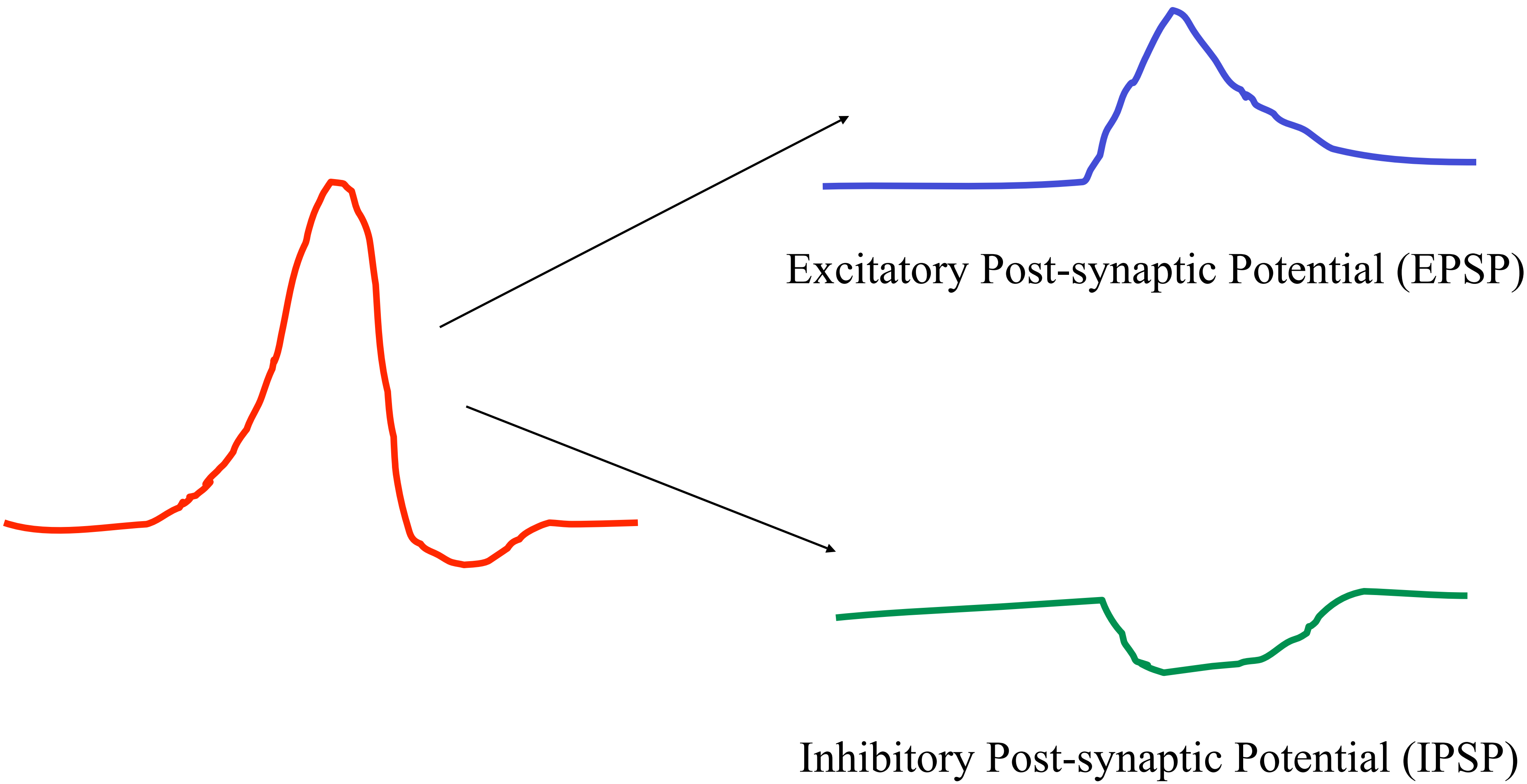


Synaptic Excitation -vs - Inhibition

- The post-synaptic receptors can gate all types of channels: K+, Na+, and Ca++, and combinations of these.
- Different neurotransmitters bind to different receptors.
- Different neuron types express different receptors and secrete different neurotransmitters.
- **<u>Excitation:</u>** If the NT is of a type that opens Na+ or Ca++ channels, then the post-synaptic neuron gets depolarized by Na+ or Ca++ influx, and the AP gets transmitted from pre- to post-synaptic neuron.
- **<u>Inhibition:</u>** But if the NT is of a type that opens K+ or CL– channels, then the post-synaptic neuron gets hyperpolarized and the AP is not transmitted.
- In <u>artificial</u> neural networks, one node can have both positive and negative weights on its outgoing arcs, but in <u>real</u> neural networks, a neuron only releases one type of neurotransmitter. Hence it can only be a promotor (excitatory) or inhibitor, **but not both**, with respect to **all** its post-synaptic neighbors.

