The ShanghAI Lectures

人工智能

上海授课

The ShanghAI Lectures by the University of Zurich
An experiment in global teaching

Today from EPFL, Lausanne (Switzerland), University of Osaka (Japan), and Salford University (UK)

10 November 2011

欢迎您参与

“来自上海的人工智能系列讲座”

Lecture 6

Building brains for bodies: Artificial Neural Networks

10 November 2011

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 上海人工智能授课

Implementation of learning in embodied systems

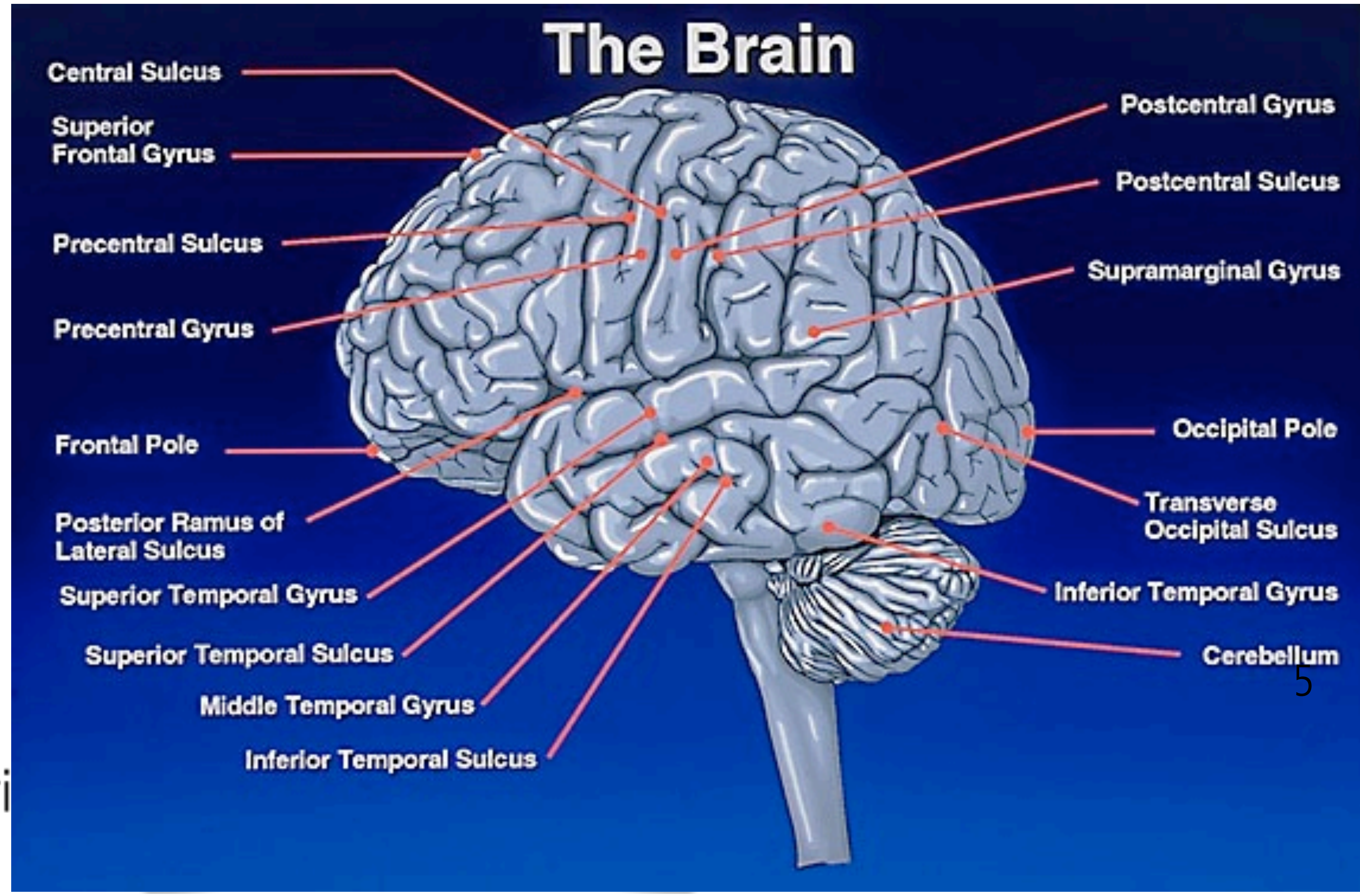
important approach:

"Artificial Neural Networks"

Building brains and bodies

Implementing learning in embodied systems for adaptive behavior: Artificial neural networks (focus box 5.1, p. 156, "How the body …")

Inspiration: The brain



Brain vs. Computer

- asynchronous
- temporarily synchronized
- massively parallel
- many (relatively!) simple processors
- no separation of data and programs
- typically distributed
- individual neurons relatively unreliable
- fault tolerant
- energy efficient
- forgetting
- high variability in performance
- intrinsic learning

- clocked
- globally synchronized
- few, highly complex processors
- separation of data and programs
- storage locations
- reliable, extremely fast
- not fault tolerant
- high energy consumption
- storage highly reliable
- constant performance
- learning can be programmed

Brain vs. Computer (complete agent)

- embedded in complete agent
- rich sensory input
- direct exposure to sensory stimulation
- etc.

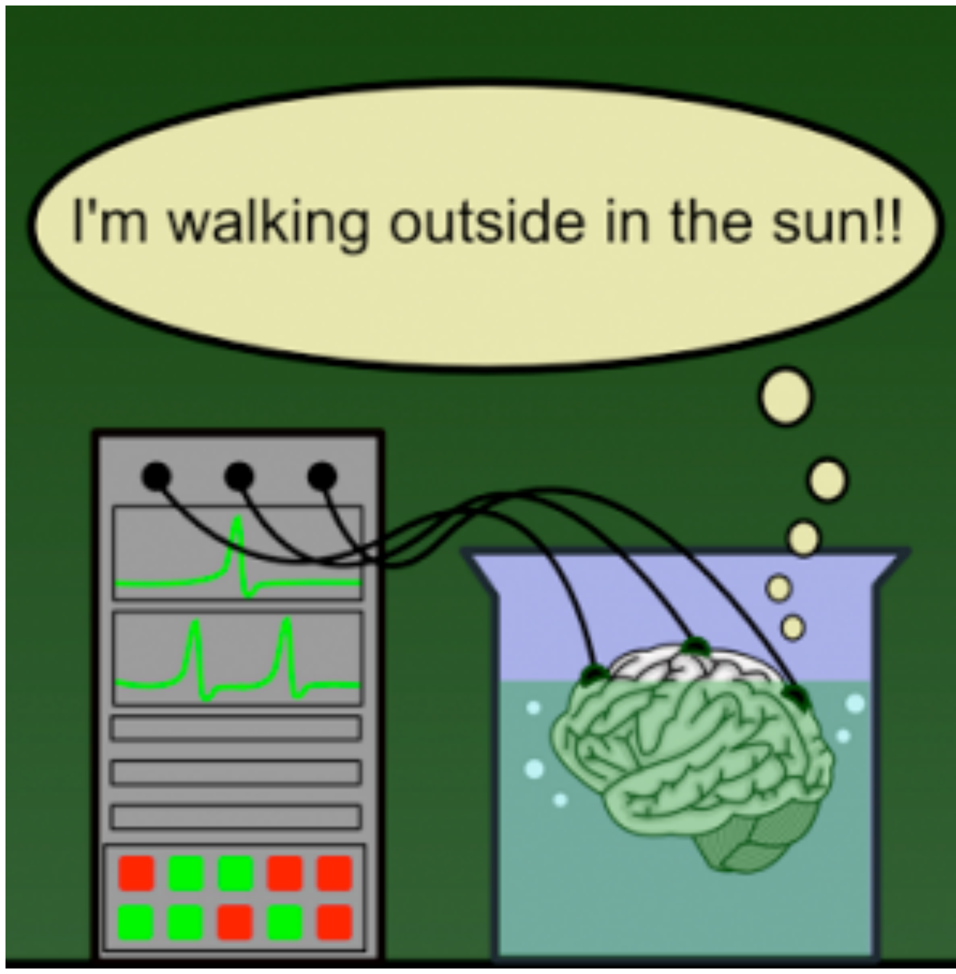
- disembodied (but cf. embedded systems)
- impoverished sensory input
- often pre-processed by humans (pushing key on keyboard)
- etc.

Brains never occur in isolation in biological systems. What about the “brain in the vat?”

Let me be clear

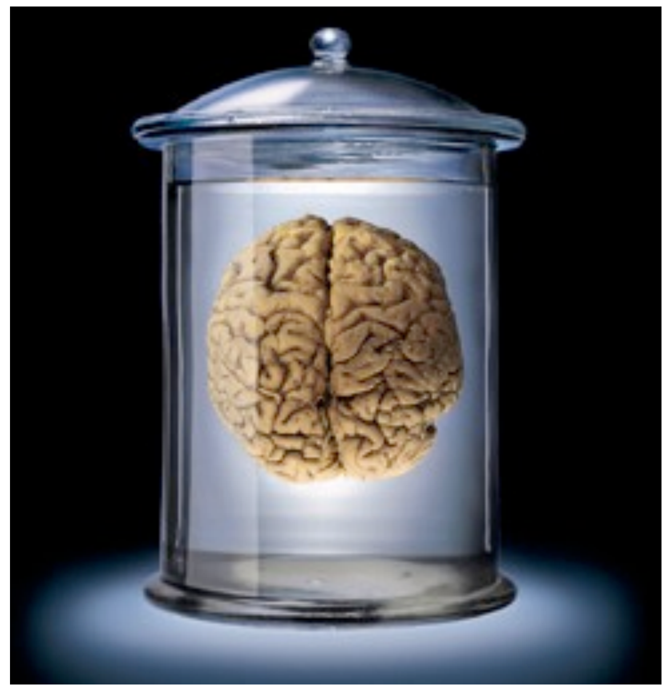
The brain is important!

but not the whole story ...

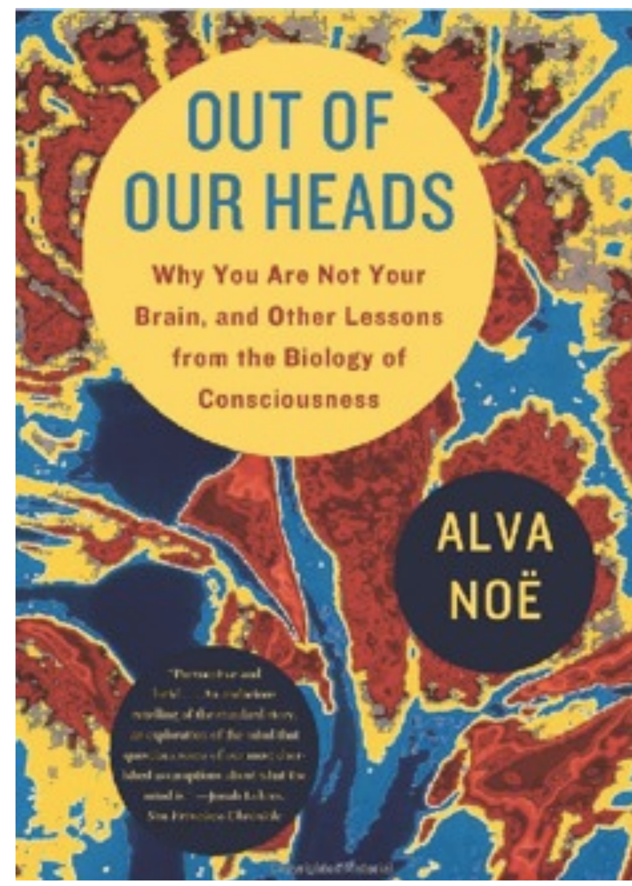


"Brain-in-a-vat"

Alva Noë, "Out of our heads - why you are not your brain", New York, Hill and Wang, 2009



- **supply energy**
- **flush away waste products**
- **complicated: providing stimulation comparable to that normally provided to a brain by its environmentally situated body**
- **—> vat will have to be something like a living body**



"Consider, first of all, that the vat or petri dish, couldn't be a mere dish or bucket as Evan Thompson and Diego Cosmelli have discussed in an essay. It would have to supply energy to nourish the cells' metabolic activity and it would have to be capable of flushing away waste products. The vat would have to be very complicated and specialized in order to control the administration of stimulation to the brain comparable to that normally provided to the brain by its environmentally situated body. If you actually try to think through the details of this thought experiment – this is something scientists and philosophers struck by the brain-in-a-vat idea almost never do – it's clear that the vat would have to be, in effect, something like a living body." (Alva Noe, Out of our heads, p. 12/13).

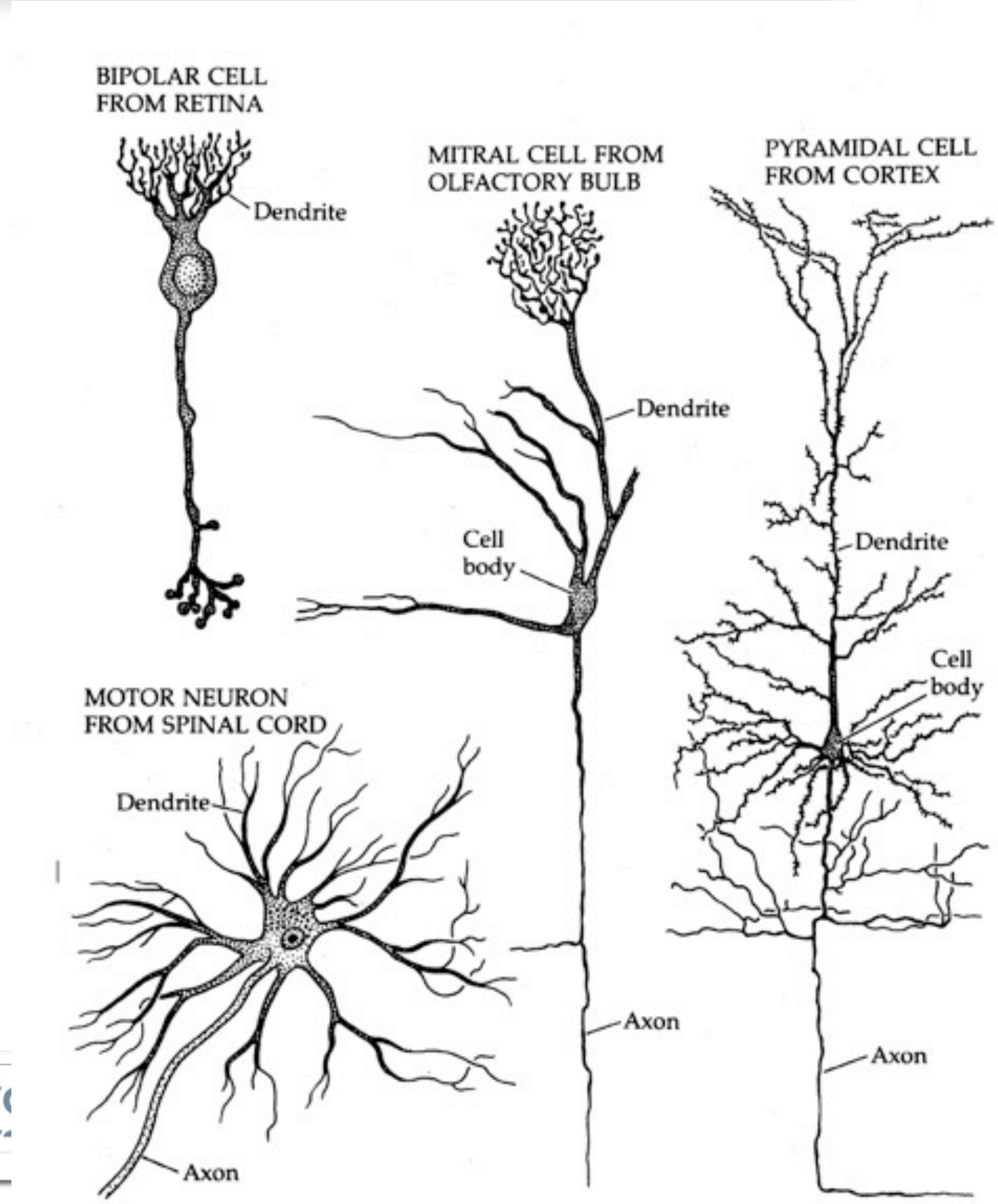
Alva Noë about the "brain-in-a-vat"

"Consider, first of all, that the vat or petri dish, couldn't be a mere dish or bucket as Evan Thompson and Diego Cosmelli have discussed in an essay. It would have to supply energy to nourish the cells' metabolic activity and it would have to be capable of flushing away waste products. The vat would have to be very complicated and specialized in order to control the administration of stimulation to the brain comparable to that normally provided to the brain by its environmentally situated body. If you actually try to think through the details of this thought experiment - this is something scientists and philosophers struck by the brain-in-a-vat idea almost never do - it's clear that the vat would have to be, in effect, something like a living body." (Alva Noe, Out of our heads, p. 12/13).