Excitatory -vs- Inhibitory Voltage Trajectories

<u>Pre-synaptic</u> <u>Post-synaptic</u>



Synapses: The Center of Learning

- Although gap junctions are much faster at transferring signals than chemical synapes, the latter are very adaptive. This is the neurobiological basis for memory and learning.
- The pre- and post-synaptic terminals involve different learning mechanisms.

Pre-synaptic: Adenyl cyclase (AC) is key. It requires 2 conditions to activate:

5) Binding of the neuromodulator **serotonin** to the pre-synaptic terminal.
6) Inflow of Ca++ (during pre-synaptic depolarization)

Post-synaptic: NMDA receptors are key. Require 2 conditions to open:

8) Binding of the neurotransmitter glutamate.
9) Depolarization of the post-synaptic cell - this expels the Mg++ ions that block the NMDA channel in the polarized state.

(A neuromodulator is a substance other than a neurotransmitter, released by a neuron at a synapse and conveying information to adjacent or distant neurons, either enhancing or damping their activities. [Wikipedia])

Coincidence Detectors

<u>Pre-synaptic</u>

- Adenyl Cyclase (AC) detects the co-occurrence of:
  1. Depolarizing of the pre-synaptic neuron (via the Ca++ signal)
  2. A general signal of "excitement" that is broadcast to many parts of the brain via a neuromodulator (in this case, serotonin).

<u>Post-synaptic</u>

- The NMDA receptor detects co-occurrence of:
  1. Firing of the pre-synaptic neuron (which eventually leads to glutamate release by pre-synaptic vesicles).
  2. Firing of the post-synaptic neuron.
     - Note: this may require that SEVERAL pre-synaptic neurons fire.

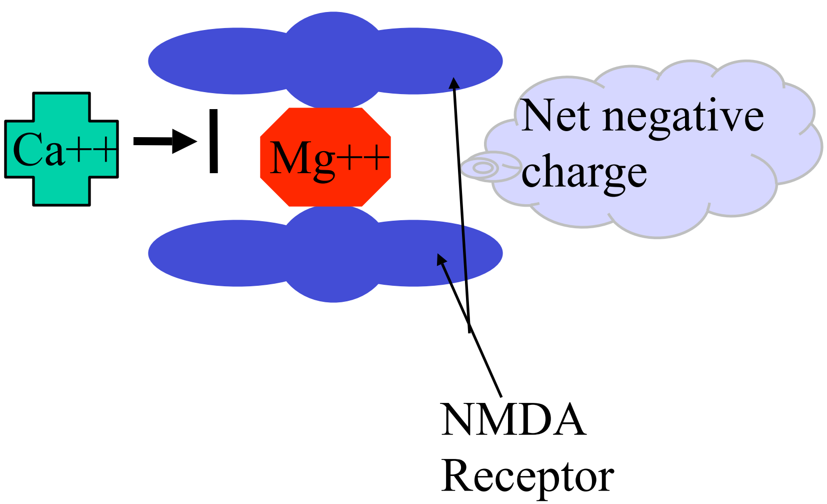
Thus, when important combinations of events occur in the world, such as the conditioned stimulus and unconditioned stimulus in a classical conditioning experiment, the brain is prepared to handle and record (i.e. remember) them!!

But the degree to which learning occurs is dependent upon the perceived excitement (i.e. interestingness) of the events.