Neurons in brain

many different types of neurons

→

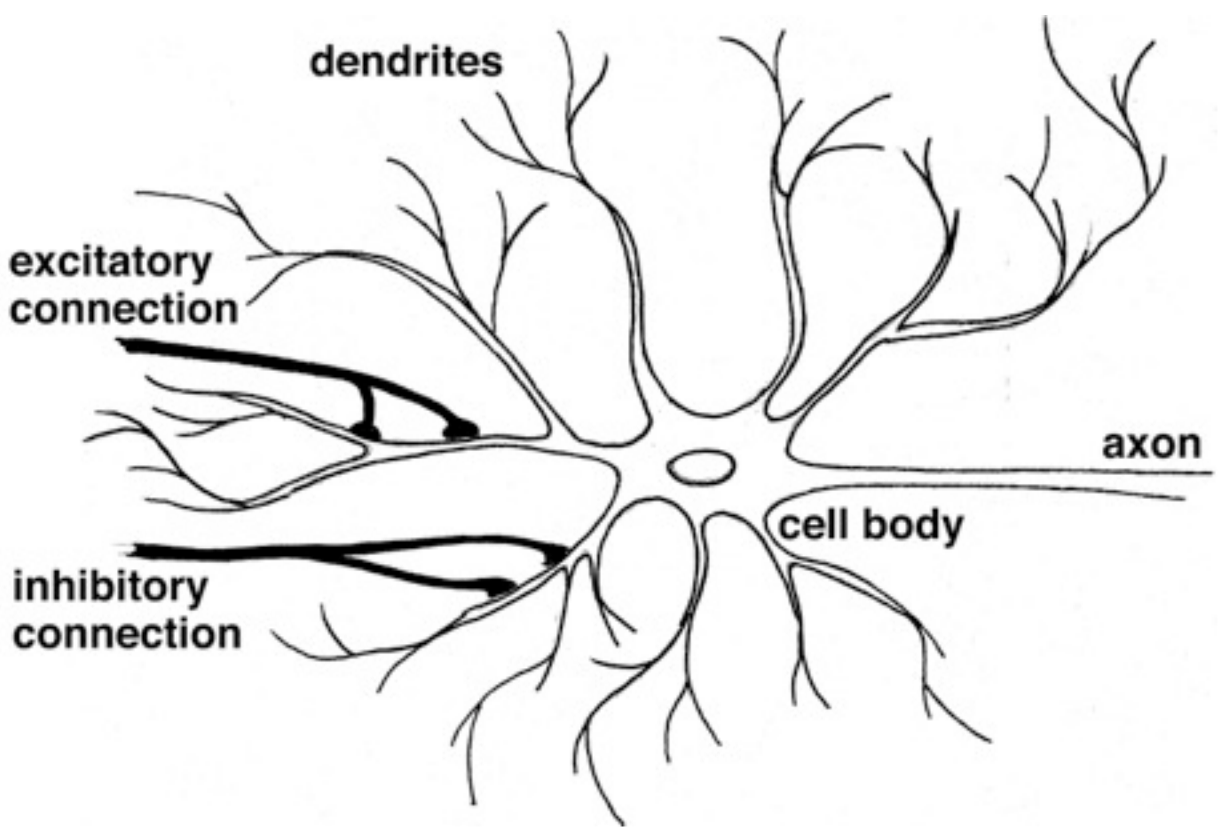


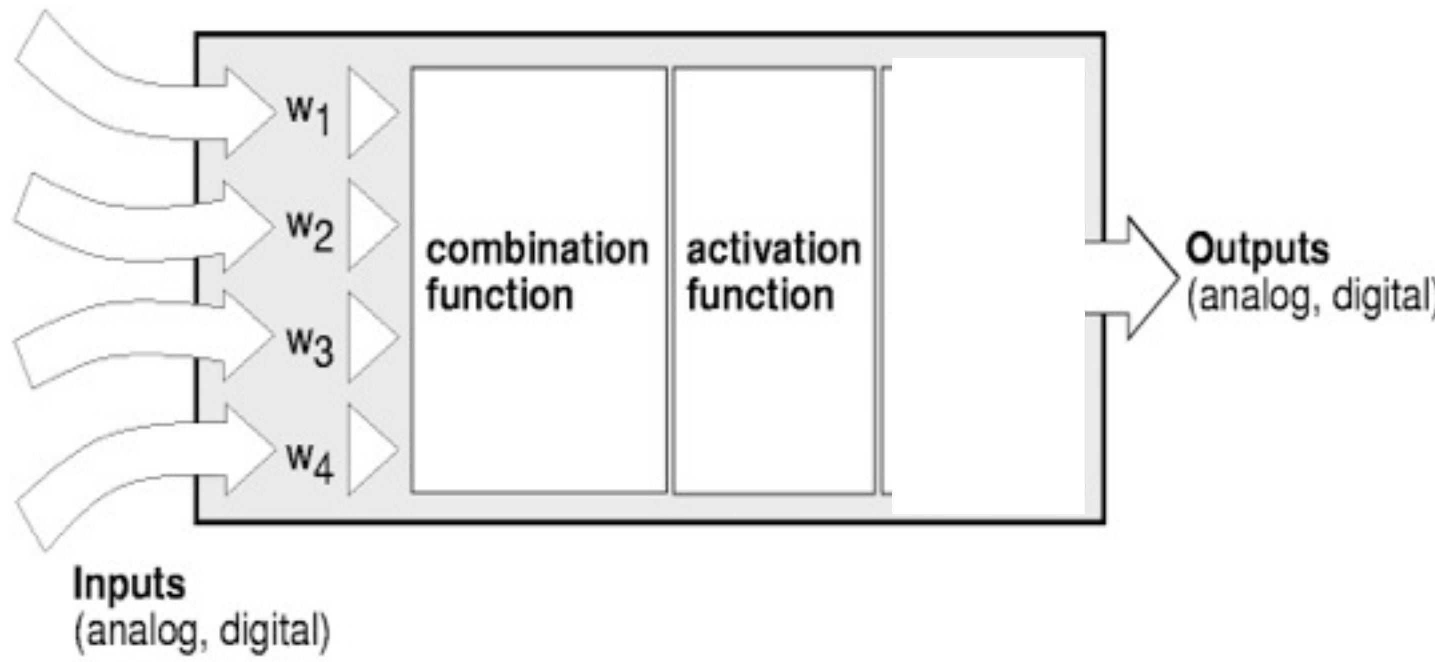
University of Zurich

complexity of individual neurons considerable and large variety
—> abstractions required

From biological neurons to abstract models

abstract artificial neurons: simple but very powerful





From biological neurons to abstract models

some important abstractions:

abstract artificial neurons: simple but very powerful

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能上海授课

Important abstractions:
activation level (a number, modeling roughly spiking frequency)
central clock
typically: chemicals ignored (but important for biological neurons)
pseudo-parallelism (simulation on largely sequential processor)
growth largely ignored/ degeneration ignored

Neural networks for robotics

popular:

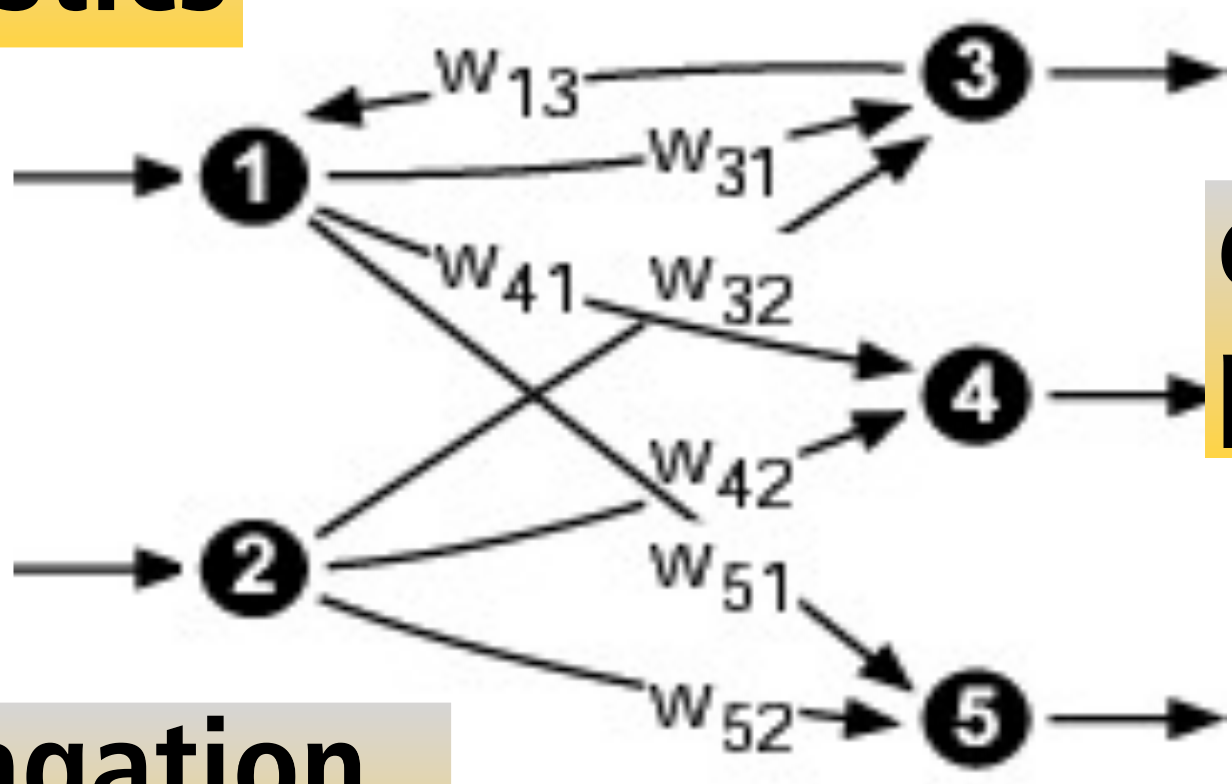
some nice properties

Ability for learning and generalization (Q: Why is generalization important?)
Graceful degradation
Coping with non-linearities
Well-investigated
Easy to build and manipulate

Designing a network: The "five basics"

node characteristics

network connectivity



embedding in physical agent

propagation rule

learning rule

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

人工智能 The ShanghAI Lectures 上海授课

Read focus box in chapter 5 of "How the body …" and attend the class on Neural Networks in the Spring term 2011 (more in-depth treatment).

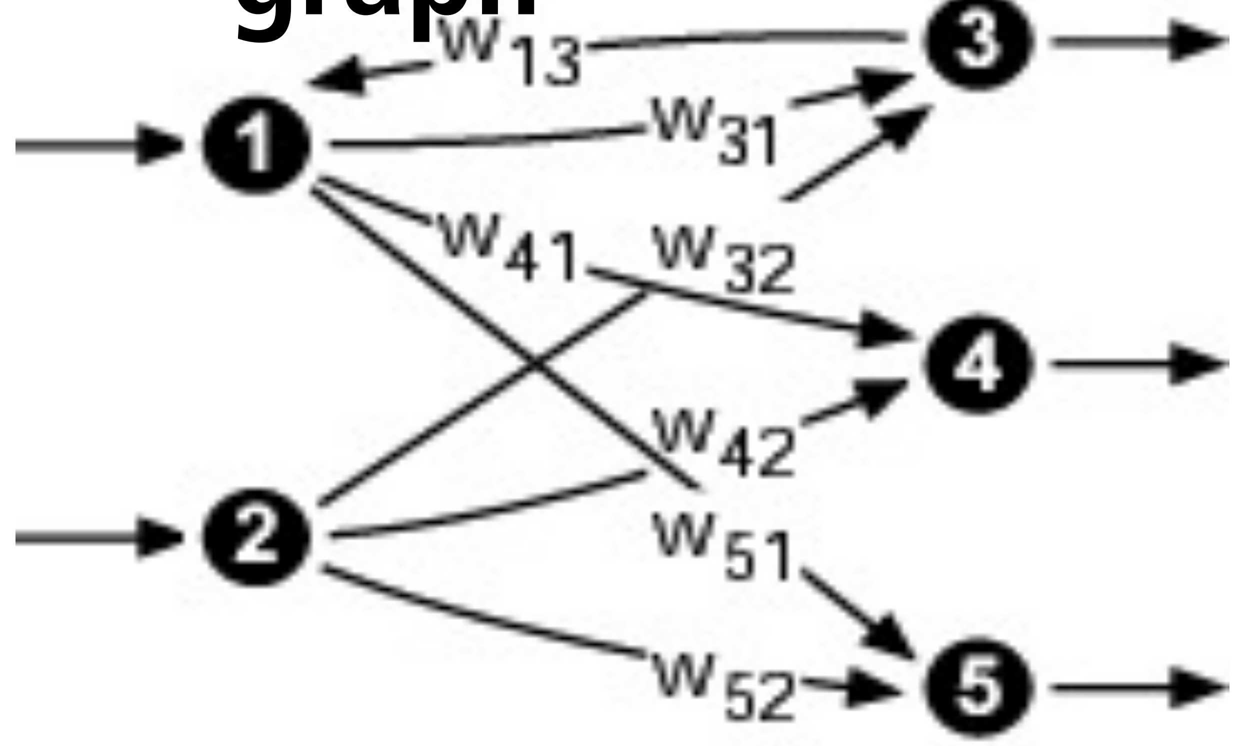
Representing connectivity

	node 1	node 2	node 3	node 4	node 5
node 1	0	0	0.8	0	0
node 2	0	0	0	0	0
node 3	0.7	0.4	0	0	0
node 4	1.0	-0.5	0	0	0
node 5	0.6	0.9	0	0	0

matrix

feedforward network with one recurrent connection

graph



Case study: Embedded artificial neural network

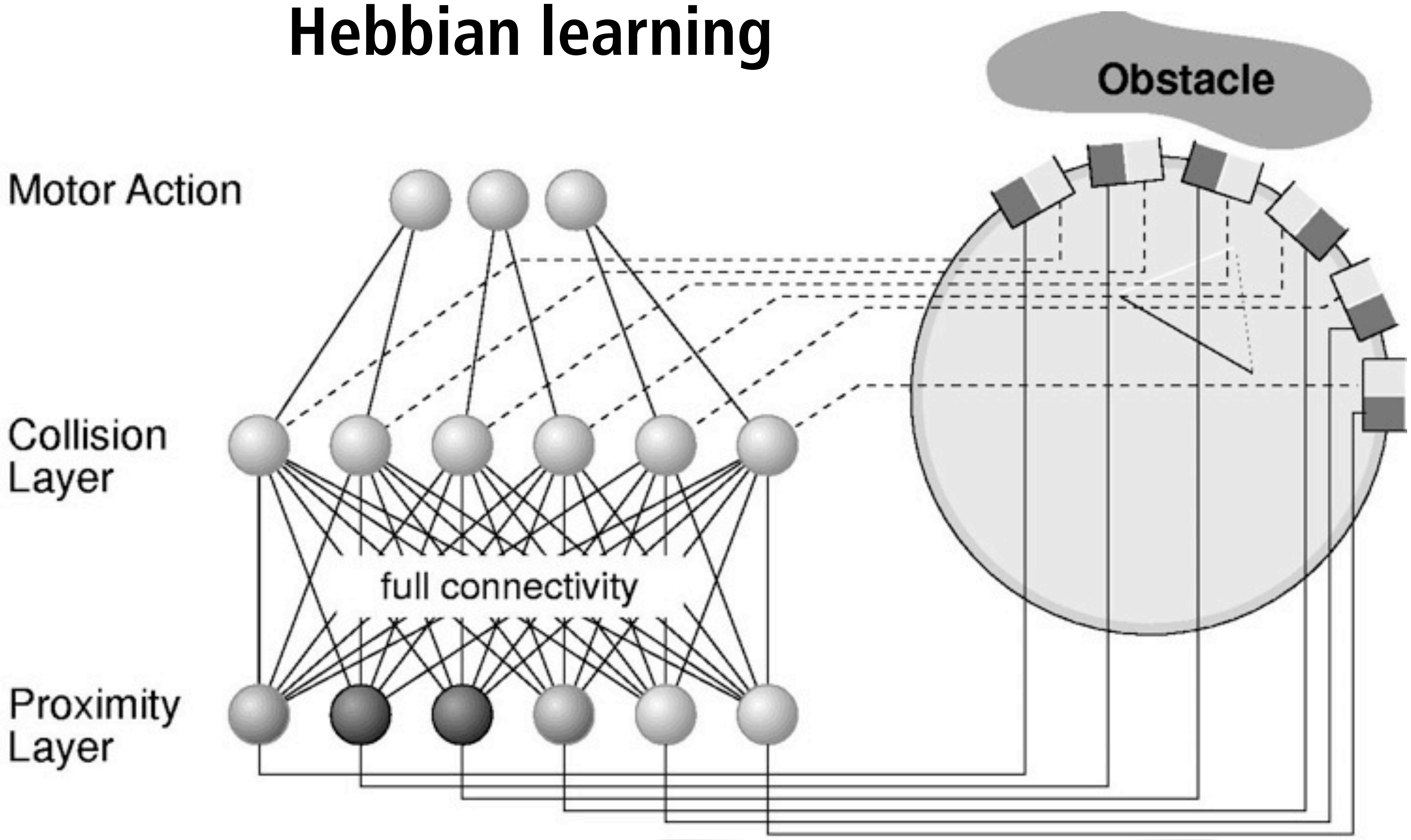
DAC: Distributed Adaptive Control

- initial design decisions:
- goal: find “food” (light source), learn to avoid obstacles, follow walls
- sensors (collision, proximity, light), actuators, default forward speed
- basic reflexes: collision on left, turn right (and vice versa), turn towards light source
- embedding of network