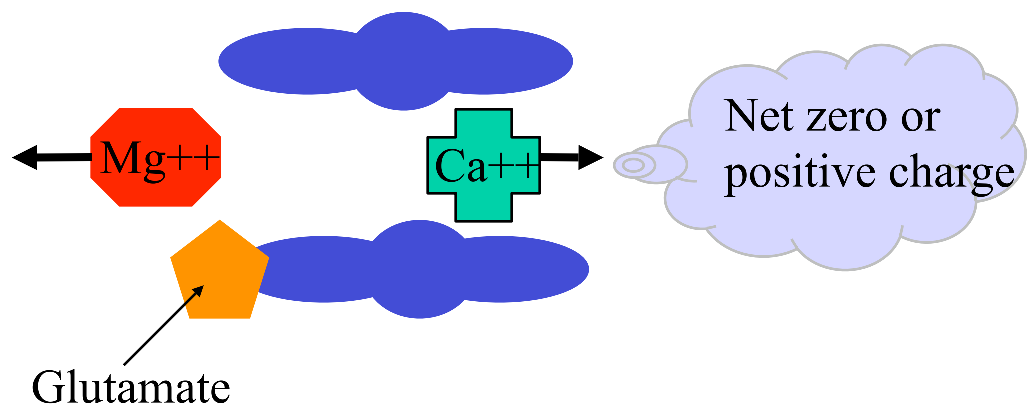
NMDA Receptors

- During the polarized state, the net charge of the neuron is negative, so this attracts the Mg++ ion, holding it in the NMDA receptor, where it blocks ion transfer.
- When the neuron becomes depolarized, the Mg++ ion is expelled.
- If, in addition, glutamate binds to the receptor, then Ca++ ions can enter.
- NMDA receptors are the main entry point for Ca++.
- Ca++ is an important **second messenger** => its presence in the cell can lead to long-term changes in the way it responds to APs (i.e. learning)