sensors: collision, light, proximity
actuators: two motors, one left, one right
--> Khepera robot
need to simulate “collision sensor” using IR sensor

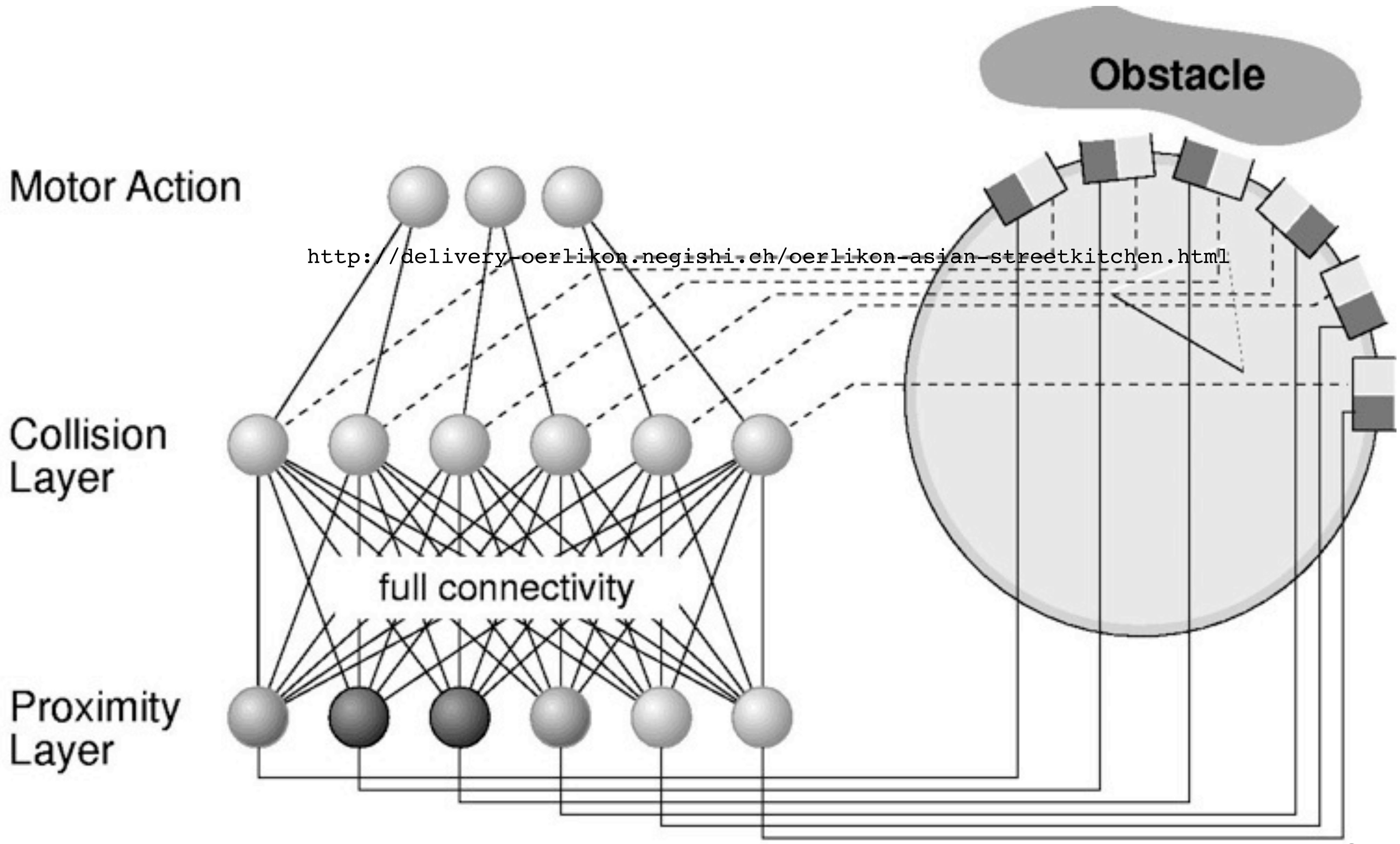
Controlling a robot with an artificial neural network

- **collision layer to motor action layer: pre-wired**
- **proximity layer to collision layer: modifiable by Hebbian learning**



collision and proximity sensors on the robot, located at the edge of the robot in pairs (of collision-proximity sensor)
fully connected feed-forward network from proximity to collision layer;
the connections from the collision layer to the motor action neurons are hard-wired and cannot be changed: they implement the three basic reflexes (reflexes are part of the value system, because we, as designer, think it is useful for the robot to be equipped with these reflexes).
Hebbian learning: simultaneous activation —> reinforce connection strength;
slogan: fire together - wire together

Environment: Arena with obstacles



http://delivery-oerlikon.negishi.ch/oerlikon-asian-streetkitchen.html

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

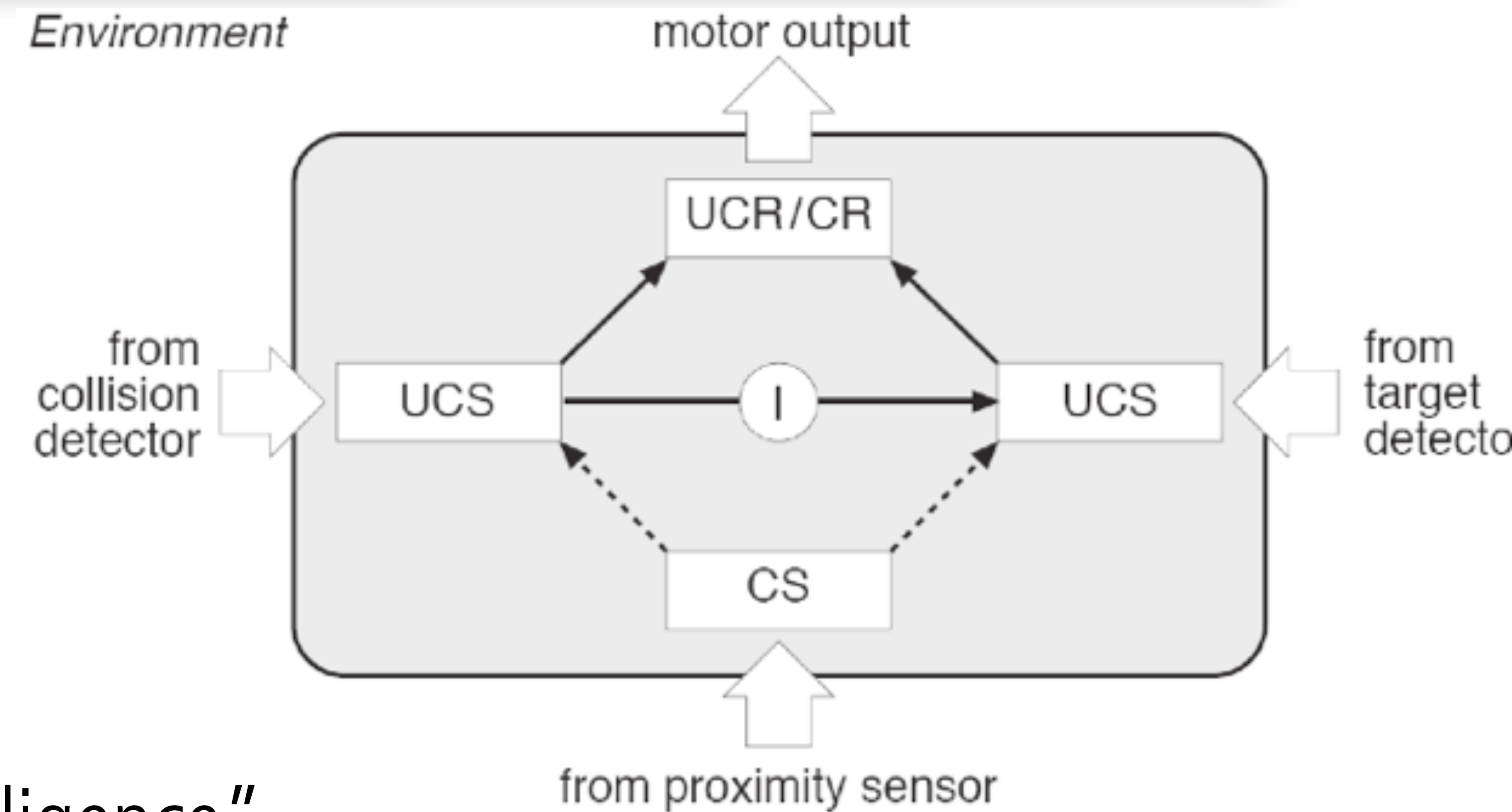
Case study: Embedded artificial neural network