Polarized (Relaxed) State



Depolarized (Firing) State



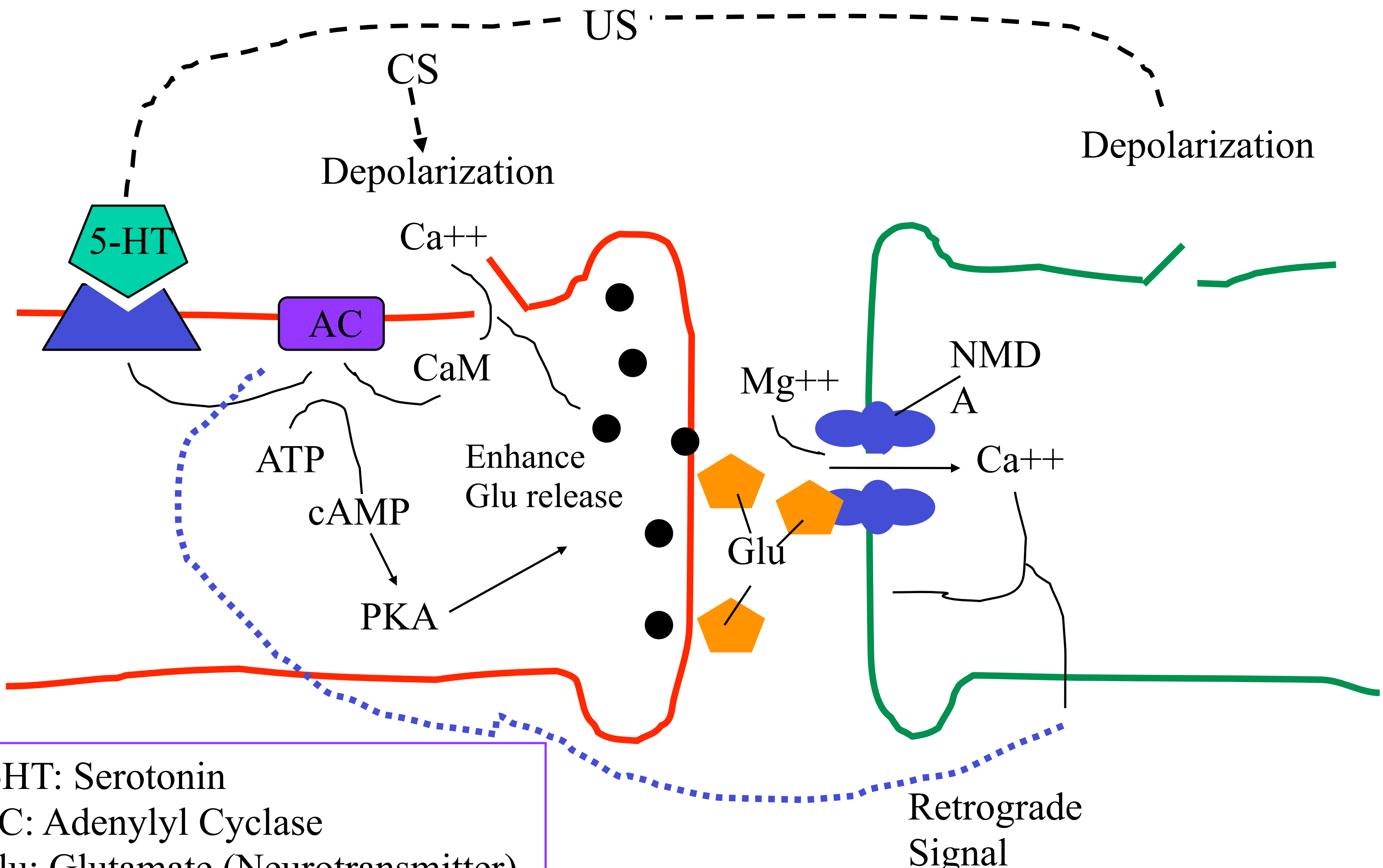
Hebb's Rule in NMDA Receptors

- "Fire together, wire together": if a pre- and post-synaptic neuron tend to fire at the same time (or the pre fires slightly before the post), then the strength of synapse between them should increase.
- <u>Fire together:</u> When pre and post fire together, then NMDA channels "detect the coincidence" and open. Ca++ then flows into the post-synaptic terminal.
- <u>Wire together:</u> Ca++ in the post-synaptic terminal stimulates the production of CaMKII and protein kinase C. From here, the details are unclear, but it appears that these:
  - Can have effects upon the K+ channels (thereby prolonging APs).
  - Can lead to the production of more non-NMDA receptors in the post-synaptic terminal.
  - Can produce retrograde chemical signals that diffuse back to the pre-synaptic cell and FURTHER increase its production of neurotransmitter.

Serotonin

- Increased serotonin + (Ca++ inflow to pre-synaptic terminal) =>