DAC: Distributed Adaptive Control

- initial design decisions:
- goal: find “food” (light source), learn to avoid obsta…
- sens…
  default forward speed
- basic reflexes: collision on left, turn right (and vice versa), turn towards light source
- embedding of network

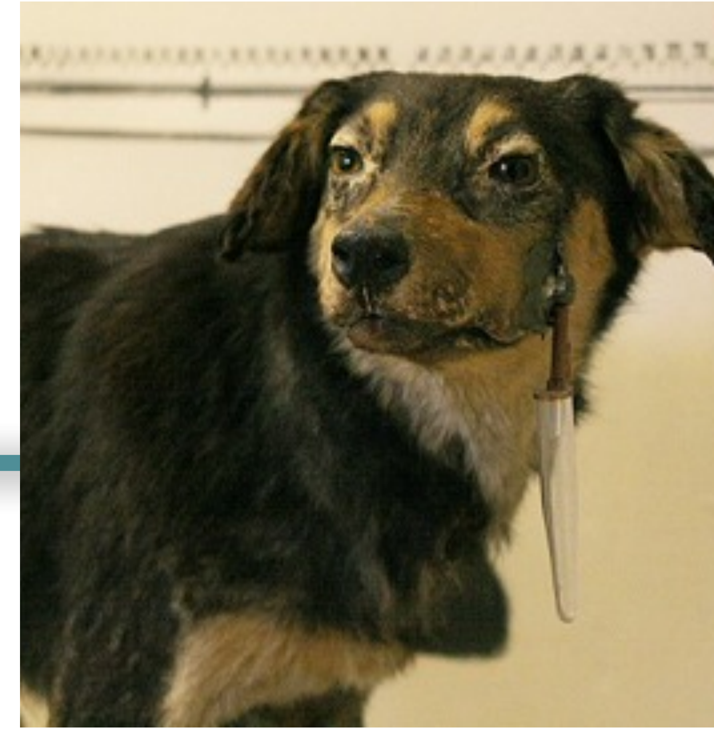
Careful: no “high-level” concepts! (“Symbol Grounding”)

DAC as a model of classical conditioning



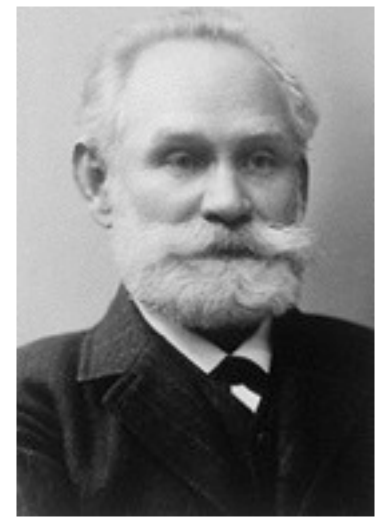
From "Understanding Intelligence"
- unconditioned stimuli: collision and light sensors (UCS)
- conditioned stimuli: proximity sensors (CS)
- unconditioned response: motor response (UCR and CR are identical)

Case study: Pavlov's classical conditioning

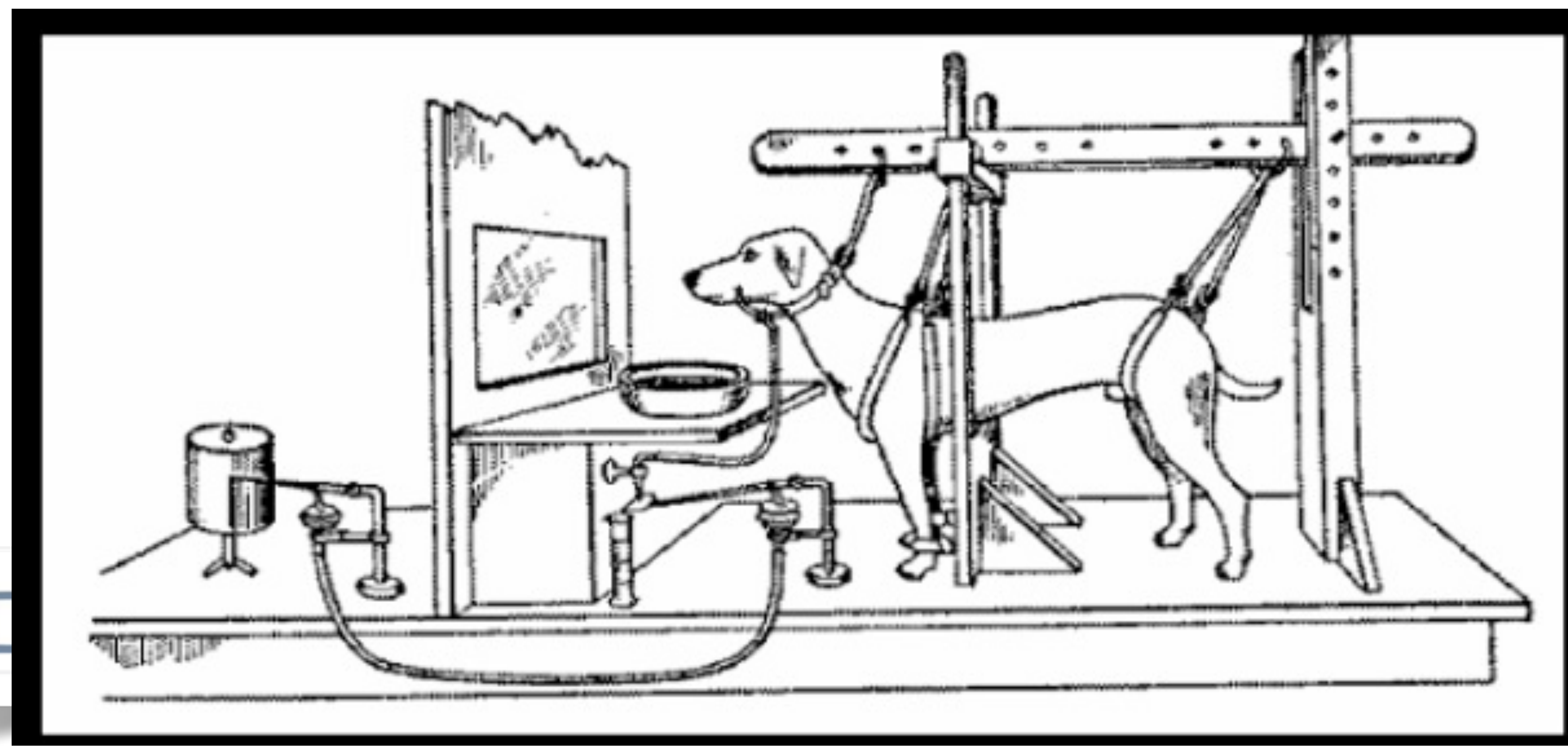


learning from well-defined experience

- unconditioned stimulus (US, e.g. food)
- conditioned stimulus (CS, e.g. sound of bell)
- unconditioned response (UR, e.g. salivation)
- conditioned response (CR, e.g. salivation)