- Increased Adenylyl Cyclase activity =>
- Increased cyclic AMP (cAMP) =>
- Increased activity of protein kinase A =>
- Increase protein phosphorylation =>
- **Reduced efficiency of K+ channels** =>
- Decreased K+ outflow during repolarization =>
- Prolonged depolarization phase =>
- Longer action potentials (APs) =>
- More time for Ca++ inflow to the pre-synaptic terminal =>
- Increased amount of neurotransmitter released (since Ca++ promotes vesicles to migrate and release neurotransmitter)
- Since the change to the K+ channels can be long-lasting, the pre-synaptic terminal changes so that all future APs are prolonged
- In short, the neural network has learned that the information encoded by the pre-synaptic neuron is significant in some way. So in the future, that same neuron will have a stronger effect upon its post-synaptic neighbors by <u>secreting more neurotransmitter</u> when it transmits an AP.
- Note: Serotonin is also important in keeping us awake!

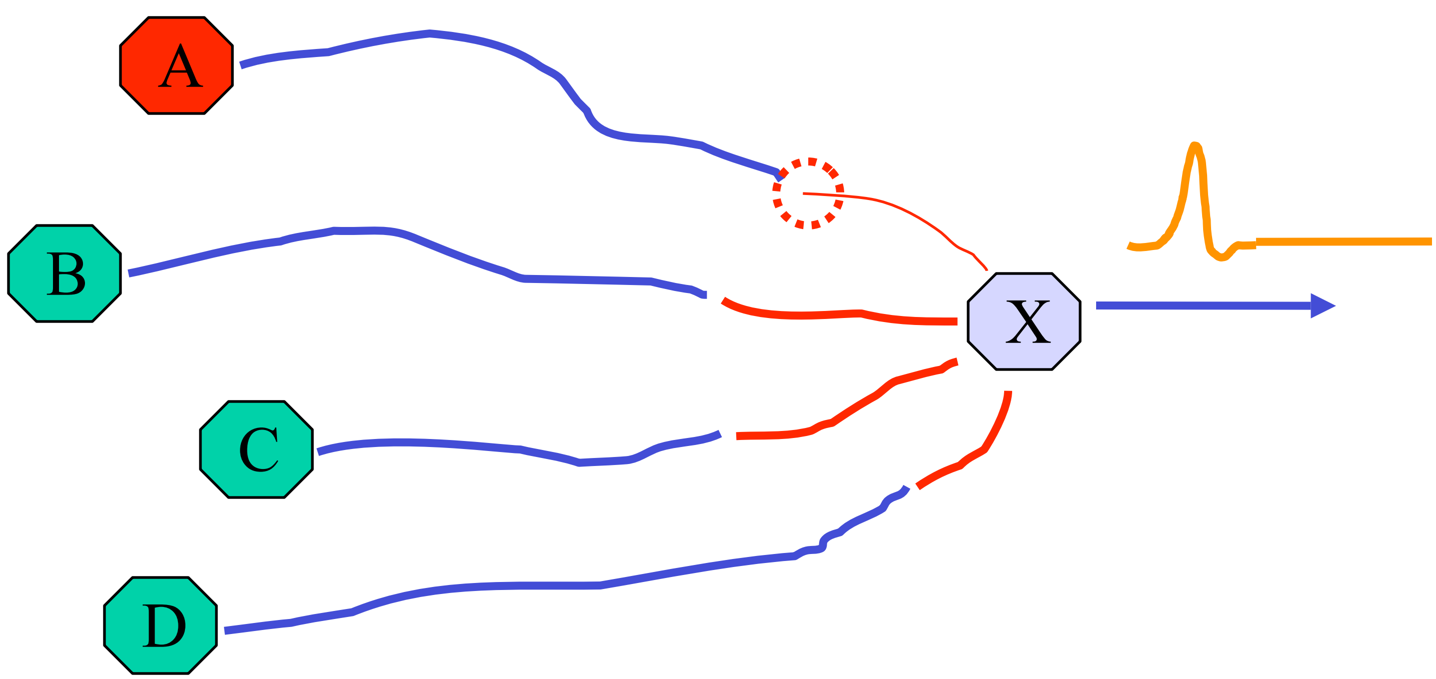
Associative Learning Overview (Carew, 2000)