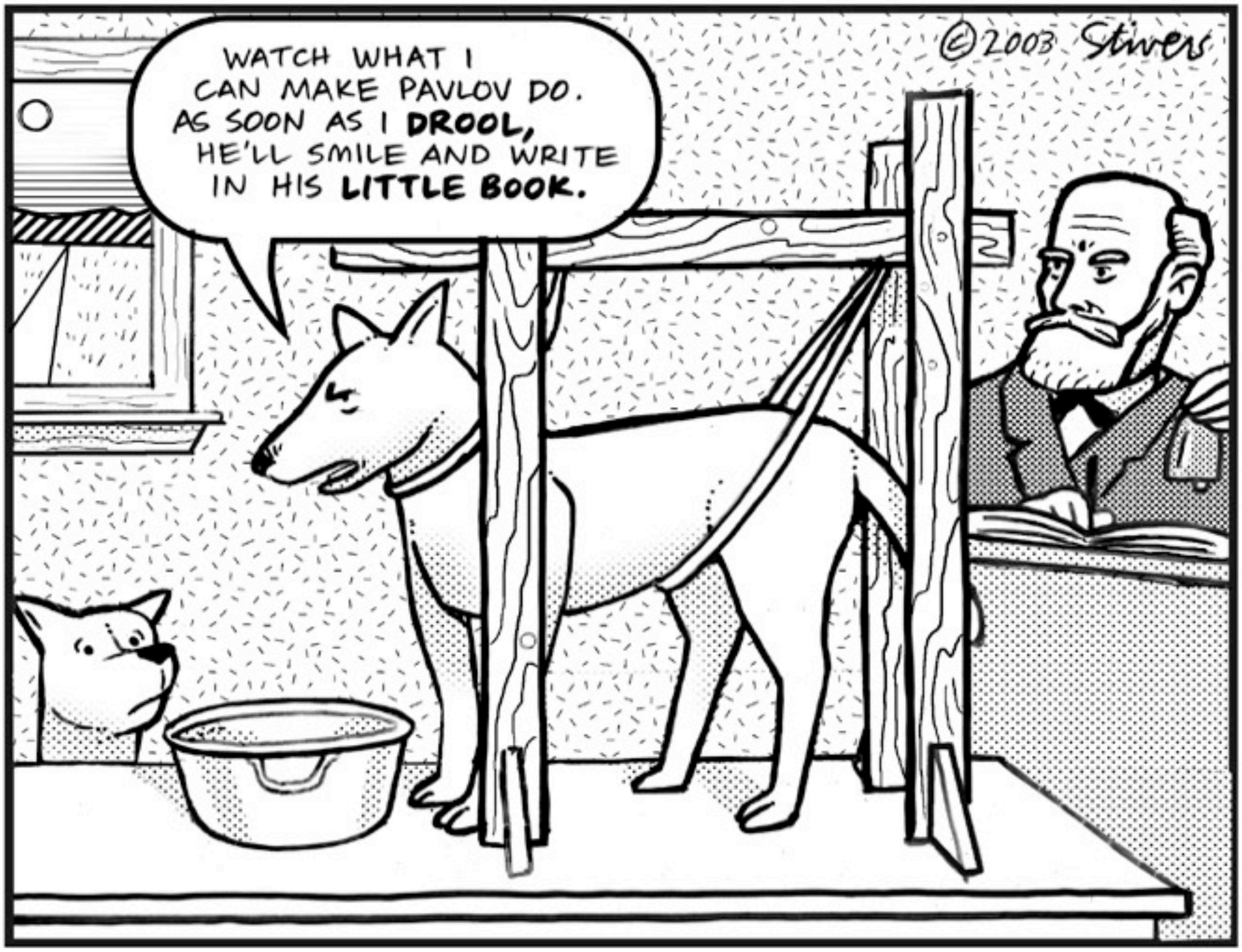
- US-CS substitution



University of Zurich

The other form of conditioning that is often discussed in the literature, is operant conditioning, where the agent has to perform a particular action such as pressing a lever. We will not discuss operant conditioning here.

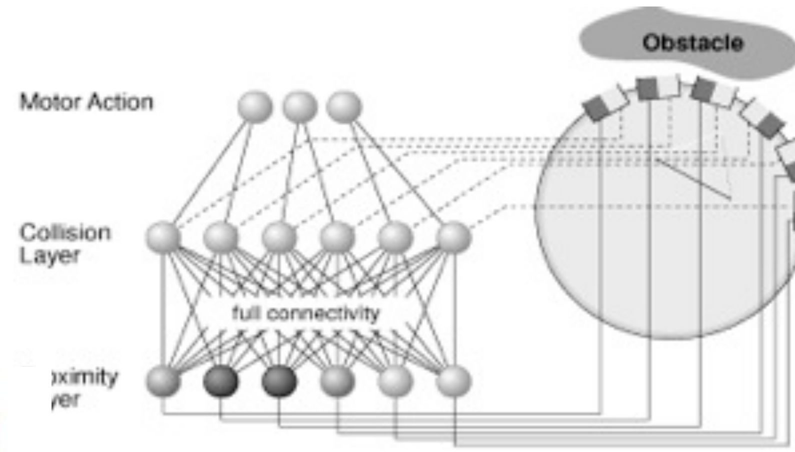
Pavlov's classical conditioning experiments



University of Zurich

DAC: Nodes characteristics, propagation and learning

$$h_i = c_i + \sum_{j=1}^{N} w_{ij} \cdot p_j$$



$$a_i = g(h_i) = \begin{cases} 0 & : h_i < \Theta \\ 1 & : h_i \geq \Theta \end{cases}$$

$$\Delta w_{ij} = \frac{1}{N}(\eta \cdot a_i \cdot p_j - \varepsilon \cdot \bar{a} \cdot w_{ij})$$

write down formulas for activation of collision layers: essentially the input is c plus the summed proximity layer activation
the nodes are “binary threshold”, we can take a threshold of 0.5; activation of collision node upon collision = 1.0
learning: Hebbian learning with learning rate eta and forgetting rate epsilon
formulas 5.7, 5.8, and 5.9, p. 161 of “Understanding Intelligence”
eta: learning rate, epsilon: forgetting rate; a: average activation in collision layer
Q: what is the forgetting good for?
Q: what happens over time?
Q: what if you have a light source along the wall?

Environment: Arena with obstacles

What happens?

Collision sensors: on/off, trigger pre-programmed reflexes

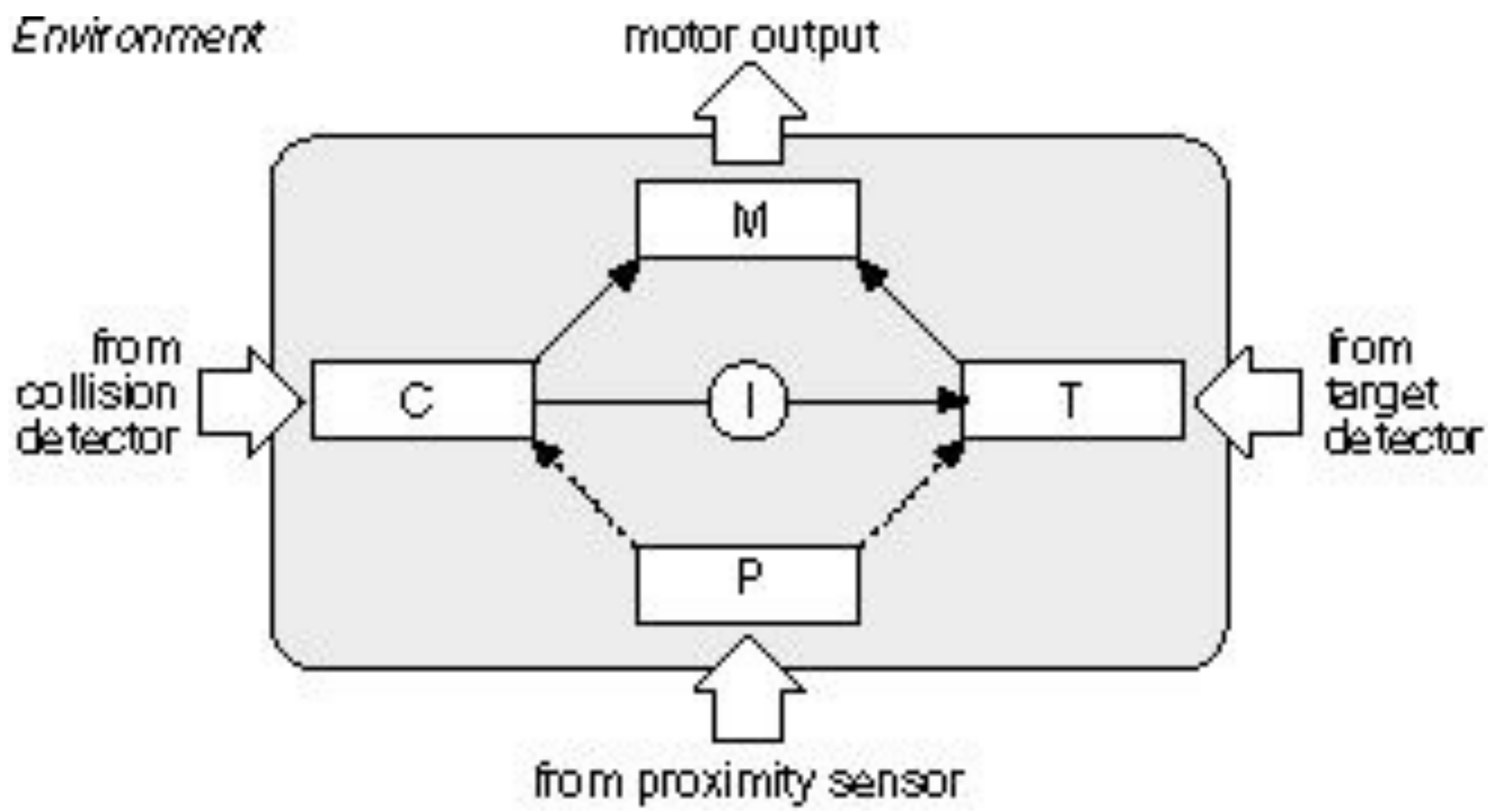
correlation with proximity sensors → partial overlap