5-HT: Serotonin
AC: Adenylyl Cyclase
Glu: Glutamate (Neurotransmitter)
US: Unconditioned Stimulus
CS: Conditioned Stimulus

Cooperation & Association

- It often takes many inputs (i.e. APs from pre-synaptic neurons) to cause the post-synaptic neuron to fire.
- Once X fires, then the A-X synapse can be strengthened by Hebbian means
- But, A-X might not have strengthened without help from B, C and D.
- However, if A-X strengthens enough, then, in the future, A alone may be sufficient to stimulate X.

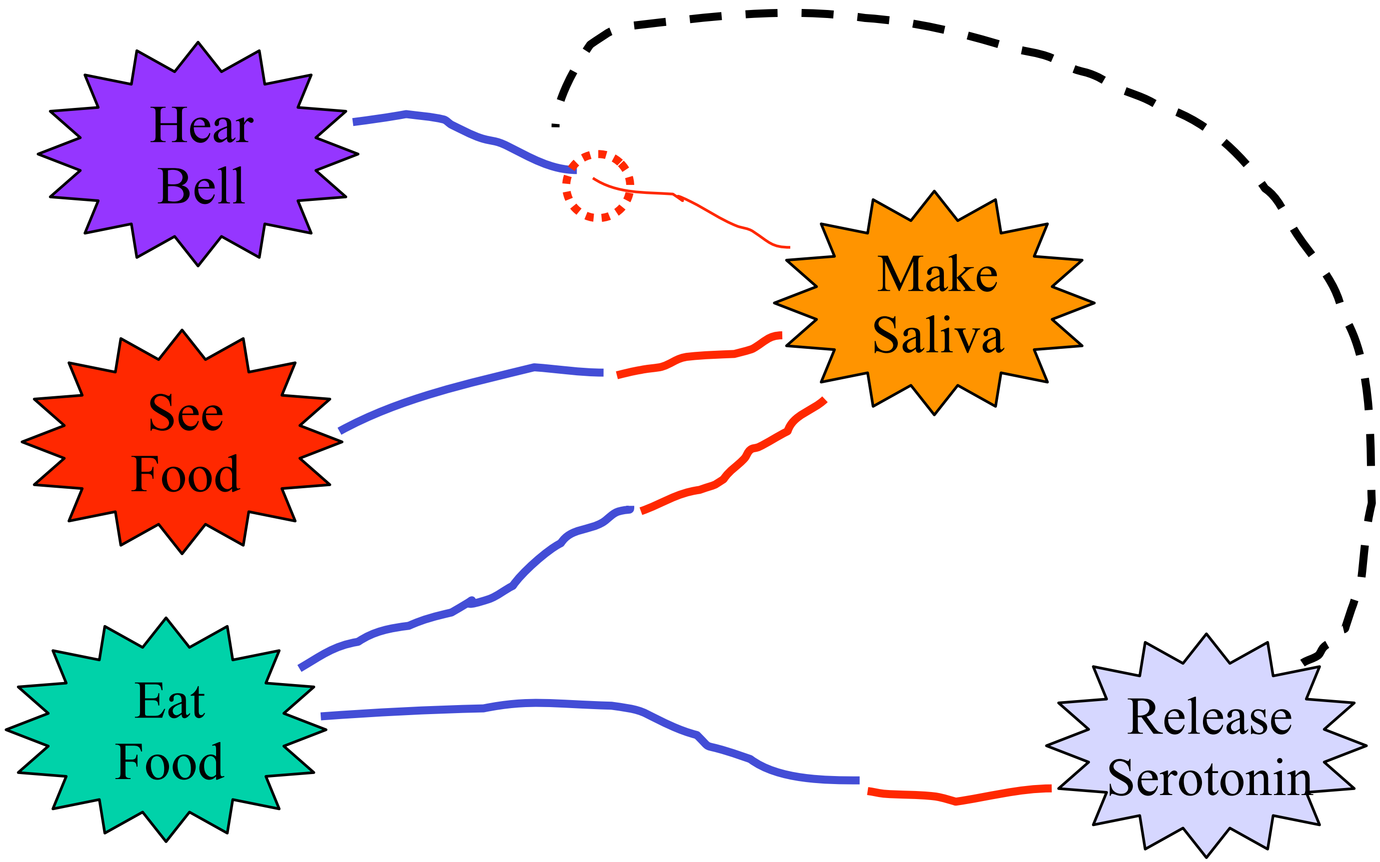


Classical Conditioning

- Basic Approach:
  - Initially: an unconditioned stimulus (US) causes an unconditioned response (UR).
  - Training: Pairing of the US with a conditioned stimulus (CS).
  - Testing: Eventually, the CS (alone) can elicit the same response, although now it is called the conditioned response (CR).
- Classic example: Pavlov's dog
  - US = sight of food
  - UR/CR = salivation
  - CS = ringing of bell
- A form of associative learning, since the organism learns to associate the CS and US.
- Can we understand these associations in terms of Hebbian learning at the neural level?

Classically-Conditioned Synapses

- The synapses from seeing & eating food to salivating are strong (often due to innate wiring).
- Eating is a very pleasurable (hence significant) experience, causing neuromodulator release (in this case, serotonin)



Hebbian Firing & Wiring

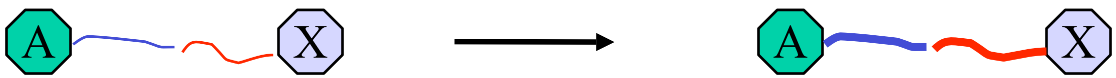
Fire together: When the bell rings and food is seen (and then eaten), salivation fires due to the strength of the see- and eat-salivate synapses. So the neurons for a) hearing a bell, and b) salivating fire at about the same time.

Wire together:

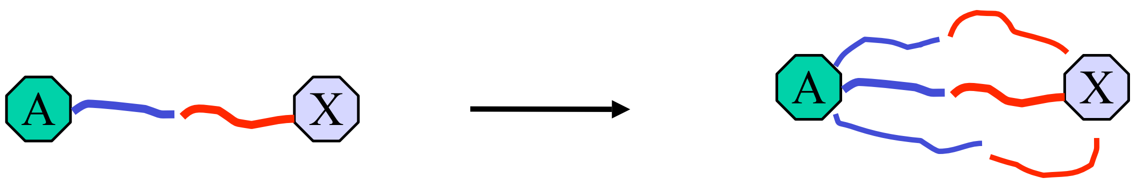
- Pre-synaptic: Serotonin is secreted because eating is a significant event, and Ca++ enters the hear-bell neuron's axons when it depolarizes/fires. So AC detects the coincidence of a) hearing the bell, and b) a significant event (via serotonin).
- Post-synaptic: In the salivation neuron, the NMDA receptors detect the coincidence of a) its own depolarization, and b) a glutamate signals coming from the see-food and eat-food neurons.
- Both coincidences increase the synaptic strength between the hear-bell neuron and the salivation neurons.

Long-Term Potentiation (LTP)

- LTP = The strengthening of a synapse.
  - Early Stage: Chemical changes in pre-synaptic and post-synaptic terminals due to AC and NMDA. Duration: minutes to hours.



  - Late Stage: Anatomical changes in synapses (via growth of new axons and/or dendrites) due to protein synthesis in pre- and/or post-synaptic neurons.



- LTD = Long-term depression = the weakening of a synapse (similar but opposite of LTP)