extraction of mutual information → "anticipation" of collision

proximity sensors start taking over from collision sensors

Robot starts behaving as if it were anticipating collisions.

DAC: Including a light source



Environment: Arena with obstacles

What happens?

Touch sensors: on/off, trigger pre-programmed reflexes

partial overlap

extraction of mutual information

“anticipation” of collision

proximity sensors start taking over from collision sensors

Video/Demo “DAC”

http://shanghailectures.org/system/files/DAC_Video_small.m4v

Summary of DAC

- learning to “anticipate” obstacles (“hallucination”?): F-O-R
- collision, proximity sensors: different physical characteristics, partial overlap: redundancy principle
- induction of information structure through physical interaction with world: principle of sensory-motor coordination/ information self-structuring
- change of sensor morphology: correlations lost
- robust behavior: noise and fault tolerance, generalization

Summary of DAC

- continuous learning, impossible not to learn
- forgetting (active forgetting: only learn when something happens)
- extremely simple, but very powerful

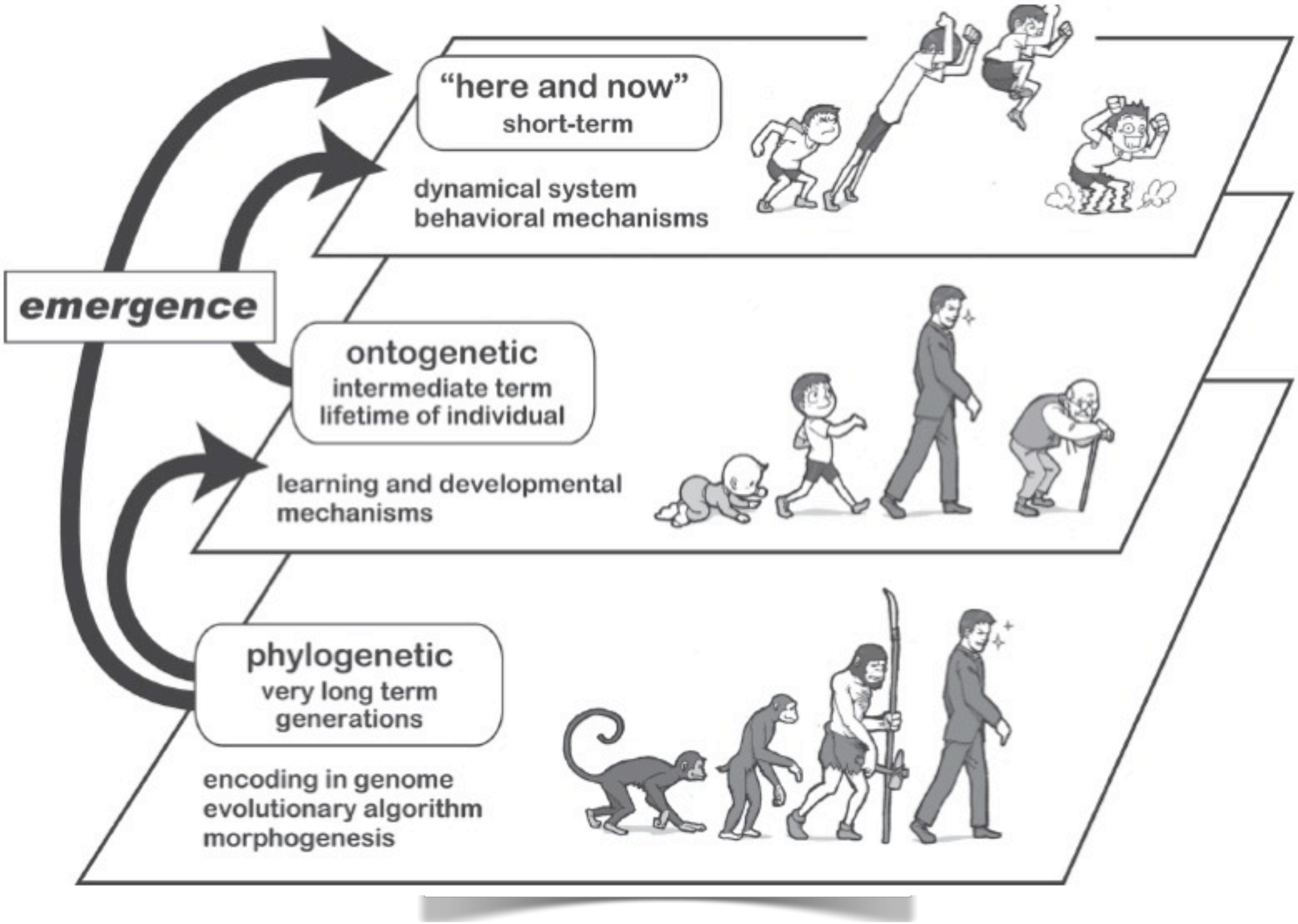
→ popularity in robotics, adaptive systems

Neural networks: Outlook

See also focus box 5.1 in "How the body …"

- more biologically plausible models, e.g. spiking neural networks, spike-time dependent plasticity
- neuro-modulator-based neural networks (chemical processes)
- growing neural networks
- multi-level simulations (including ion-channels), e.g. Henry Markram, EPFL, Switzerland

Time perspectives



Assignments for next week

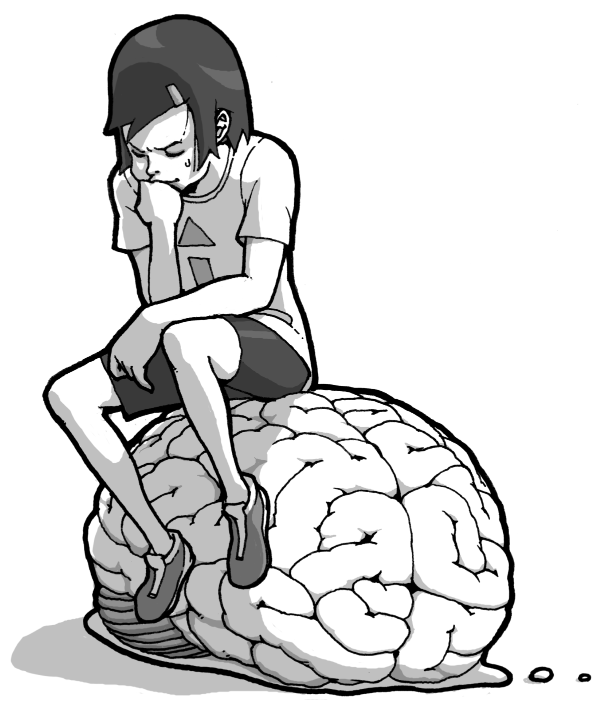
- Read chapter 5 of “How the body …

End of lecture 6

Thank you for your attention!

stay tuned for lecture 7

"Development: From locomotion/movement to cognition"



University of Zurich

lab